Department Chemie
der Technischen Universität München

Biochemical and NMR Characterization of Aggregating Proteins and their Interactions with Molecular Chaperones

Saravanakumar Narayanan

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende: Univ. Prof. Dr. Sevil Weinkauf

Prüfer der Dissertation: 1. Univ. Prof. Dr. Johannes Buchner

2. Univ. Prof. Dr. Bernd Reif

Humboldt Universität, Berlin.

3. Univ. Prof. Dr. Horst Kessler

Die Dissertation wurde am 17.09.2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 19.10.2004 angenommen.

கற்றதனா லாய பயனென்கொல் வாலறிவன்
நற்றாள் தொழாஅர் எனின் (திருக்குறள், 1:1:2, கி.மு.31)

What Profit have those derived from learning, who worship not the good feet of God
(Thirukkural, 1:1:2, B.C. 31)

To my family

கல்வி கரையில கற்பவர் நாள்சில
மெல்ல நினைக்கின் பிணிபல – தெள்ளிதன்
ஆராய்ந் தமைவுடைய கற்பவே நீரொழியப்
பாலுண் குருகின் தெரிந்து.

(நாலடியார்,2:14:5, கி.பி.300)

Learning has no bounds; the learners days are few. Would they calmly reflect, diseases are many. Let carefully investigate and make familiar with those things which are essential.

(Naladiyar, 2:14:5, a Tamil ethical literature, A.D 300)

Abstract

In this thesis, the mechanism of protein aggregation and the interactions of aggregating proteins with molecular chaperones are investigated using solution and solid-state NMR spectroscopy. The influence of chaperones on the protein aggregation is studied for the two systems, Sup35/Hsp104 and Aβ(1-40)/αB-crystallin. Protein aggregation itself is investigated using the model systems Aβ(1-40) and PI3-SH3.

The yeast prion system (Sup35/Hsp104) was investigated using various NMR (Diffusion, Saturation Transfer Difference, etc.) and biophysical techniques (Equilibrium dialysis, CD spectroscopy, etc.). We could show that Hsp104 specifically interacts with critical oligomers of Sup35 and that this interaction modulates the aggregation characteristics.

In the literature it has been shown that the interaction of the Alzheimer's disease causing amyloid peptide (Aβ(1-40)) with a small heat shock protein, αB-crystallin, increases the neurotoxicity in cultured neurons. We could identify the chemical groups of Aβ(1-40) that are involved in this interaction using STD NMR. As a consequence of the interaction, we found that Met35 in Aβ(1-40) becomes oxidized. The reason for the increased neurotoxicity is discussed based on these findings. Using biochemical, NMR and biophysical approaches (Isothermal titration calorimetry, CD spectroscopy and Colorimetric and a NMR redox assay that developed during this thesis), we could characterize specific redox properties of Aβ(1-40) and αB-crystallin. We could show that the interaction of αB-crystallin with Aβ(1-40) is copper modulated and redox driven. A simple concept for the implication with respect to neurotoxicity is presented. For the two chaperone systems, we observe a general trend for ligand chaperone interactions, which involves redox activity of chaperones. This is discussed at the end of the thesis.

The molecular mechanism of early steps in fibril formation is studied using the Aβ(1-40) peptide. Upon variation of the salt conditions, different oligomeric states of Aβ(1-40) could be generated that are analyzed by solution-state NMR experiments. We find that, in addition to specific peptide-peptide contacts, interactions of anions with selective amino acids of Aβ(1-40) are key for the aggregation process.

The SH3 domain of phosphatidylinositol 3'-kinase (PI3-SH3) forms amyloid fibrils when incubated at low pH for long periods of time. Early stages of fibril formation were studied using solution-state NMR, whereas the fibril structure was probed by solid-state NMR. By solid-state NMR, we find that the side chain of H25 is involved in fibril stabilizing interactions. By solution-state NMR, we show that modification of PI3-SH3 using Diethyl pyrocarbonate (DEPC) has an influence on the distribution of soluble and oligomeric PI3-SH3. Even though that the modified protein is more stable, aggregation is facilitated. Amino acids which specifically affected as a function of pH are identified using ^{1}H, ^{15}N correlation experiments. Structural models for protein aggregation are discussed in the context of domain swapping of SH3 domains.

Parts of this thesis have been or will be published in due course:

1. **Saravanakumar Narayanan**, Benjamin Bösl, Stefan Walter, Bernd Reif "Importance of low-molecular weight species for prion propagation in the yeast prion system Hsp104/Sup35" *Proc.Natl.Acad.Sci. USA*, (2003) **100** (16), 9286-9291.

2. Salvador Ventura, Jesus Zurdo, **Saravanakumar Narayanan**, Matildo parreno, Ramon Mangues, Bernd Reif, F.Xavier Aviles, Christopher Dobson, Luis Serrano "Sequence triggers for protein aggregation and amyloid formation- An SH3 case" *Proc.Natl.Acad.Sci. USA*, (2004) **101** (19), 7258-7263.

3. **Saravanakumar Narayanan,** Bernd Reif "Characterization of chemical exchange between soluble and aggregated states of β-Amyloid by solution state NMR upon variation of salt conditions" *Biochemistry, (2004) (In press)*

4. **Saravanakumar Narayanan,** Wilbert Boelens, Wilfried de Jong, Bernd Reif "Mechanistic understanding of the increased neurotoxicity of β-amyloid induced by the small heat shock protein αB crystallin" *Manuscript prepared for the Submission to Nature Structural and Molecular Biology.*

5. **Saravanakumar Narayanan,** Stefan Walter, Bernd Reif "Biochemical and structural investigations of the interactions between the molecular chaperone Hsp104 and the critical oligomers of yeast prion protein domain of Sup35p" Manuscript prepared *for the Submission to Biochemistry.*

6. **Saravanakumar Narayanan,** Salvador Ventura, Bernd Reif "Solution and solid-state NMR investigations of mechanism of fibril formation of Pi3-SH3" (in preparation)

Acknowledgements

First, I would like to express my deep gratitude to my Ph.D supervisor, **Prof. Dr. Bernd Reif.** His encouragement and guidance were the driving forces of my success. His scientific discussions and motivation were the essential ingredients for my research work. Under his shelter, I enjoyed moving around various places to learn new things and also to present my research work in conferences. I delighted my scientific freedom under his tutelage. His trust and concern on me fueled my interests. Bernd is an inspiring person who influenced me with his dedication, devotion and determination towards research work.

My sincere thanks to my head of the department, **Prof. Dr. Horst Kessler** who provided excellent working ambience and state-of-art NMR facilities. His suggestions shaped me a lot. He was very caring and helpful. I now remember, how to hold a laser-pointer without making circles around the target, during talks.

I am indebted to **Dr. Stefan Walter** who made my hands deft at protein chemistry work. I enjoyed his guidance as well. Working with him was a pleasant experience. I am obliged to **Prof. Dr. Johannes Buchner** for providing me excellent protein chemistry facilities. His appreciation of my work after my talk at the *Euresco* meeting was a memorable incident.

I thank **Prof. Dr.Michael Sattler** for the facilities at EMBL-Heidelberg. **Mr. Gunter Stier** is to be acknowledged, as he taught me molecular cloning and fast protein expression. He showed me how to move various projects at the same time efficiently.

I am grateful to **Prof. Dr. Hartmut Oschkinat** for the NMR-facilities at FMP-Berlin. I have to thank **Dr. Christian Freund** for allowing me to use his protein lab facilities during my stay at FMP-Berlin. **Dr. Peter Schmieder** at FMP - Berlin and **Dr. Gerd Gemmecker** at TU-Munich were helpful for providing me ample NMR measurement time for my research work.

Prof. Dr. Steffen Glaser was helpful for providing me MOCCA XY(16) and SIAM-TACSY pulse sequences for my research work and **Dr. Rainer Haeßner** for providing me the computer facilities and timely help. **Dr.Anette Diehl's** suggestion to my thesis was useful.

Acknowledgements

Index

Index

Index

Abbreviations

1D, 2D, 3D	One dimensional, Two dimensional, Three dimensional
Å	Angstrom
AAA	ATPases associated with various cellular activities
AD	Alzheimer's disease
ATP	adenosine tri-phosphate
APP	amyloid precursor protein
COSY	correlated spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CSA	chemical shift anisotropy
CT	constant time
CW	continuous wave
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DOSY	diffusion ordered spectroscopy
E. coli	*Escherichia coli*
EM	electron microscopy
EMBL	European Molecular biology laboratory
Eq.	equation
ERAD	endoplasmic reticulum associated degradation
ESI	electron spray ionization
ESR	electron spin resonance
FMP	Forschungsinstitut für Molekulare Pharmakologie
FT	Fourier transform
GSH	glutathione (reduced)
GSSG	glutathione (oxidized)
HP-SEC	high performance size exclusion chromatography
HSP	heat shock protein
HSQC	heteronuclear single quantum correlation
INEPT	insensitive nucleus enhancement by polarization transfer
ITC	isothermal titration calorimetry
kDa	kilo Dalton
MALDI	matrix-assisted laser desorption/ionization
MAS	magic angle spinning

Abbreviations

MBP	maltose binding protein
MDC	Max-Delbrück Centre for Molecular Medicine
MHz	megahertz
MOCCA	modified phase-cycled carr-purcell type
MREV	Mansfield-Rhim-Elleman-Vaughan
MWCO	molecular weight cut-off
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
PCR	polymerase chain reaction
PDB	protein data bank
PDI	protein disulfide isomerase
PI3	phosphotidyl-inositol 3
ppm	parts per million
RDC	residual dipolar couplings
r.f.	radio frequency
RP-HPLC	reversed phase HPLC
SAR	structure activity relationship
SDS-PAGE	sodium dodecyl sulphate- polyacrylamide gel electrophoresis
SH3	Src-homologue-3
SS-NMR	solid state NMR
STD	saturation transfer difference
ThT	Thioflavin-T
TOCSY	total correlation spectroscopy
TPPM	two-pulse phase modulation
TROSY	transverse relaxation optimized spectroscopy
TUM	Technical university of Munich
UV	ultraviolet
WATERGATE	water suppression by gradient- tailored excitation
YPD	yeast extract/peptone/dextrose

1.0 Introduction

1.1 Protein misfolding and disease

The "Central dogma of molecular biology" suggests that the flow of functional information travels from replicating DNA to transcribing RNA and to translating protein (*1*). According to Francis Crick, "Once information has passed into protein, it cannot get out again"(*2*).

DNA $\xrightarrow{\text{Transcription}}$ **RNA** $\xrightarrow{\text{Translation}}$ **Protein**

Figure 1. 1. The flow of genetic information as described by Crick in 1958

But, discovery of retroviruses opened a new door to the RNA world and this RNA to DNA cycle revised the "central dogma" (*2, 3*). Identification of proteins that carry disease-causing information, further revised the "central dogma" in such a way that a protein itself can elucidate biological activity without nucleic acid connections (*4*).

A human body may contain as many as 100,000 different types of proteins with varied structures to execute indispensable biological functions. The wide variety of highly specific structures that result from protein folding and that bring the key functional groups into close proximity has enabled living systems to develop astonishing diversity and selectivity in their underlying chemical processes.

The change from the regular arrangement of a rigid structure to the irregular, diffused arrangement of the flexible open chain causes the protein to denature, has been known for about 74 years (*5*). In-depth knowledge of protein folding came from the proposed "Thermodynamic hypothesis", which stated that the inter-atomic interactions and hence the amino acid sequences dictate the folding of the translated linear strand of amino acids into a fully functional three dimensional protein (*6*). Immediately, a mathematical model, "Levinthal's paradox", contradicted the existence of protein folding pathways that are governed by amino acids. It was based upon the fact that it would take an astronomical length of time to search for a correct energy minimized fold of a protein, if it undergoes random search (Fig.1.2a) (*7*). But, after enormous theoretical efforts combined with experimental data, the solution to the celebrated "Lavinthal's paradox" has emerged, considering biased search on energy landscapes (Folding funnels) on which the folding reaction occurs (Fig.1.2b) (*8*).

a. Levinthal's paradox   b.   The 'pathway' solution to the   c. idealized smooth funnel landscape
                               random search problem   for the fast folding pathway
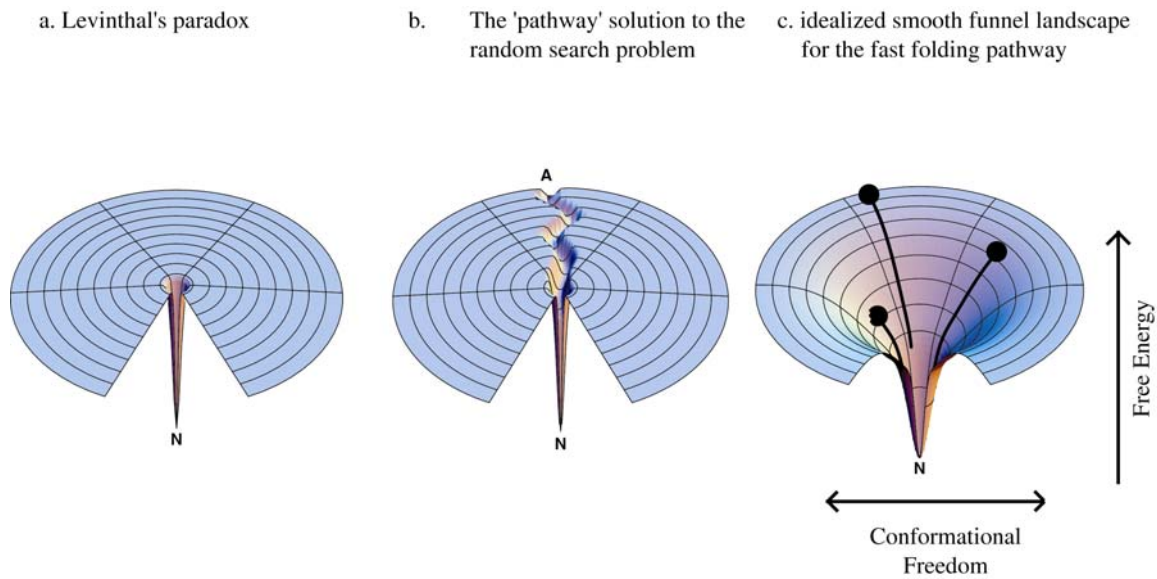
Figure 1.2. Pictures illustrate how a protein could undergo a fast-folding; a. The Levinthal 'golf-course' landscape. The native conformation is depicted as "N". The chain searches for Native conformation randomly; b. A pathway is assumed to lead from a denatured conformation A to the native conformation N, so conformational searching is more directed and folding is faster than for random searching; c.The chain forms increasing numbers of intrachain contacts, and lowers its internal free energy, its conformational freedom is also reduced. (K. A. Dill & H. S. Chan, Nature Struct. Biol. (1997), 4, 10)

According to the concept of a "funnel shaped energy landscape", the driving force for a protein to fold to its native state originates from a strong slope of the energy landscape toward native conformation (Fig.1.2c). However, this can be counteracted by roughness of the energy landscape that could render the folding reaction less effective, a phenomenon called "frustration", which eventually leads to kinetic traps and production of misfolded proteins (Fig. 1.3).

These kinetically trapped intermediates with only partially formed structures can have significant lifetime and hence associate and form ordered aggregates. That such complexities are seen even in the benign environment of dilute solution of a pure protein, e.g., hen lysozyme (*9*), suggests that they are even more likely to occur in crowded environment of the cell.
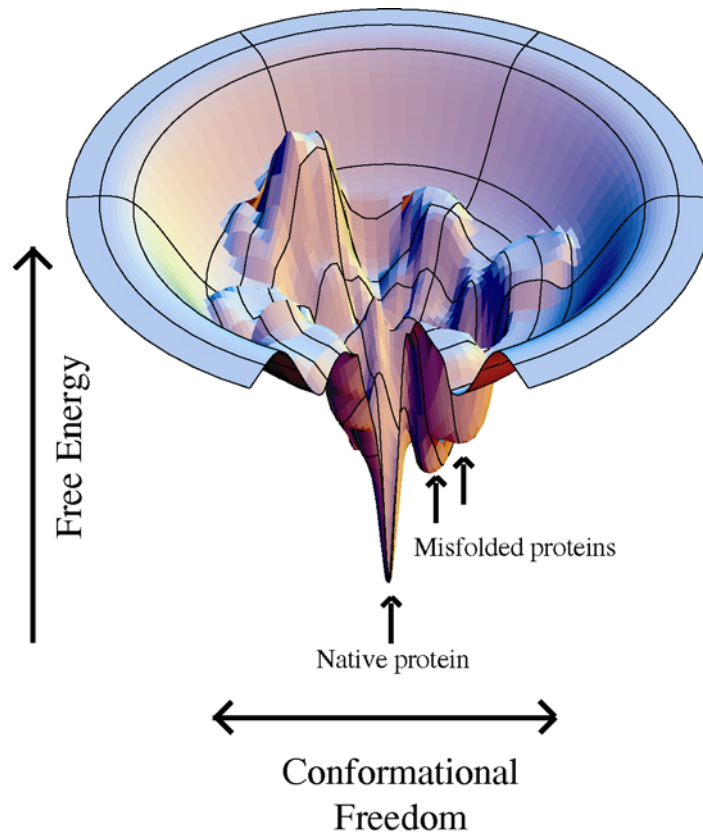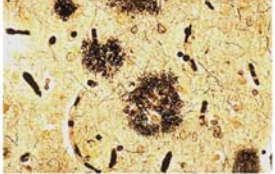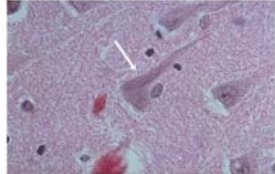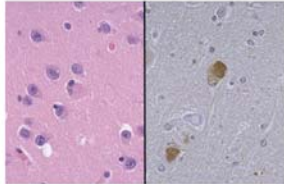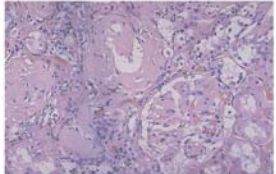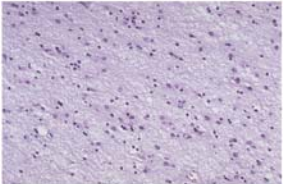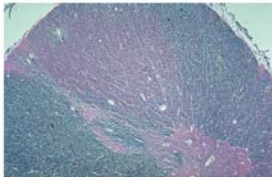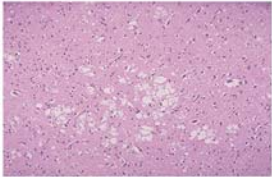
Figure 1.3. Denatured molecules at the top of the funnel might fold to the native state by a myriad of different routes, some molecules fold on different routes and may be kinetically trapped in a misfolded state.

Only the correctly folded proteins have long-term stability in biological milieu and are able to interact selectively with their natural partners. It is therefore not surprising that the failure of proteins to fold correctly (e.g., P53 protein), or their inability to remain in the soluble form (e.g., β-amyloid), is the origin of a wide variety of pathological conditions (*10*). Indeed, a huge variety of previously unrelated diseases share the feature of misfolded and aggregated protein deposits (Figure 1.4) . These disease-causing proteins have various three dimensional folds with nothing in common in terms of sequences, but instead converge in a unique fibrillar structure called "amyloidosis" (*11*).

Disease	Protein	Location	Pathology
Alzheimer's disease	Amyloid β-Protein	Neurons and glias	
Tauopathies	Tau	Neurons	
Parkinson's disease	α-Synuclein	Neurons	
Kidney amyloidosis	β-Amyloid	Renal	
Huntington's Disease	Huntingtin (polyglutamine)	Neurons/ glias	
Amyotrophic lateral sclerosis	Superoxide dismutase	Neuron & spinal cord	
Creutzfeldt-jakob /scarpie	Prion protein	Neurons and glias	

Figure 1.4. Some human diseases characterized by progressive misfolding and aggregation of proteins (Figures reproduced from http://medlib.med.utah.edu/WebPath/TUTORIAL/CNS/CNSDG.html (Internet pathology laboratory for medical education, Florida state university college of medicine, USA)

Amyloidosis have traditionally been defined as diseases in which normally soluble proteins accumulate in various tissues as insoluble deposits. They are made up of long un-branched filaments of approximately 10 nm diameter. The X-ray diffraction pattern indicates that the fibrils have extensive beta sheet character, and that these sheets run perpendicular to the fibril axis to generate what is described as a cross β-structure (*12*) (Fig. 1.5).
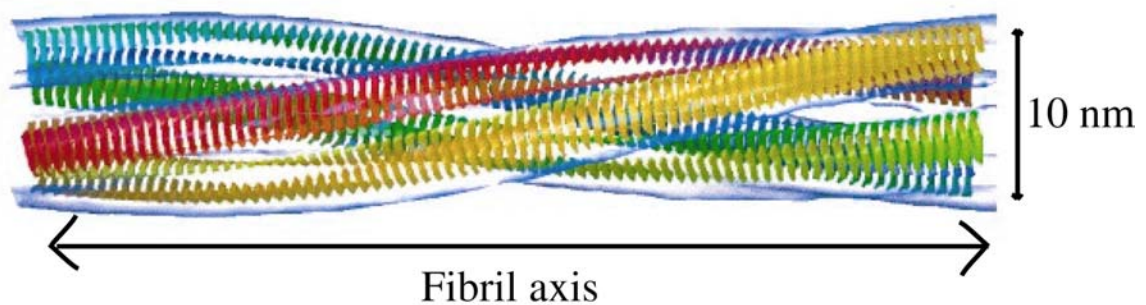
Figure 1.5. Computer generated molecular model of Insulin Fibrils (Courtesy: Birkbeck Crystallography amyloid research laboratory, U.K.)

There are striking similarities in fibril structure of different peptides and proteins. The first phase in amyloid formation seems to involve the formation of soluble oligomers. The earliest species visible by electron microscopy generally resemble small, bead-like structures. These early preformed aggregates then transform into species with more distinctive morphologies, often called protofibrils. These structures are commonly short, thin, sometimes curly, fibrillar species that are thought to assemble into mature fibrils, perhaps by association accompanied by some degree of structural re-organization (Fig. 1.6).
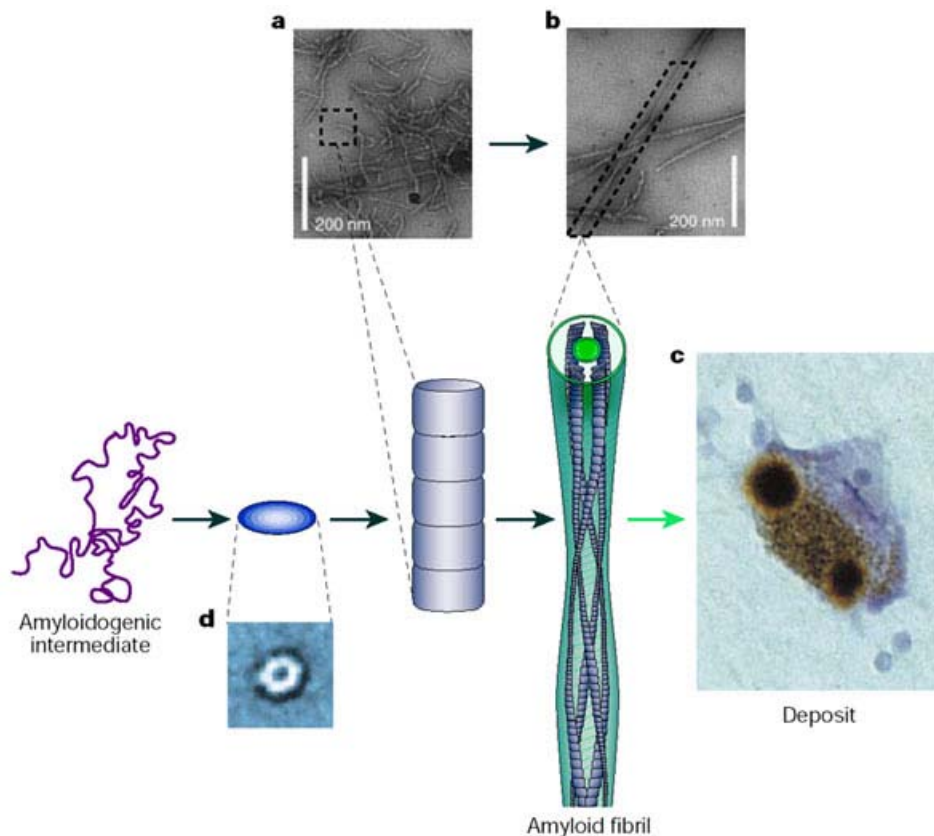
Figure 1.6. A schematic representation of the general mechanism of aggregation to form amyloid fibrils.

a) Misfolded proteins associate with each other to form small, soluble aggregates that undergo further assembly into protofibrils or proto-filaments. b) Mature fibril formation occurs. c) The fibrils often accumulate in plaques or other deposits d) Some of the early aggregates seem to be amorphous in nature, although others form ring shaped species with diameters of approximately 10 nm (Picture adapted from Dobson, C.M., Nature,(2003) , 426, 884)

1.2 Preventing protein misfolding and aggregation *in vivo*

In eukaryotic cells, the vast majority of proteins are synthesized and folded either by freely floating ribosomes in the cytosol or by ribosome-translocon complexes that are bound to the Endoplasmic recticulum (ER) (*13*). During translation, the folding information encoded in the amino acid sequence becomes available in a vectorial fashion, i.e., synthesis and translocation proceed from the amino-to the carboxy-terminus of the polypeptide. The polypeptide exit tunnel is largely formed by RNA with significant contributions from several ribosomal proteins. The length of the tunnel from the site of peptide synthesis to the exit in ribosomal subunit is about 100 Å (*14*). The tunnel is largely straight and the average diameter is 15 Å (Fig. 1.7).
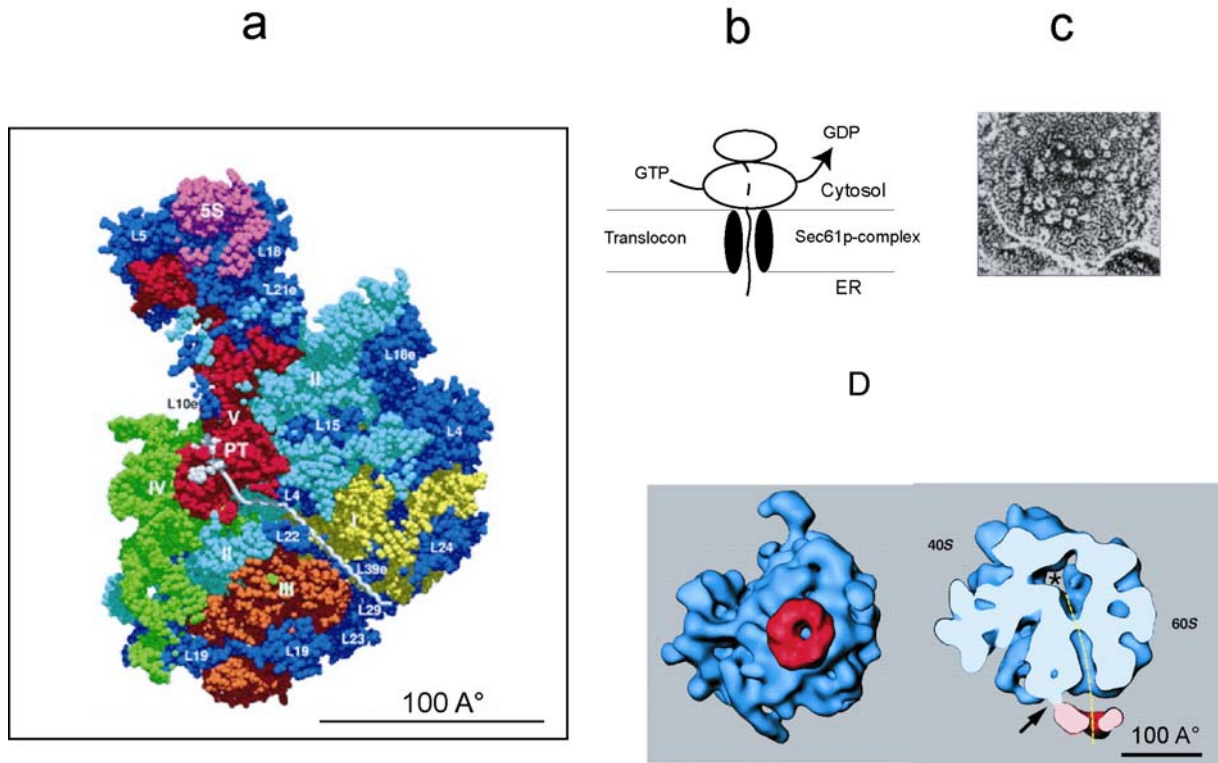
Figure 1.7. a) The polypeptide exit tunnel surface is shown with backbone atoms of the RNA color coded by domains. Domains I (yellow), II (light blue), III (orange), IV (green), V (light red), 5S (pink), and proteins are blue. A possible trajectory for a polypeptide passing through the tunnel is shown as a white ribbon. PT, peptidyl transferase site. (Picture adapted from Nissen, P., et al., Science, 2000, 289, 920) ; b) A cartoon representation of translocation of synthesized nascent polypeptide chain from ribosome-translocon complex into endoplasmic reticulum; c) Structure of the translocon complex in native mammalian ER membranes as viewed by freeze fracture electron microscopy (picture adapted from Kent, E.S., et al., Cell, 1998, 92, 381); d) Three-dimensional reconstruction pictures of the ribosome-translocon (Sec61) complex in a surface representation. The stem connecting ribosome and the translocon complex is indicated by arrow. (Picture adapted from Beckmann, R., et al., Science, 1997, 278, 2123)

The tunnel surface is largely hydrophilic and includes exposed hydrogen bonding groups from the bases, backbone phosphates and polar protein side chains (Fig. 1.7). There are no patches of hydrophobic surface in the tunnel which is large enough to form a significant binding site for hydrophobic sequences in the nascent polypeptide. Also, the 15 Å wide tunnel is expected to prohibit protein folding beyond helix formation inside the ribosome. The 100 Å (~130 Å in case of ribosome-translocon complex) length of the channel can accommodate 35 amino acid residues or an α-helix of 65 residues (*14*). Since the formation of stable tertiary structure is a cooperative process at the level of protein domains (50 to 300 amino acids), an average domain can complete a fold only when its entire sequence has emerged from the ribosome. It takes more than a minute to synthesize a 300-residue protein in eukaryotes (*15*). As a consequence, many nascent chains expose non-native features for a considerable length

7

of time and are prone to aggregation. This tendency to aggregate is thought to be greatly increased by the close proximity of nascent chains of the same type in polyribosome complexes. These combined effect of nascent chain formation and molecular crowding necessitate a stringent quality control machinery within the cell compartment.
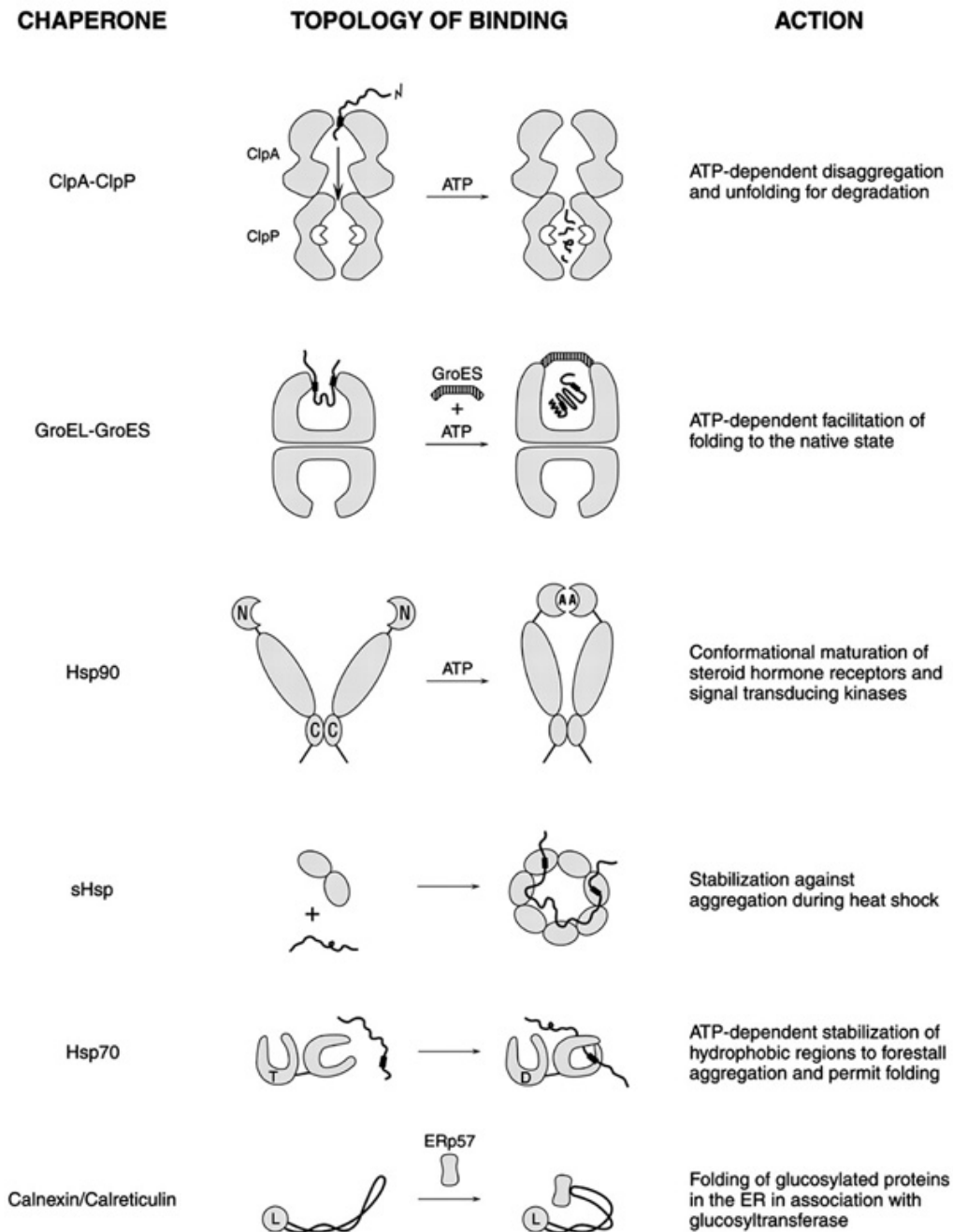
Figure 1.8. Different classes of chaperones and their actions . (picture adapted from Horwich, A.L., et al., EMBO Reports, 2001, 2, 1068)

The group of proteins that assist in the correct folding or assembly of other proteins in vivo, are termed "Molecular chaperones". Many of them are also designated as "Heat shock proteins", as they are normally overexpressed when the cells are shifted to high temperature or heat shock response. According to their molecular weight, molecular chaperones are divided into several classes or families (*16*). Each of them has its own mode of action and own topology of binding. The classification is depicted in the Fig 1.8. Chaperones usually form homo-or hetero-oligomers. Together with a different class of proteases, the proteasomes, they maintain the quality control machinery. Generally, the quality control machinery consists of combination of a several distinct processes: recognition, refolding, retention, release and degradation. Although general principles of quality control are similar in both cytosol and ER, fundamental differences exist. The common eukaryotic cytosolic and ER chaperones are listed in the table 1.1 (*17*).

Table 1.1   Classification of Chaperones

Organelle	Target polypeptide	Chaperone	Role
ER	Nascent secretary protein	Bip (GRP78)	Completion of translocation and folding
	Unfolded proteins	GRP72, GRP94	Completion of folding
	Unfolded glycoprotein	Calreticulin, Calnexin	Completion of folding (lumen & membrane)
	MHC class 1 heavy chains	P88	Stabilization
	Proline containing proteins	Cis/trans prolyl isomerase	Catalysis of slow protein folding reactions
	Proteins with disulfides	PDI	Formation and rearrangement
Cytosol	Steroid receptor	Hsp70, Hsp90	Stabilize inactive receptor
	Unfolded proteins	Hsp40	Cooperate with Hsp70 to stabilize non-native forms
			Binds to intermediates and promote folding
	Unfolded/misfolded proteins	TRiC	

Most of the cytosolic and ER chaperones act by holding nascent and newly synthesized chains in a state competent for folding upon release into the medium. Alternatively, large cylindrical chaperonin complexes provide physically defined compartments for an unfolded protein or protein domain to fold while being sequestered into cytosol or into ER compartments. In addition to the above mechanism, the ER also contains chaperones and folding enzymes that

are unique, such as calnexin and calreticulin and the family of thiol-disulfide oxidoreductases. In eukaryotic cytosol and ER, the protease complex responsible for most of the protein degradations is the 26S proteasomes. But, in ER, there also exist an additional degradation pathway for the glycosylated proteins termed " endoplasmic reticulum associated degradation (ERAD) pathway".
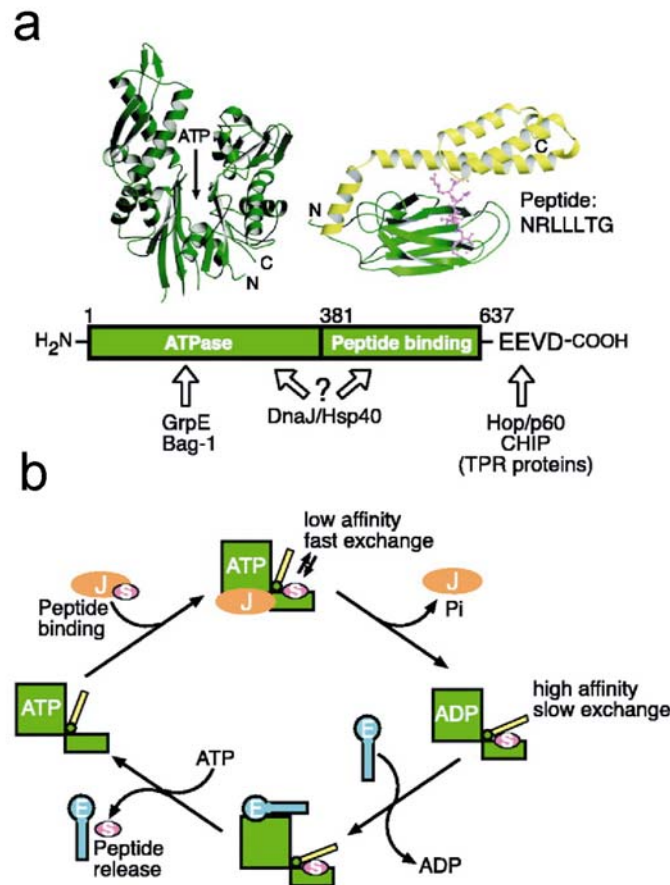
Figure 1.9. a) (Top) Structures of the ATPase domain and the peptide-binding domain of Hsp70 shown representatively for E. coli DnaK. The alpha helical latch of the peptide binding domain is shown in yellow and a ball-and-stick model of the extended peptide substrate in pink. ATP indicates the position of the nucleotide binding site. The amino acid sequence of the peptide is indicated in single-letter code (D, Asp; E, Glu; G, Gly; L, Leu; N, Asn; R, Arg; T, Thr; and V, Val); a) (Bottom) The interaction of prokaryotic and eukaryotic cofactors with Hsp70 is shown schematically. Residue numbers refer to human Hsp70. Only the Hsp70 proteins of the eukaryotic cytosol have the COOH-terminal sequence EEVD that is involved in binding of tetratricopeptide repeat (TPR) cofactors; b) Simplified reaction cycle of the DnaK system with DnaK colored as in (a). J, DnaJ; E, GrpE; S, substrate peptide. GrpE is drawn to reflect the extended shape of the protein. Not all substrates are presented to DnaK by DnaJ. The intermediate DnaK-DnaJ-substrate-ATP is probably very transient, as this is the fastest step of the cycle . (Picture adapted from Hartl, F.U .and Hayer-Hartl, M., Science, (2002),295,1852)

The general quality control mechanism in both cytosol and ER initially start with recognition of a stretch of hydrophobic amino-acid side chains. In the case of eubacterial Hsp70 (DnaK in eubacteria / BiP in eukaryotes), together with co-chaperone of the Hsp40 (DnaJ in eubacteria) and nucleotide exchange factor, GrpE, specifically recognize a ~7 residues long hydrophobic core region with leucine and isoleucine residues. Exposed hydrophobic amino acid side chains in conjugation with an accessible polypeptide backbone are the common structural features in DnaK recognition. The binding sites occur statistically every ~40 residues in proteins and are recognized with affinities of 5 nm to 5 μM (*15*). Rapid peptide binding occurs in the ATP bound state of DnaK and stable holding is achieved by hydrolysis of bound ATP. The NH_2- terminal J domain of DnaJ binds to DnaK and accelerates hydrolysis of ATP. The initial binding of DnaK to the substrate and subsequent conformational change of DnaK upon hydrolysis of ATP are the major driving factors for Hsp70 action (Fig. 1.9, (*15*)). Upon rebinding of ATP, the DnaK-peptide complex dissociates, completing the cycle.

A similar ATP driven mechanism also has been proposed for the other class of cytosolic molecular chaperones, termed as "chaperonins". The chaperonins are a conserved class of large double ring complexes of ~800 kDa, enclosing a central cavity (Fig. 1.10).
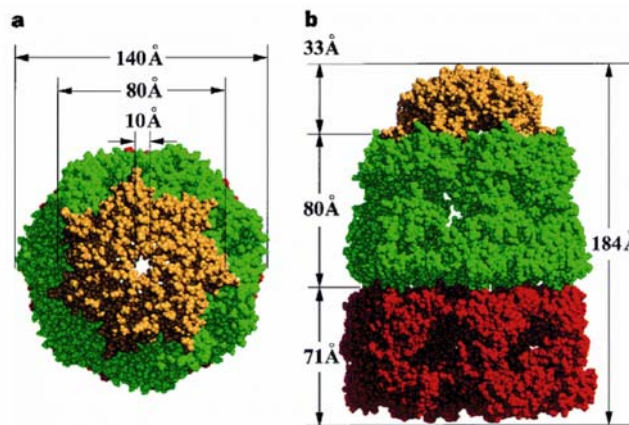
Figure 1.10. Overall architecture and dimensions of the GroEL–GroES complex; a, Van der Waals space-filling model of the entire complex in a top view looking down from the GroES-binding (cis) side; b, Side view. The complex is colour coded as follows: trans GroEL ring, red; cis GroEL ring, green; GroES, gold.(picture adapted from Xu, Z., et al., Nature, 1997, 388, 741)

The chaperonin GroEL generally cooperate with cofactors of the GroES, Hsp10 family of proteins, and is ATP-dependent. Substrate binding of the GroEL-GroES complex ("Anfinsen cage") is associated with rigid body movements of the peptide-binding apical domains of the

bound GroEL ring, elevating and twisting their hydrophobic surfaces away from the central cavity and replacing the lining of the cavity with hydrophilic surfaces. These movements result in rapid release of the substrate protein from the cavity walls into an encapsulated chamber, where it has no access to other monomers with which it could aggregate. The walls of the cavity that are hydrophilic in character, may encourage productive folding. After the interactions with GroEL, the substrate's hydrophobic surfaces are buried and its hydrophilic surfaces are exposed, as in native proteins. This folding-active ternary complex has a finite lifetime (12 seconds at 23°C) before the hydrolysis of ATP in the proximal ring and the subsequent binding of ATP to the opposite ring cause it to dissociate, releasing both GroES and the polypeptide substrate (*18*).

ATP binds with positive cooperativity within one GroEL ring, but between rings with negative cooperativity, so that ATP effectively occupies only one ring at a time. This establishes the asymmetry of the system, which is reinforced by the nucleotide requirement for GroES binding. Non-native polypeptide binds to the open ring of an asymmetric complex (first panel, Fig 1.11), and GroES binding to the same ring forms a folding-active complex and triggers polypeptide folding (second panel). The lifetime of this complex is determined by ATP hydrolysis in the GroES-bound ring, which weakens the interaction between GroES and GroEL (third panel). Binding of ATP and polypeptide to the opposite ring (fourth panel) then discharges GroES and the polypeptide, either in the native state (N) or one committed to it (I_c) or in a still non-native state (I_{uc}) that can rebind to GroEL and try again to refold (fifth panel). Binding to an open GroEL ring may be associated with an unfolding action. Note that the rings oscillate back and forth as polypeptide-accepting and then folding-active, a function of the asymmetric binding of ATP/GroES. cis, binding of GroES and polypeptide to the same GroEL ring (Fig. 1.11).
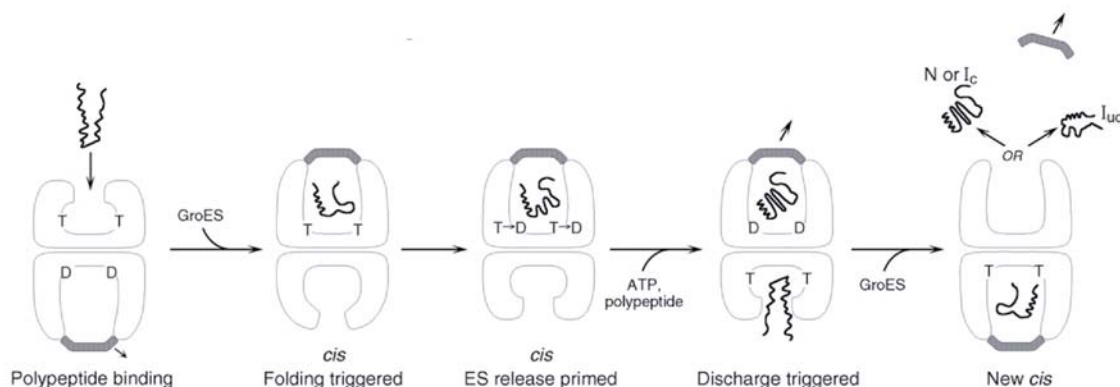
Figure 1.11. GroEL-GroES mechanism (ATP is designated as T and ADP as D)

Apart from the above ATP dependent chaperone assisted refolding mechanisms, ER uses an alternate retention mechanism termed as calnexin / calreticulin cycle in order to assist the folding of glycoproteins. This unique mechanism involves formation of a calnexin and calreticulin complexes together with thiol-disulfide oxidoreductase (eRP57) (Fig. 1.12). Two functionally independent ER enzymes mediate the on-off cycle in this chaperone system. Glucosidase II is responsible for dislocating the substrate glycoprotein from calnexin or calreticulin by hydrolysing the glucose from the monoglucosylated core glycan. UDP-glucose: glycoprotein glucosyltransferate (GT), on the other hand, is responsible for reglycosylating the substrates so that it can reassociate with calnexin or calreticulin. Reglucosylation by GT happens only if the glucoprotein is incompletely folded. The folded protein can only exit the cycle when GT fails to reglucosylate it. The glucose acts as a selective tag for incompletely folded protein- it is, in a sense, "a stamp of disapproval" that tells the system that the protein is not yet ready to be deployed. The cycle continues till the protein attains folded structure or gets degraded by 26S proteasomes before transported via translocon complex. The folded proteins are transported out of the ER to the golgi apparatus.
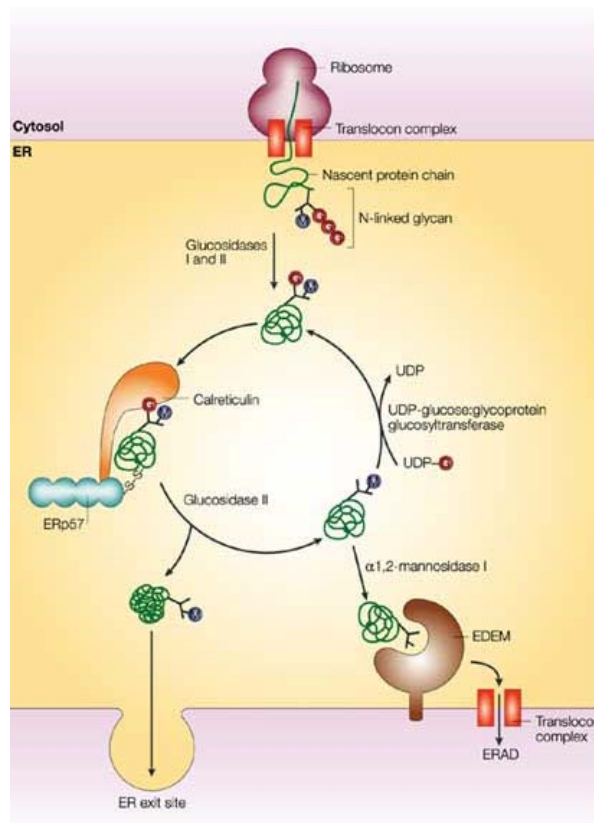
Figure 1.12. Calnexin /Calreticulin cycle

The 26S proteasome complex consist of a variety of different subunits with varied proteolytic activity. The complex is ubiquitous in eukaryotic cells, working as natural machinery for degradation of proteins. Regarding the structure as well as the function of the 26S proteasome, the macromolecule can be subdivided into two parts. Four ring-like structures form the 20S proteasome, a cylindrically-shaped molecule with a length of 15 nm and a diameter of 11 nm, enclosing three cavities. The central cavity contains the active sites of the proteasome, where peptide chains are cleaved into pieces. Attached to the 20S core particle are one or two so-called 19S complexes, resulting in a length of 30 nm or 44 nm for the entire particle. It is assumed, that it contains the substrate binding site, recognizing proteins subjected to degradation. Additionally, a protein has to be unfolded prior to cleavage, because only the unfolded peptide chain can enter the reaction compartment of the 20S core(Fig. 1.13).
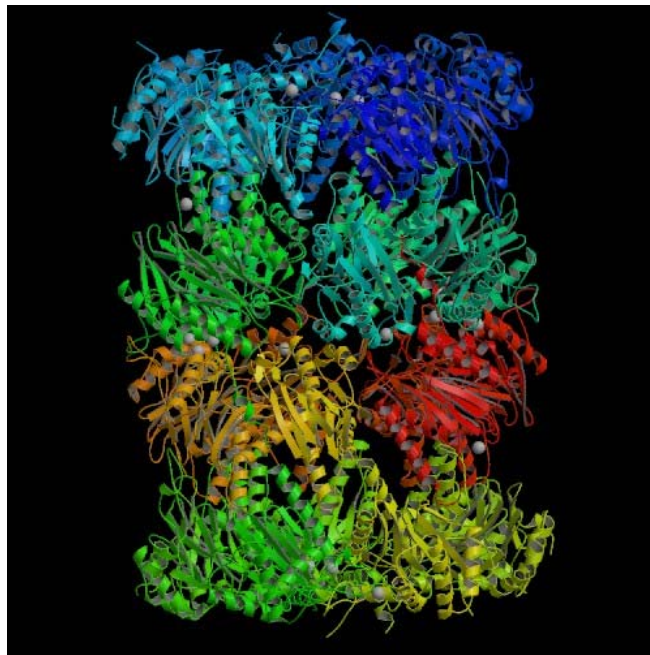
Figure 1.13. Structure of 20S proteasome (PDB:1RYP)

Apart from the degradation in the lysosome, most misfolded peptide escape either from cytosolic chaperones or from ER chaperones and those undergo calnexin/calreticulin cycle, are subjected to undergo a stringent degradation procedure in cytosol , termed as "ubiquitin-proteasome pathway". The misfolded proteins are poly-ubiquitinated by ubiqutin activating and conjugating enzymes (E1,E2,E3) eventually recognized and subsequently degraded by 26S proteasome complex (Fig. 1.14).
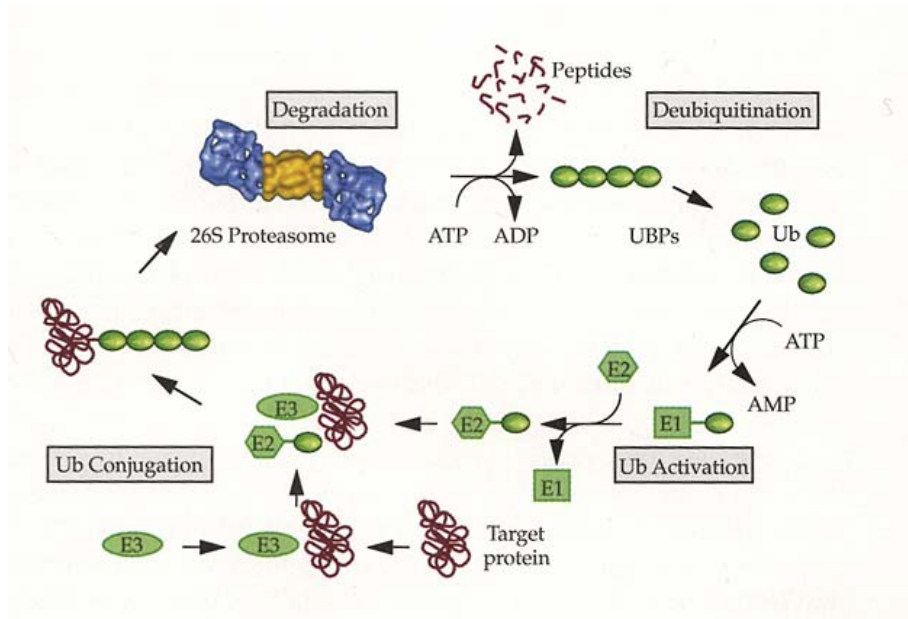
Figure1.14. Ubiquitin-Proteasome pathway

1.3   Models of fibril formation

Understanding the mechanism of fibril formation should prove useful in understanding the mechanism of several diseases. Currently, there are four popular mechanistic models describing amyloid fibril formation. The first is referred to as templated assembly (TA), wherein a pre-assembled nucleus binds to a soluble state (S) peptide in a random coil conformation in a rapid pre-equilibrium step (*19*). This is followed by a rate-determining structural change to add the peptide to the growing end of the fibril or filament, presumably as part of the β-sheet-rich quaternary structure (Fig. 1.15a). The second model is the monomer-directed conversion (MDC) model, which implies that a monomeric peptide can adopt a conformation called the A state that is analogous to the conformation adopted in the fibril. In the rate-determining step, this structured monomer (A) then binds and converts an S state monomer, resulting in an A state dimer (Fig. 1.15b). The dimer then dissociates and the structured monomers rapidly add to the end of the growing fibril (*20*). The third model, the often invoked nucleated polymerization (NP) mechanism (Fig. 1.15c), is characterized by the rate-limiting formation of a nucleus resulting from an equilibrium between monomers that are and are not assembly competent. Once a nucleus is formed, assembly occurs by the addition of assembly competent monomers to the growing end of the fibril (*21*).
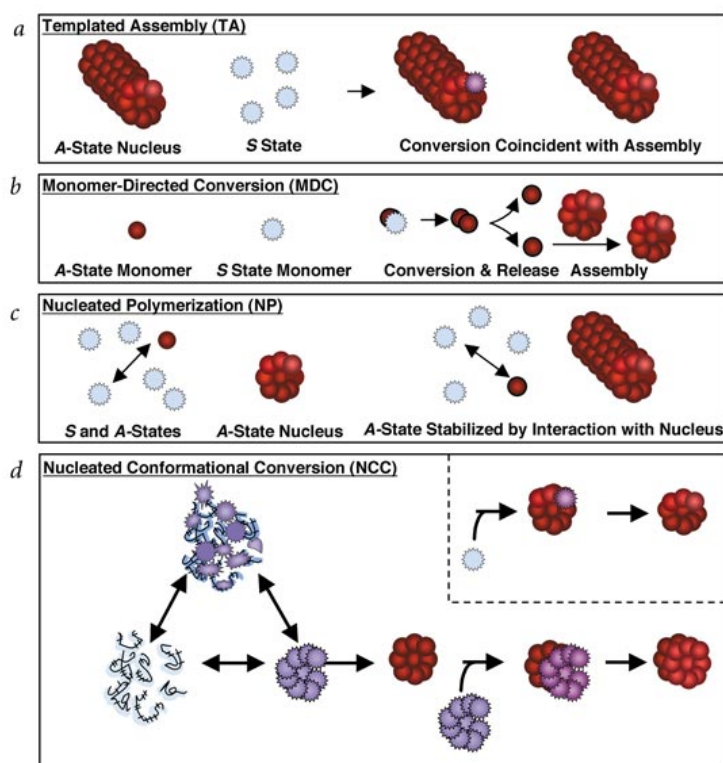
Figure 1.15. Existing models for conversion of amyloidogenic polypeptides into fibrils.

A recent model, nucleated conformational conversion (NCC) proposed that nuclei form by conformational rearrangements within structurally dynamic oligomers lacking a defined quaternary structure (Fig. 1.15d). Once nuclei are formed, they interact with structurally flexible oligomers with a distribution of subunits, adding a group of subunits to the fibril end simultaneously (*22*).

1.3.1   Theory of fibrillogenesis

Protein aggregation occurs due to the differential rate of clearance pathway of the amyloidogenic proteins and their production within the cell compartments. In-depth analysis of physics of fibril formation in relation to the concentration of proteins would eventually unfold the biology of fibril formation and will also be useful in developing techniques to identify them. Fibrillization of many proteins is controlled by two kinetic parameters: the nucleation rate (K_n) and the elongation rate (K_e). A kinetic theory of fibrillogenesis was established which relates the concentration of the amyloidogenic protein and the rate of fibrillogenesis (*23*). Fibrillogenesis differs depending upon whether the total protein concentration, C, is above or below a certain threshold critical "micelle" concentration (cmc), c*. These micelles may serve as nuclei for the fibril formation. For C > c*, the rate of elongation is independent of C, as a reversible equilibrium is established between monomer

and micelles. On the other hand, for C < c*, the initial rate of elongation is proportional to C (Fig. 1.16). To determine c*, the crossover between the domains where the initial growth rate is proportional to the initial concentration and where it is independent of the initial concentration has to be determined.
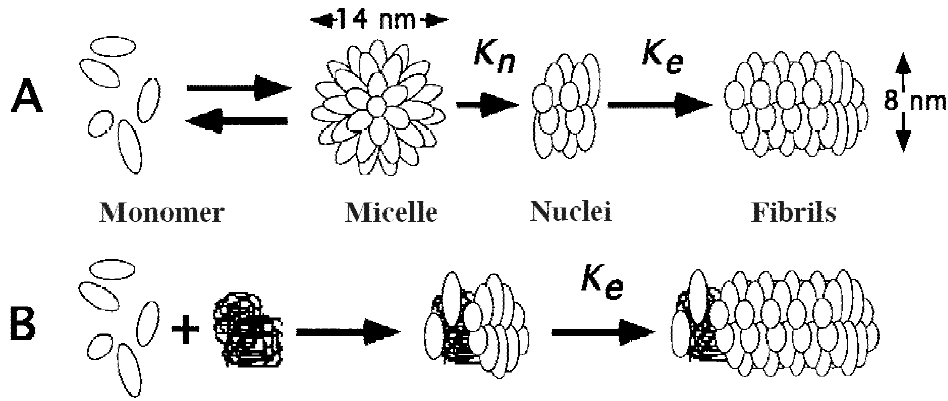
Figure 1. 16. A) Homogeneous nucleation and growth of fibrils for C > c ; B) Heterogeneous nucleation and growth of fibrils for C < c* (seeded growth)*

Above the critical micelle concentration, the concentration of fibrils as a function of time ($\dot{Np}$) can be expressed as a combination of nucleation (K_n) and elongation rates (K_e) and concentration c, of free monomers. The nucleation rate is the average number of nuclei produced by a single micelle per unit time and the elongation constant is the coefficient of proportionality between the number of monomers attached per unit time to each fibril.

$$\dot{Np} = cK_e N_{p-1} - cK_e N_p + K_n M \delta_{pn0} \quad [1]$$

The first term on the right hand side of equation 1 describes the creation of the fibrils of size p from those of size p-1 by monomer binding. The rate of such binding is proportional to the concentration of monomers, c. The second term accounts for the reduction in p-mer concentration due to the conversion of (p+1)-mers. The third term describes the generation of the n_0-mer nuclei from micelles. Monomer binding is irreversible and the breaking or merging of fibrils is neglected in above equation. Two state variables c and M are related according to the principles of thermodynamic equilibrium between the monomers and the micelles, assuming two-state model of micelle formation.

Here, all micelle have same aggregation number (m_0) (*23*), and it can be shown that

$$m_0 M = c^* \left(\frac{c}{c^*} \right)^{m_0} \quad [2]$$

The probability of finding the species Np can be related to the Gibb's free energy, ΔG,

$$[Np] = [Np_0] \exp[-\Delta G(p)/RT] \qquad [3]$$

At equilibrium, [Np] = [Np-1] and the critical micellar concentration is connected to the slope of the free energy, $d\Delta G/dp$, which is related to the ratio of the rate constants into and out of the state for a given monomer concentration.

$$\frac{[Np]}{[Np_{-1}]} = \exp\left\{-\left[\frac{\Delta G(p)-\Delta G(p-1)}{p-(p-1)}\right]\Big/RT\right\} = \exp\left(-\left\{\frac{d\Delta G}{dp}\right\}\Big/RT\right) \qquad [4]$$

The changeover in rates is therefore related to the change in slope of the free energy barrier, and a barrier that is linear with size gives a constant rate ratio (*24*) (Fig. 1.17).
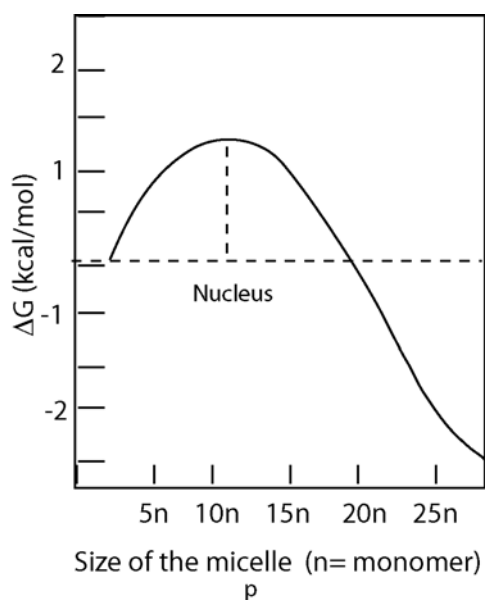
Figure 1.17. Free energy of the aggregate DG is shown on the vertical axis, where as the size of the aggregate is shown on the horizontal axis. Polymerization requires that the aggregate size pass through the maximum free energy

When the turning point is sufficiently sharp, the implication is that there is one state with a particularly small population that will represent the rate-limiting step for the reaction. The bottleneck is a thermodynamic nucleus, a necessary species but very scarce in the reaction path, which is least stable. When the small concentration of nuclei effectively form a barrier to further growth, then the rate of the formation of polymers is set by the population of nuclei and the rate of elongation of the nuclei themselves.

1.4 *in vitro* **methods to monitor protein association and protein-protein interactions**

Exact site-specific molecular mechanism and kinetics of amyloidogensis can be studied effectively through *in vitro* methods only, though laborious mutagenesis study would also help to explore the interactions. Several techniques are useful to follow the oligomerization and fibril formation. Certain analytical and biochemical methods are useful to monitor oligomerization and others are effective in monitoring protein-protein interactions. Here are some characteristic properties to monitor the substrate-substrate or protein-substrate interactions.

1.4.1 Monitoring aggregation

a) Dye Binding
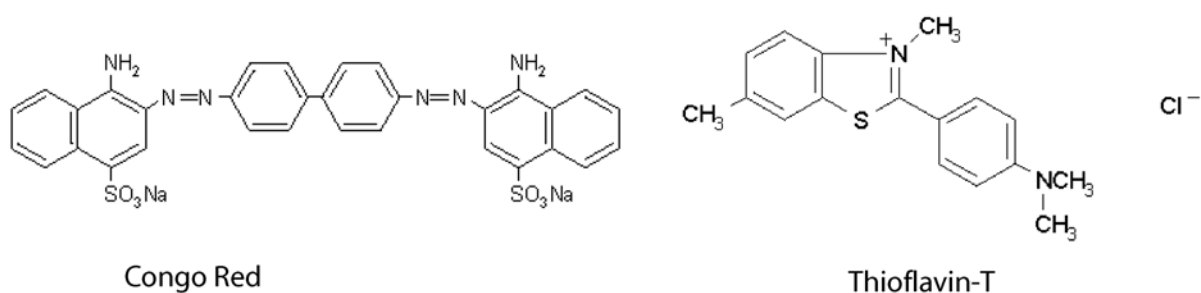
Figure 1.18. Commonly used amyloid staining dyes

Investigation of amyloid fibril formation requires not only the ability to distinguish the characteristic ordered extended β-sheet structure from the amorphous aggregates, but quantification of the amyloid form as well. Congo red (CR) and Thioflavin-T (ThT) are commonly used to stain amyloid fibrils both *in vitro* and in brain pathology (Fig. 1.18). Upon binding, characteristic spectral alterations occur for a variety of amyloid fibrils that do not occur on binding to the precursor monomers or amorphous aggregates. Upon binding to amyloid fibrils, a spectral shift of 100 nm for excitation (λ_{ex} = 342 nm to 442 nm) and 50 nm for emission (λ_{em} = 430 nm to 482 nm) occurs in the case of ThT (*25*). This large fluorescence excitation spectral shift allows selective excitation of amyloid fibril-bound Thioflavin-T and makes the quantification possible (Fig. 1.19).
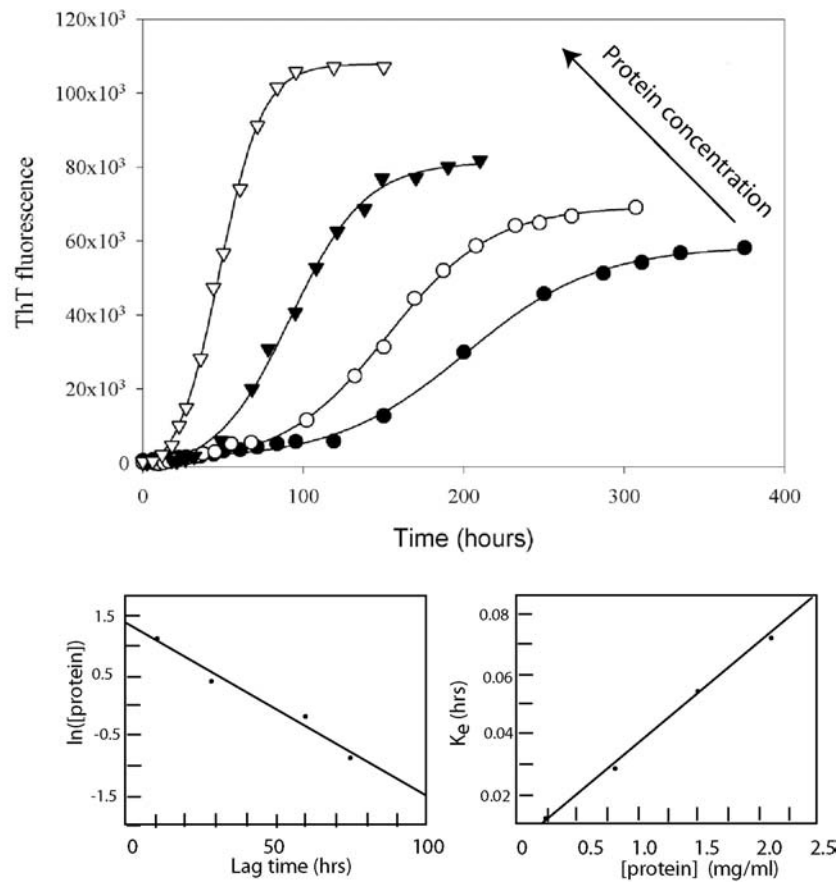
Figure 1.19. The effect of protein concentration on the kinetics of fibril formation as monitored by Thioflavin T fluorescence

At a given protein concentration, the changes in ThT fluorescence intensity follow a characteristic sigmoidal curve, *i.e.* an initial lag phase (nucleation), a subsequent exponential growth phase (elongation), and a final equilibrium phase (saturation) are observed (Fig. 1.19, top). From a concentration-dependent experiments, an inverse linear relation can be obtained between the lag time vs concentration of the amyloidogenic protein and there is also a linear dependence of the first-order rate constant for fibril growth (elongation) on the protein concentration (Fig. 1.19 bottom).

b) Hydrodynamic radii

As the protein aggregates, the mean square radius of the protein increases, assuming the protein behaves like a hard sphere (*26*).

$$R_G^2 = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}^2$$

where R_{ij} is the distance between the centers of spheres i and j for the aggregates containing n spheres. The value R_G is termed as radius of gyration, it relates to the hydrodynamic radius by a dimension less factor ρ, i.e. $R_G = R_H \cdot \rho$. For random coil, the value ρ ranges from 1.5-1.8 and for the hard sphere, it is measured as 0.778. The hydrodynamic radius includes effects arising from both shape and hydration. Several analytical techniques are based on hydrodynamic radii measurements. Analytical techniques that measure sedimentation or centrifugation coefficients (ultracentrifugation), molecular weight (Size exclusion chromatography), aggregate scattering factor (Quasi or dynamic light scattering), small angle X-ray scattering (SAXS), affinity capillary electrophoresis (ACE), diffusion coefficient (DOSY-NMR), polarization and anisotropy (fluorescence depolarization techniques) and photo bleaching-recovery all depend on the hydrodynamic radius of proteins. The molecular weight of the oligomers can be deduced by the following equations.

$$M = \left(KT/6\pi\eta FD \right)^3 \left[4\pi N_A /3 \left(\upsilon_2 + \delta_1 \upsilon_1 \right) \right]$$

Where M is the molecular weight, K is the Boltzmann constant, T is the absolute temperature, η is the viscosity of the solution, N_A is Avogadro's number, υ_2 and υ_1 are the partial specific volumes of the molecule and solvent water respectively, δ_1 is the fractional amount of water bound to the molecule (hydration number). F is a shaper factor or perrin factor, which is defined to be the ratio of the friction coefficient of the molecule (f) to that of a hard sphere (f_0) with equivalent mass and partial specific volume. For protein shapes modelled as rotational ellipsoids, F can be expressed in terms of the axial ratio p (p=b/a, with b being the equatorial radius and a being the semi axis of revolution). For prolate ellipsoid (p<1),

$$F = f/f_0 = (1-p^2)^{1/2} \Big/ \left\{ p^{2/3} \ln\left[\left\{ 1 + \left(1-p^2\right)^{1/2} \right\} / p \right] \right\}$$

a

For oblate ellipsoid (p>1),

$$F = f/f_0 = (1-p^2)^{1/2} \Big/ \left\{ p^{2/3} \tan\left[p^2 - 1 \right]^{1/2} \right\}$$

b
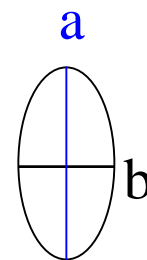
Clearly, the calculation of molecular mass requires, η, $\upsilon 2$, $\delta 1$ and F to be known. The viscosity of water can always be corrected for various temperature as per *CRC Handbook of Chemistry and Physics* and can be closely approximated by the following equations,

$$\log_{10}\eta = \frac{1301}{998.333 + 8.8185(t-20) + 0.00585(t-20)^2} - 3.30233 \text{ in the range } 0 \le t \le 20^\circ C$$

and

$$\log_{10}\frac{\eta}{\eta_{20}} = \frac{1.3272(20-t) - 0.001053(t-20)^2}{t+105} \text{ in the range } 20 \le t \le 100^\circ C,$$

partial specific volume (volume/molecular weight) is calculated from the amino acid composition and volume is computed by summing the mean volumes of each residue in the protein as described in the ref. (*27*). Hydration numbers in the range of 0.3 – 0.4 gm water per gram of protein are common for most of the proteins. Finally, the diffusion coefficient is an important property for the molecular weight calculation. To simply monitor the molecular weight of the oligomers as a function of time, correction to the viscosity and diffusion co-efficient are sufficient. Diffusion coefficient (D) is inversely proportional to hydrodynamic radii, $D = K_B T / 6\pi\eta R_h$, measuring diffusion coefficient either from light scattering or ultracentrifugation experiments, calculates the average molecular weight of the oligomeric species during association. Hydrodynamic radius increases as the molecule associates (Fig. 1.20).
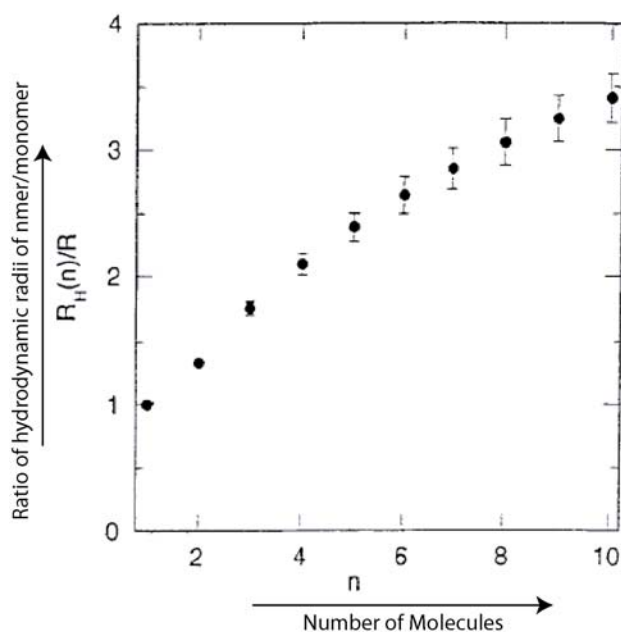
Figure 1.20. Ratio of hydrodynamic radii as number of molecules in the association sphere.

c Microscopic techniques

Seeing is believing. The advent of several advanced digital imaging technologies for the measurement and efficient algorithms for the image-processing make the microscopic techniques amenable for monitoring the fibril formation and morphology. Total internal reflection microscopy, electron microscopy (EM, tunnelling-EM, cryo-EM, etc.), atomic force microscopy are some of the few existing technologies. (Fig. 1.21).
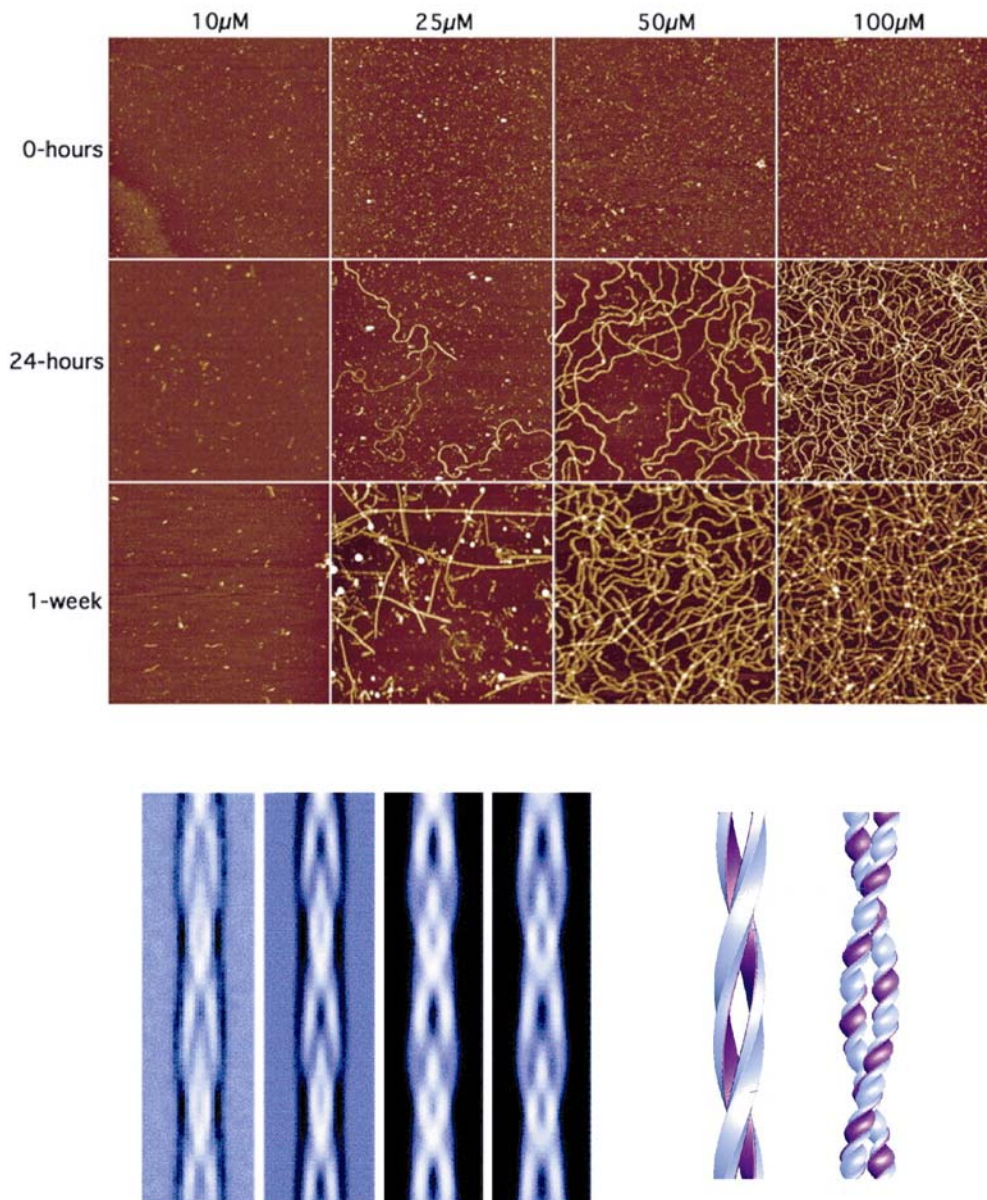
Figure 1.21. Monitoring fibril formation of β-amyloid through AFM, as a function of time and concentration (Top) (Picture adapted from Stine, Jr, W.B., et al., J.Biol.Chem. (2003), 278, 11612); Bottom: Observing fibril morphology of insulin amyloid fibrils through cryo-EM and computed generated models for fibril packing (Picture adapted from Jiménez, J.L., et al., Proc.Natl.Acad.Sci.(2002),99,9196)

d   Secondary structural changes

Proteins that contain various secondary structural elements form fibrils that contain characteristic cross-β-structure.   Hence, monitoring the changes in secondary structural elements is also a way to follow the fibril formation. Circular dichroism spectroscopy (CD) and Fourier-transform infrared spectroscopy (FT-IR) are commonly-used spectroscopic methods apart from NMR, for this purpose. Far-UV CD spectroscopy rely on the principle that the local secondary structure of the peptide backbone creates a chiral environment which rotates the plane of polarized light in a biased way (*28*). Depending upon the rotation, the mean molar CD is calculated.

$$\theta_{mr} = \theta_d \cdot \frac{M}{C \cdot l \cdot n_r} \deg.mol.cm^{-1}.$$

$$\theta_d = \frac{2.303}{4} \cdot (A_L - A_R) \cdot \frac{180}{\pi} \deg$$

Where, C=Concentration (Mol), l= length of the path, (cm), n_r=number of residues, M = molecular weight, (A_L-A_R)= difference between left and right polarized light (observed value). Simply monitoring the (A_L-A_R) as a function of time or concentration is merely enough to follow the kinetics of aggregation (Fig. 1.22). In CD spectroscopy, intensity at 217 nm corresponds to the typical β-sheet arrangement. Minima at 208 nm and 222 nm correspond to helical propensity. Several deconvolution methods are used to extract the exact percentage of secondary structural elements (*29*).
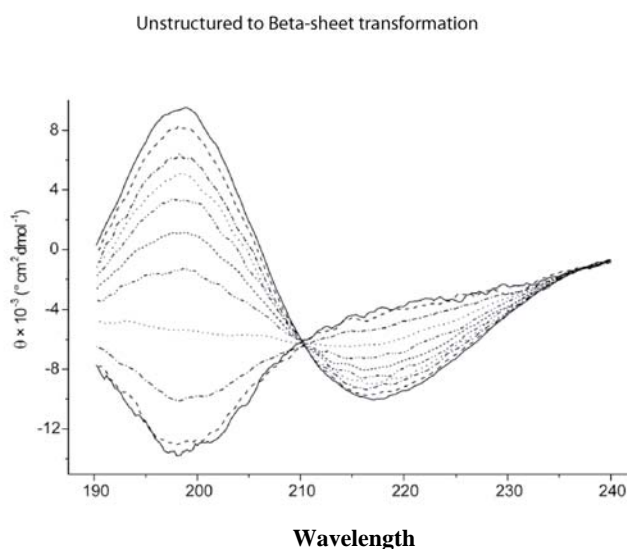
Figure 1.22. A typical secondary structure transition from randomcoil (Minimum at 198nm) to the β–sheet (minimum at 218 nm) for a peptide as a function of concentration .(Figure courtesy: Sandro Keller, FMP-Berlin)

In Fourier transform infra-red spectroscopy (FT-IR), the absorption of energy corresponding to the vibration states of the molecule is measured. If there are N uncombined atoms free to move in three dimension, (3N - 6) vibrational degrees of freedom exist and each absorb at unique wave number between 10,000-200 cm^{-1} (1/wavelength) depending upon the energy of vibration. Infrared spectroscopy is well-established as the method for the analysis of protein secondary structure, in solution and in fibrils. There is a wealth of information that can be used to derive structural information by analyzing the shape and position of bands in the amide I region of the spectrum (*30*). The amide I band results from the C=O stretching vibration of the amide group coupled to the bending of the N-H bond and the stretching of the C-N bond. These vibrational modes, present as infrared bands between approximately 1600-1700 cm^{-1}, are sensitive to hydrogen bonding and coupling between transition dipole of adjacent peptide bonds and hence are sensitive to secondary structure. The presence of a number of amide I band frequencies have been correlated with the presence of α-helical, anti-parallel and parallel β-sheets and random coil (Fig. 1.23). The increase in β-sheet conformation as a function of time is an easy way to monitor fibril formation in solution.
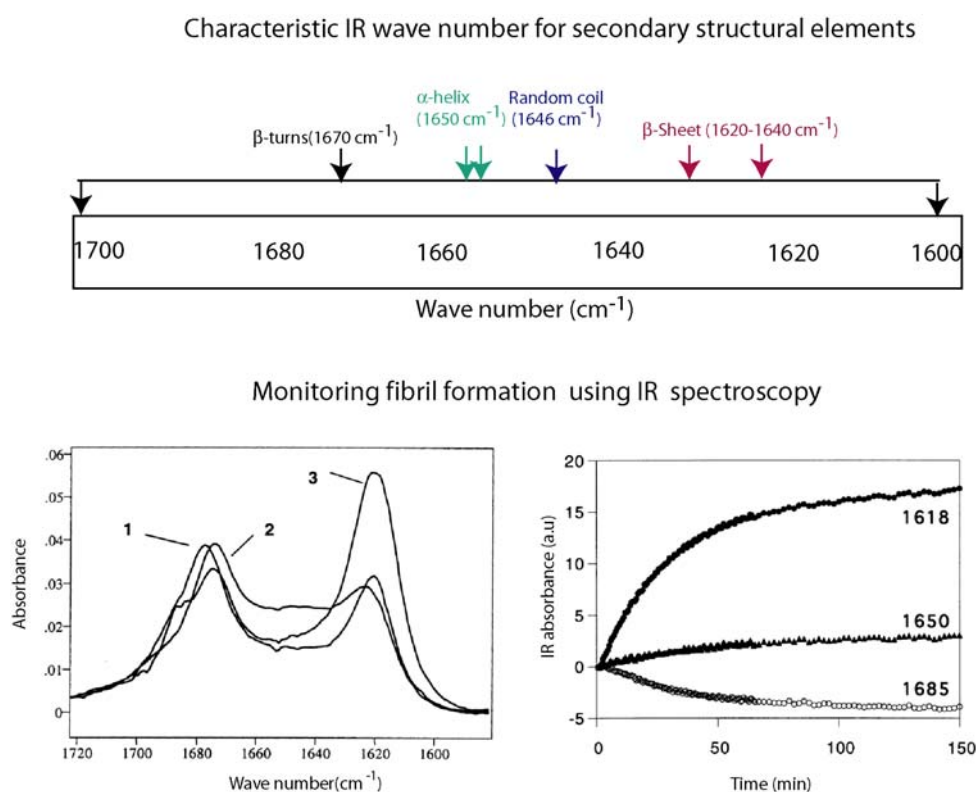
Figure 1.23. Characteristic IR absorption for various secondary structural elements (top). Monitoring IR wavelength as a function of time (bottom) (Adapted from Wilkins, D.W., et al., Eur.J.Biochem.. (2000), 267, 2609)

e   Molecular sizes

Chromatographic separation of macromolecules on the basis of size is referred to as gel filtration or size-exclusion chromatography. Concentration-dependent protein association has been studied in the past using large zone gel filtration chromatography (*31*), while modern techniques use molecular size as a marker to monitor aggregation. A novel method uses stabilization of oligomers using Photo-Induced Cross linking of Unmodified Proteins (PICUP), to study the distribution of oligomers prior to aggregation (*32, 33*). This method enables formation of covalent bonds between closely interacting polypeptide chains without any prior chemical modification or spacers. It uses the energy transfer from transient Ruthenium(III) ion to the proteins which generates stable carbon radical in peptide bonds at the time of irradiation and stabilise the oligomers by making covalent C-C bonds. Considering the formation of oligomers of the nth order from the collision between the $(n-i)^{th}$ species and i^{th} species, then the resulting equation is

$$X_i + X_{n-i} \rightarrow X_n \qquad \text{for i =1 to n-1.} \qquad [1]$$

Because X_n is further consumed to form larger oligomers, the concentration of X_n changes with time according to the following equation

$$[\dot{X}_n] = \sum_{i=1}^{n-1} k_{i,n-i}[X_i][X_{n-i}] - \sum_{i=1}^{\infty} k_{i,n}[X_i][X_n] \quad [2] \quad \text{where } K_{i,n}, K_{i,n-i} = \text{cross linking rate constants}$$

Integration of equation [2] over a function time yields the distribution of the oligomers at any time point during the reaction. The efficiency of cross-linking increases as the population of oligomers increases. Plotting the population of oligomers versus their abundance provide the information about stable oligomers. Control experiments have to be performed to rule out non-specific   interaction of monomers with cross-linkers(Fig. 1.24).
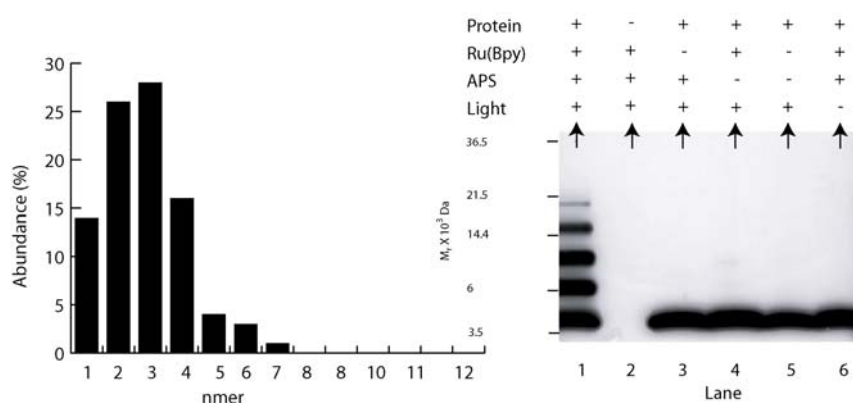
Figure 1.24. Asymmetric gaussian distribution of oligomeric abundance (Left) and SDS-PAGE after PICUP reaction. The gel is representative of each of five independent experiments (Right) (Pictures adapted from J.Biol.Chem. (2001), 276,35176)

1.4.2 Monitoring protein-substrate interactions

Outlined here are the methods that are important to probe the *in vitro* interactions between proteins and their ligands. Library-based methods like phage display, yeast-hybrid systems and other genetic methods are not discussed.

a   Thermodynamics

Every molecular interaction either generates or absorbs small amount of heat. Isothermal titration calorimetry (ITC) measures directly the energy associated with a chemical reaction triggered by the mixing of two components (*34*). A typical ITC experiment is carried out by the stepwise addition of one of the reactants (~5-10 μL per injection) into the reaction cell (~1.5mL) containing the other reactant. The interaction that occurs between protein and ligand on each injection either releases or absorbs a certain amount of heat (q_i) proportional to the amount of ligand that binds to the protein in a particular injection (v x ΔLi) and the characteristic binding enthalpy (ΔH_b) for the reaction: $q_i = v \times \Delta H_b \times \Delta L_i$, where v is the volume of the reaction cell, ΔL_i is the increase in the concentration of bound ligand after the ith injection. Considering the protein ligand interaction as

$$[M]_f + [L] \xrightarrow{\ Ka\ } [ML] \qquad [1]$$
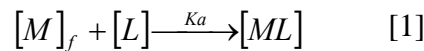
where $[M]_t$ is the total protein concentration including bound and free fractions, K_a is the binding constant, and [L] is the free ligand concentration. Considering Θ as fraction of protein sites occupied by ligand, i.e. $\Theta = \dfrac{[M]_{bound}}{[M]_{total}}$, then the concentration of bound sites in the protein is defined as $n\Theta M_t$. The association constant can be represented as

$$K = \frac{\Theta}{(1-\Theta)[L]} \qquad [2]$$

The known experimental quantity is the total ligand concentration, rather than the free ligand concentration. Hence the free ligand concentration can now be replaced as the difference in concentration of total ligand concentration and the bound ligand concentration.

$$[L] = [L]_t - n\Theta M_t \qquad [3]$$

Rearranging equation [2] and [3], yields a quadratic equation for Θ

$$\Theta^2 - \Theta\left[1 + \frac{[L]_t}{nM_t} + \frac{1}{nKM_t}\right] + \frac{[L]_t}{nM_t} = 0 \qquad [4]$$

Solution to the above quadratic equation gives the value of the fractional saturation (Θ), where, n_i is the number of binding sites. The total heat content q_i of the solution contained in the v at fractional saturation Θ is,

$$q_i = n\Theta M_t \Delta H v \qquad [5]$$

where ΔH is the molar heat of ligand binding. Solving equation [4] and substituting the value of Θ into eq. (5) gives

$$q_i = \frac{nM_t \Delta H v}{2}\left[1 + \frac{[L]_t}{nM_t} + \frac{1}{nKM_t} - \sqrt{\left(1 + \frac{[L]_t}{nM_t} + \frac{1}{nKM_t}\right)^2 - \frac{4[L]_t}{nM_t}}\, \right] \qquad [6]$$

The process of fitting experimental data involves initial guesses of n, K, and ΔH and calculation of $\Delta q(i)$ for each injection and comparison of these values with the measured heat for the corresponding experimental injection. The data analysis is performed using designated software (Fig. 1.25). From ΔH, other thermodynamic parameters like ΔG and ΔS can be calculated as $\Delta G = -RT \ln K_a$ and $\Delta S = \dfrac{\Delta H - \Delta G}{T}$. Repeating the experiment at various temperature the specific heat capacity can also be calculated as $C_p = \left(\dfrac{\partial H}{\partial T}\right)$.
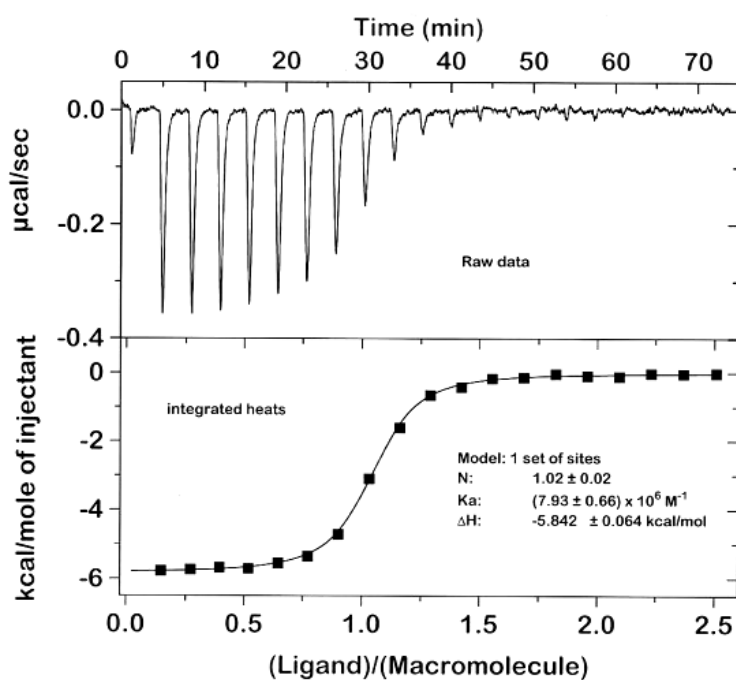
Figure 1.25. A typical Isothermal titration calorimetry (ITC) raw data (top) and its curve fitting (bottom) (Fisher, H.F. and Singh, N., Methods Enzymol , (1995), 259, 194)

b   Indirect methods

Various techniques like, radio-immunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), surface plasmon resonance (SPR) (indirect immobilization and subsequent evanescent wave signal detection) use indirect way of monitoring protein-ligand interactions. For higher affinity, RIA and ELISA are useful, whereas for the detection transient binding (10^{-5} M to 10^{-7}M) SPR techniques are preferable. RIA is based on the principle of differential labelling. Unlabelled analyte competes with radio-labelled analyte (usually labelled with I-125) for the limited binding sites available on the protein. At the end of the reaction, the labelled analyte-protein complex is separated and measured for radioactivity. The amount of radioactivity separated is compared with the values of known analyte standards and the concentration of analyte present in the sample (Fig 1.26a). ELISA rely on the detection of substrate complexes with primary antibody-labelled protein and the interaction is detected using secondary antibody, which can recognize the primary antibody (Fig. 1.26b).
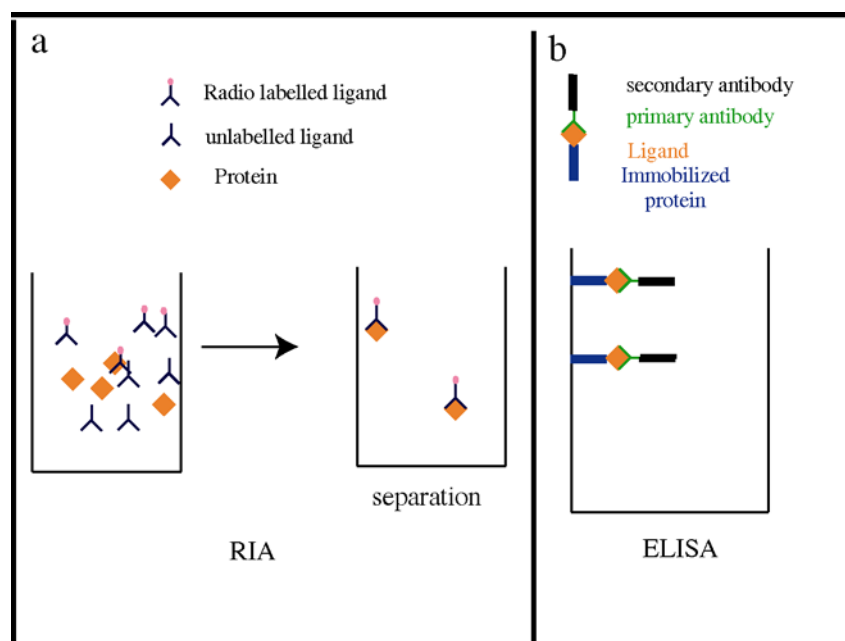
Figure 1.26. Principle of RIA and ELISA

Surface plasmon resonance uses the basic principle of opto-electronics. An evanescent (decaying) electrical field associated with the plasma wave travels for a short distance (~300 nm) into the medium from the metallic film (*35*).   If the surface is immersed in an aqueous buffer (refractive index or μ ~1.0) and the protein (μ ~1.33) binds to the surface, this results in an increase in refractive index which is detected by a shift in the angle called as, ϕ_{spr}. The

sensitive photo-detector measures very small changes in ϕ_{spr}. The change is quantified in resonance units or response units (RUs) with 1 RU (R) equivalent to a shift of 10^{-4} (ϕ_{spr}) degrees (Fig. 1.27). The association (k_1) and dissociation (k_{-1}) rate constants are obtained by fitting the nonlinear equation

$$R_t = R_{max}\left[1 - e^{-(k_1C + K_{-1})(t-t_0)}\right]$$

where R_t is the response at time t, R_{t0} is the amplitude of the initial response. R_{max} is the maximum response, C is the concentration of the analyte in the solution. The ratio of k_1 and k_{-1} provides association constant.
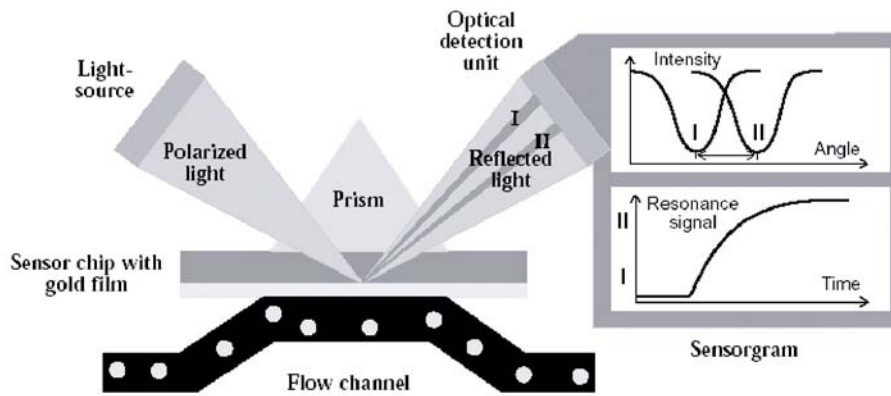
Figure 1.27. Principle of Surface plasmon resonance

c Fluorescence-based experiments

Fluorescence resonance energy transfer (FRET) is a technique based on a columbic interaction between an optically excited donor fluorophore and an acceptor fluorophore in its electronic ground state. The donor and acceptor fluorophore are attached to the protein and the substrate, respectively (*36*). After excitation, the donor fluorophore transfers the excitation energy to the acceptor. The rate constant for this energy transfer can be calculated from Förster's formula:

$$K(r) \propto K\frac{k^2\phi}{r^6\tau_d}\int\frac{f_D(v)\varepsilon_A(v)}{v^4}dv$$

Here, r is the distance between donor and acceptor, $f_D(v)$ and $\varepsilon_A(v)$ are the fluorescence spectrum of the donor and the absorption spectrum of the acceptor, respectively. κ is an orientation factor which accounts for different angles between donor and acceptor molecule

dipole moments, ϕ_D is the quantum yield of the donor fluorescence and τ_D the corresponding lifetime. Since all parameters except r can be determined or estimated before an experiment, FRET allows distances to be measured in specific labelled biomolecules. The range that is measurable by this method typically covers distances from 10 Å to 100 Å.
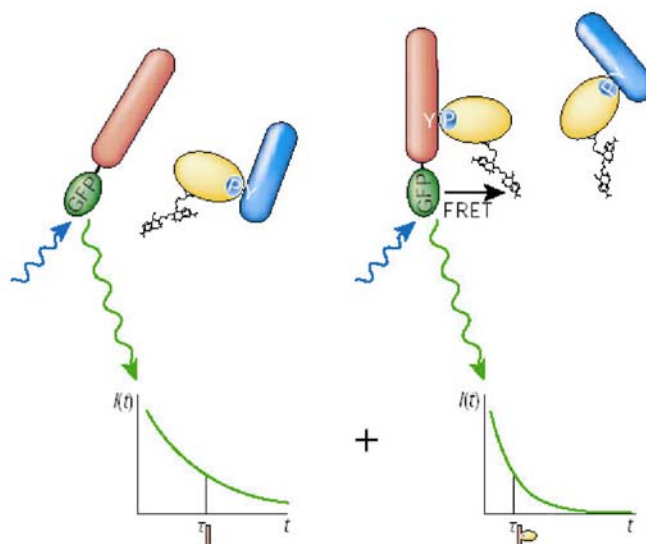
Figure 1.28. Principle of FRET, the fluorescence resonance energy transfer from the GFP fluorescence to the acceptor dye in the substrate molecule(right) compared to the control experiment (left). The lifetime correlation experiment can also indicate that a sharp decrease in fluorescence lifetime shows interaction with partners.

Luminescence resonance energy transfer (LRET) using lanthanide chelates is a technique which measures distances between donor and acceptor molecules. It has a number of advantages over conventional FRET methods, as the method exploits the long decay time of lanthanides. Hence, distances of over 100 Å can be determined, with less uncertainty than standard FRET methods. Another method is also in use, that is based on bioluminescence, termed as Bioluminescence resonance energy transfer (BRET). This method does not rely on the incident radiation, but on its own bioluminescence property. In practice, one of the two proteins is genetically fused to the blue light emitting *Renilla* luciferase and the other one to a blue light absorbing yellow fluorescent protein. If the two hybrid proteins interact, the excitation energy of the luciferase may be transferred to the fluorescent protein, resulting in an easily detected yellow-shift in the luminescence spectrum (*37*).

1.5   NMR spectroscopy

1.5.1   Principles of NMR spectroscopy

Nuclei with half-integer spin angular momentum (I=1/2n, where n=non-zero integer) generate a nuclear magnetic moment, μ. This nuclear magnetic moment is collinear with the spin angular moment vector, $\mu = \gamma I$, by convention, $\mu_z = \gamma I_z$, in which γ is characteristic for each nucleus and is termed "magnetogyric ratio". The spin quantum number is quantized and adopts (2I+1) possible values. Hence $I_z = \hbar m$, where $\hbar = \dfrac{h}{2\pi}$ and h is the Planck's constant.

Therefore, the magnetic moment is defined as, $\mu_z = \gamma \hbar m$. In the presence of an external magnetic field, the energy of spin states are no longer degenerate and split according to the Zeeman interactions h, $E = -\mu \cdot B$. Since the energy is a scalar product between the external magnetic field and the nuclear magnetic moment, the minimum interaction energy is reached, when the projection of μ on B is maximized. Also, $E = -\gamma \hbar m B_o$, where B_0 is the static magnetic field strength.

After inserting the NMR sample into the static magnetic field B_0, the spins in the ensemble of molecules in the NMR tube are either in the α- or β-state. The precession of nuclear spins around the axis of the static magnetic field with a *z*-axis projection being parallel or antiparallel to B_0, respectively. The population of the α- and β-states according to a Boltzmann distribution can be described as

$$N(\alpha/\beta) = \exp^{\{\mu B_0/(kT)\}} \approx 1 + \mu B_0/(kT)$$

The population difference between the two states are fairly small (~N(α/β)=1.0001/1). Nevertheless, the small population difference produces an effective magnetization along the *z*-axis. However, there is no magnetization observed in the *x,y*-plane since the individual precession frequencies are uncorrelated. In order to measure the precession frequencies, the system is disturbed and brought into a non-equilibrium state which allows transverse (*x* or *y*) magnetization to be monitored. The frequency which exactly matches (resonates) with the energy difference of the two states is absorbed. This frequency is termed as Larmor frequency,

$$v_0 = \frac{\gamma B_o}{2\pi}$$
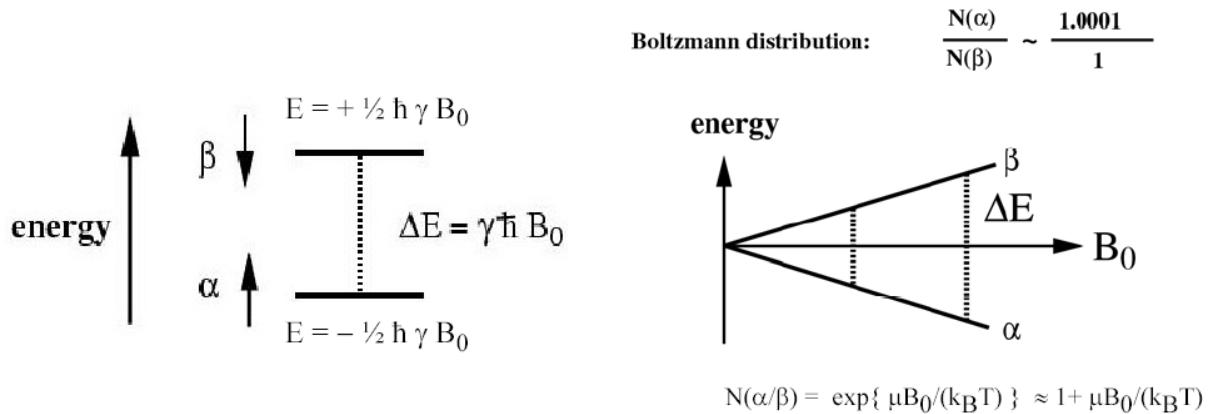
Figure 1. 29. Principle of NMR: Quantized energy level absorption and change in population of spin states.

The signal-to-noise (S/N) can be represented as

$$S/N \approx N\gamma_I\gamma_S^{(3/2)}B_0^{(3/2)}(NS)T_2^{(1/2)} \quad \text{where}$$

S/N = signal-to-noise ratio

N = number of spins (sample concentration)

γ_I = gyromagnetic ratio of excited spins

γ_S = gyromagnetic ratio of detected spins

B_0 = static magnetic field strength (i.e. 600 MHz 1H)

NS= number of scans

T_2 = transverse relaxation time

The separation of the energy levels defines the population difference (according to Boltzmann statistics) which provides the magnetization that is measured in an NMR experiment. Since ΔE is very small, NMR is a rather insensitive spectroscopic method, and optimizing the signal-to-noise ratio is always a critical issue for NMR. Notably, there exists a local magnetic field which depends upon the shielding tensor, σ_{kk}. For the case of spherical symmetry,

$$\sigma_{kk} = \frac{4\pi e^2}{3mc^2}\int r\rho(r)dr \, ,$$

the integral value denotes the probability to find an electron within the radius of r from the centre of the nucleus. Hence, the Larmor frequency changes according to the electronic environment as

$$v_0 = \frac{\gamma(1-\sigma_{kk})B_o}{2\pi}$$

This gives rise to different NMR resonance frequencies for nuclei attached to different chemical groups (*38*).

1.5.2. Communication within NMR signals.

The power of NMR spectroscopy relies on the transfer of magnetization from one nucleus (I) to another (S) and this communication can occur either through bond (via scalar coupling) or through space (through NOE). As NMR deals with energy of the spin systems, it is more appropriate to study the change in population and hence the observed resonances in terms of change in energy of the system. The Hamiltonian ($\mathcal{H}$) refers to the observable corresponding to the total energy of a system. At time *t,* physical state of various systems will be characterized by the correspondent wave functions $\psi(r,t)$, where the *r* denotes the position of the coordinates relevant to the system under study. The time variance of the function $\psi^k(t)$ will be determined by the Schrödinger equation

$$\frac{d}{dt}\left|\psi^k(t)\right\rangle = -i\,\mathcal{H}(t)\left|\psi^k(t)\right\rangle$$

Introducing the complete set of orthonormal functions ϕ_n, the wave functions $\psi^k(t)$ may be written as $\psi^k(t) = \sum a_n^k(t)\phi_n$. The term a_n^k denote probability amplitudes of kth system of the ensemble to be in the respective states ϕ_n. $\left|a_n^k(t)\right|^2$ represents the probability that a measurement at time *t* finds the kth system of the ensemble to be in particular state. Hence, the value of $\left|a_n^k(t)\right|^2$ should be 1 for all k values. In quantum mechanics, the state of the spin-system is described by the density operator $\rho(t)$. The density operator can be defined by the density matrix elements $\rho_{mn}(t)$.

$$\rho_{mn}(t) = \frac{1}{N}\sum_{k=1}^{N}\left\{a_m^k(t)a_n^{k^*}(t)\right\}$$

Clearly, the matrix element $\rho_{mn}(t)$ is the ensemble of average *probability* amplitudes, a_n^k. The quantity $\rho_{mn}(t)$ represents the probability that a system chosen *at random* from the ensemble at time *t* is found to be in the state ϕ_n. The density operator is hermitian and normalized.

Diagonal elements of density matrix are non-negative and their summation value is one. The equation of motion for the density operator can be represented as

$$\frac{d}{dt}\rho(t) = -i \ [\mathcal{H}(t), \rho(t)]$$

This differential equation, called, Liouville-von Neumann equation or simply density operator equation, is of central importance for calculating the dynamics of quantum mechanical systems. By selecting suitable rotating frame, the Hamiltonian can often be made time-independent within finite segments of time, and the evolution can be expressed by a sequence of unitary transformations of the type.

$$\rho(t+\tau_1+\tau_2) = \exp\{-i\,\mathcal{H}_2(t)\tau_2\}\ \exp\{-i\,\mathcal{H}_1(t)\tau_1\}\rho(t)\ \exp\{+i\,\mathcal{H}_1(t)\tau_1\}\ \exp\{+i\,\mathcal{H}_2(t)\tau_2\}$$

with the propagators $\exp\{+i\,\mathcal{H}_k(t)\tau_k\}$. This equation applies to any sequence of intervals τ_k in which time-independent average Hamiltonian can be defined.

From the Hamiltonian point of view, the single spin ½ nucleus in a static field is considered as

$$\mathcal{H} = \omega_0 I_z + \omega 1\left(I_x \cos\left(\omega_{rf} + \phi\right) + I_y \sin\left(\omega_{rf} + \phi\right)\right)$$

where the first term represents the influence of static magnetic field and the second term represents the rf pulse. The effective Hamiltonian can be written as, $\mathcal{H} = \mathcal{H}_z = \sum_{i=1}^{N} \omega_i I_{iz}$, in which ω_i is the Larmor frequency. It can be represented in terms of rotational frame during free precession as offset of spin, Ω_i. The rf pulses applied down a specific axis induce rotations in a plane orthogonal to that axis. $I_z \xrightarrow{\alpha I_{\pm x}} I_z \cos\alpha \mp I_y \sin\alpha$. For the coherence transfer mediated via bond electrons between the interacting nuclei, the Hamiltonian can be represented as summation of free precession laboratory frame Hamiltonian and scalar coupling constant (J_{IS}),

$$\mathcal{H}_z = \sum_{i=1}^{N} \omega_i I_{iz} + 2\pi \sum_{i=2}^{N} \sum_{j=1}^{i-1} J_{ij} I_i I_j$$

Pulse sequence elements can be combined to produce more complex sequences designed to perform specific tasks. The sensitivity in NMR is dependent upon the magnetogyric ratio and natural abundance of the nucleus with half-integral spin quantum number. As both these terms are superior in protons over any other nuclei, transfer of magnetization from proton to any other nuclei can be performed by complex pulse sequences. A unique basic element in all hetero-nuclei experiments involve in transfer of magnetization from proton to the respective

hetero-nucleus (^{15}N, ^{13}C, etc.) is the INEPT building block (Insensitive nuclei enhancement through polarization transfer) (Fig. 1.30).
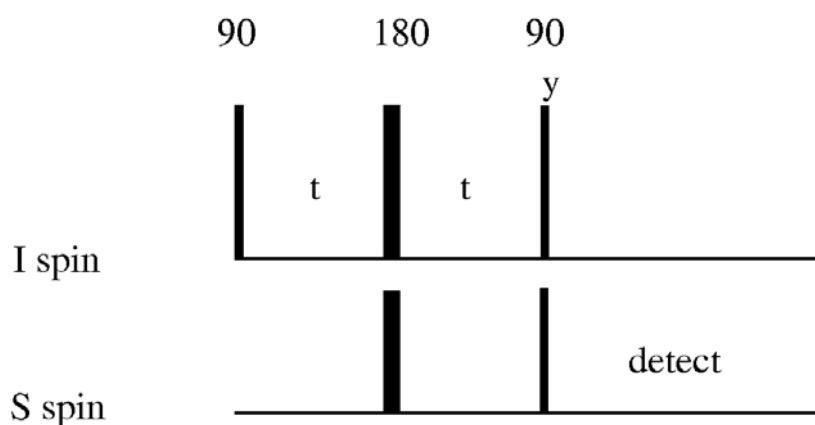
Figure 1.30. A basic element, INEPT pulse sequence used for magnetization transfer. phases are x otherwise mentioned.

In INEPT transfer, the magnetization transfer occurs as follows

$$I_z \xrightarrow{\frac{\pi}{2}I_x - t - \pi(I_x + S_x) - t} \left\{ I_y \cos(2\pi J_{IS}t) - 2I_x S_z \sin(2\pi J_{IS}t) \right\} \xrightarrow{\frac{\pi}{2}(I_y + S_x)} \left\{ Iy \cos(2\pi J_{IS}t) - 2I_z S_y \sin(2\pi J_{IS}t) \right\}$$

If the delay $t=1/4J_{IS}$, then the final signal is given by $-2I_zS_y$, where the antiphase magnetization has been transferred to the S spin. It is interesting to know that under the effect of scalar coupling Hamiltonian, the magnetization can transfer from the spin I_y to spin S via an unique antiphase magnetization term $2I_xS_z$. In a nutshell, the macroscopic magnetization on I_z is transferred to the transverse plane of S nuclei by the action of rf pulses and this process is communicated through direct-bond between the I and S.

The z-magnetization of different nuclei exchanged between spins on the way to thermal equilibrium by means of the dipolar interaction between different nuclei is called the nuclear Overhauser effect (NOE). NOE effect is responsible for the magnetization transfer through space for the nuclei that are in close proximity. Considering the transfer of magnetization from the nucleus I to the nucleus S, whose magnetogyric ratio are γ_I and γ_S, respectively, the communication of NMR signals via NOE is proportional to the square of the magnetogyric ratio of the individual nucleus and time of NOE build-up, but inversely proportional to the sixth power of the distance between the dipoles under solution condition.

1.5.3 NMR properties to monitor aggregation

Oligomerization involves changes in hydrodynamic radii and molecular weight of protein. Translational diffusion and relaxation are the two important parameters that are dealing with dynamic systems like protein-protein interactions and protein movements in various time scales. Translational diffusion is defined as translational motion in the absence of a concentration gradient. The diffusion coefficient (D) of translational motion is indicative of the hydrodynamic properties of molecules. The pioneering work of Stejskal and Tanner showed that the diffusion coefficient of molecules in solution can be measured using pulse field gradient (PFG) NMR methods (*39*). In the experiment (90°-τ(PFG)-180°-τ(PFG)-acq), the strength of the PFG pulse is increased in successive experiments. The echo amplitude [I] is differentially attenuated in each spectrum due to translational diffusion, which is related to the diffusion coefficient according to

$$ I = I(0)\exp\left[-(\gamma \delta G)^2 \left(\Delta - \frac{\delta}{3} \right) D \right], \; D = \left(\frac{kT}{6\pi \eta R_H} \right) $$

where η = viscosity, R_H = hydrodynamic radius, γ = ^{1}H magnetogyric ratio, δ= PFG duration, G= gradient strength, and Δ = time between PFG pulses . The stimulated echo experiment has been designed in order to avoid T_2 relaxation effects by storing the magnetization along the Z axis during T, so that the relaxation depends primarily on T_1, which is usually much longer than T_2 for proteins (Fig. 1.31). Longer diffusion times (Δ) can be used for larger proteins without significant signal loss from relaxation.
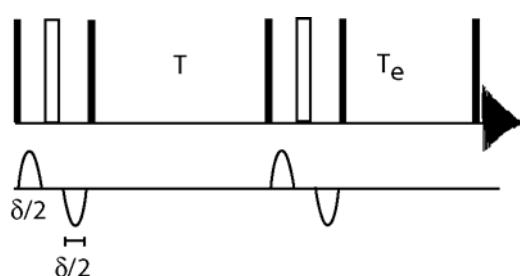
Figure 1.31. Stimulated-echo pulse sequence with bipolar gradient pulses. T_e refers longitudinal eddy current delay. Open wide boxes indicate 180 ° pulses and narrow closed boxes denote 90° pulses. Gradient pulses are indicated in Gaussian shaped pulses.

Though the average molecular weight can be calculated from diffusion coefficient, changes in the state of aggregation can be monitored by introducing small organic molecule as an internal standard in order to correct the change in viscosity as a consequence of protein aggregation (*40*). Hence, the hydrodynamic radius of the peptide during protein association

can be monitored by monitoring hydrodynamics of the peptide ($R_H{}^{peptide}$) as a function of time.

$$R_H^{peptide} = R_H^{ref} \times \frac{D_{ref}}{D_{peptide}}$$

The improved diffusion experiments, such as heteronuclear-stimulated echoes (*41*), diffusion exchange (*42*), DOSY-HSQC (*43*) provide more information regarding protein-protein association.

1.5.4 Relaxation properties and motion

The ^{15}N NMR relaxation studies have emerged as a powerful approach for the determination of the motional properties of molecules in solution (*44*). In an ^{15}N relaxation experiment, one creates non-equilibrium spin order and records how this relaxes back to equilibrium. The relaxation of spin orders is related to rotational motions of the amide group. This includes the overall internal motions of the macromolecule. The effect of the motions on relaxation is described by a spectral density function J(ω). This function describes the relative distribution of the frequencies of the rotational diffusive motions. The spectral density can be represented in terms of tumbling motion of the molecule, $J(\omega) = J(\omega) = \dfrac{\tau_c}{1 + \omega^2 \tau_c^2}$, in which, τ_c is the rotational correlation time of molecule. In slow motion limit ($\omega\tau_c \gg 1$), the spectrum density function is compared to the correlation time. The correlation time is a time of the movement of the dipolar field when the molecule rotates. For a spherical molecule of radius (*a*), rotating in a viscosity (η), τ_c is defined as, $\tau_c = \dfrac{4\pi a^3 \eta}{3kT} = \dfrac{V\eta}{kT}$, where V is the volume of the molecule.

As the volume is proportional to the molecular weight, the rotational correlation time is also proportional to the molecular weight. As a rule of thumb, the rotational correlation time (in nanosecond) of a molecule in aqueous solution at room temperature is about half its molecular weight in kDa. $\tau_c \cong Mol.wt/2$.

At equilibrium, the magnetization of the ^{15}N spins is aligned along the external field, and this alignment can be inverted by radio frequency pulses. The spins will relax back to equilibrium along the direction of the magnetic field with a longitudinal relaxation time, T_1, or the longitudinal relaxation rate, R_1^N. It can be shown that the T_1 relaxation rate is proportional to the square of the dipolar field strength times the spectral density function. The magnetization can also be oriented perpendicular to the external magnetic field (xy plane). The relaxation time of this spin order back to equilibrium is called the transverse relaxation time, T_2, and the

rate can be denoted as R_2^N. Both of these relaxation parameters are described in terms of spectral density function,

$$R_1^N = D_{IS}^2\left[J(\omega_N - \omega_H) + 3J(\omega_N) + 6J(\omega_N + \omega_H)\right] + CSA^2 J(\omega_N)$$

$$R_2^N = 0.5d^2\left[4J(0) + J(\omega_N - \omega_H) + 3J(\omega_N) + 6J(\omega_N + \omega_H)\right] + \frac{c^2}{6}\left[4J(0) + 3J(\omega_N)\right] + R_{ex}$$

In these relations, D_{IS} represents the strength of the dipolar interaction between the proton and nitrogen, and CSA represents the chemical shift anisotropy. These are the main relaxation mechanisms for ^{15}N in proteins. ω_N and ω_H are the resonance frequencies of ^{15}N and ^1H. These relaxation times are effective in magnetization transfer. The build up of antiphase magnetization in the INEPT is described as

$$\langle 2I_y S_z \rangle(T) = \sin(\pi J_{IS} T)\exp\left(-R_2^I T\right)\langle I_x \rangle(0) \text{ , in which}$$

$$R_2^I = \frac{2}{5}\left[\frac{2}{9}(\gamma_I B_o \Delta\sigma_I)^2 + \frac{1}{2}(\hbar\gamma_I\gamma_s/r_{IS}^3)^2\right]\tau_c + \frac{1}{2T_{IS}} + \frac{1}{2T_{2I}}$$

where T is the magnetization transfer period. r_{IS} is the distance between the two nuclei involved. $\Delta\sigma_I$ is the CSA of the nucleus I, B_0 is static magnetic field, γ_I and γ_S are the magnetogyric ratios of I and S, respectively. T_{2I} and T_{1S} account for the transverse relaxation of spin I and the longitudinal relaxation of the spin S.

A third relaxation parameter is the so-called heteronuclear nuclear Overhauser effect (NOE). This is measured by saturating the proton (^1H) signal and observing changes in the ^{15}N signal. The rate at which this occurs is the heteronuclear cross relaxation rate, R_{HN} (H_Z <---> N_Z); for long proton saturation, it reaches the steady state NOE (H_Z <---> N_Z) value. The relaxation of the spin orders is due to rotational diffusive motions of the nitrogen atom and the orientation of its chemical bonds (NH bond) relative to the external field. These are essentially the three useful measurements to probe mobility of macromolecules. In terms of spectral density functions, heteronuclear NOE can be represented as (*38*).

$$NOE(H_z \leftrightarrow N_z) = 1 + \frac{\gamma_H D_{IS}^2}{\gamma_N R_N(N_z)}\left[6J(\omega_N + \omega_H) - J(\omega_N - \omega_H)\right]$$

The equations indicate that ^{15}N relaxation is primarily sensitive to very fast motions in the order of the frequencies of ^1H and ^{15}N as well as the sum and difference thereof. However, the transverse relaxation rates can also be influenced by slow conformational exchange, R_{ex}. For two exchanging sites, $R_{ex} = \Delta\omega^2 p_1 p_2 \tau_{ex}\left[1 - (2\tau_{ex}/\tau_{cp})\tanh(\tau_{cp}/2\tau_{ex})\right]$, where $\Delta\omega$ is difference in chemical shift, p_i and p_2 are the population of nuclear spin state 1 and 2, respectively. τ_{ex} is the reduced lifetime of the exchanging sites. τ_{cp} is the delay between pulses

in the spin-echo experiment. As R_{ex} affects only those residues that have different chemical shifts in the two species, site specific information can be derived. With the measurement of the three relaxation parameters, the longitudinal relaxation rate, R_1^N, the transverse relaxation rate, R_2^N, and the heteronuclear NOE(H_Z<->N_Z), many useful information like equilibrium constant, relative population of oligomeric species, molecular motions can be derived (*45*).

1.5.5   NMR based screening in drug discovery

Measurement of changes in chemical shifts is an easy way to observe site-specific protein-ligand interactions. By mapping the ligand induced chemical shift perturbations of the protein, the location of the binding site can be obtained. Structure activity relationships (SAR) by NMR was the first method for NMR screening (*46*). SAR by NMR relies on a fragment-based approach, wherein a large library of small molecules is screened using 2D ^{1}H-^{15}N correlation spectra of the target protein as a readout. From changes in protein amide chemical shifts, one can readily identify whether the binding has occurred for one or more components in a mixture of compounds. This technique is also useful for exploiting site-specific protein-protein interactions.

In combination with other biophysical techniques like ITC experiments and FRET techniques, cooperative or allosteric site specific binding can be obtained (*47*). Measuring intermolecular NOE of these differentially labelled protein is another way to achieve site-specific interaction. This intermolecular NOE experiment relies on cross-correlated relaxation, where the protein magnetization can be transferred to the ligand and *vice versa*. Editing and filtering techniques, transfer NOE, reverse NOE pumping, saturation transfer difference NMR are complementary techniques relying on the same principle which allow one to explore the specific intermolecular NOE signals (*48, 49*).

1.6   References

1.   Crick, F. H. C. (1958) *Symp.Soc.Exp.Biol. 12*, 138-163.
2.   Crick, F. H. C. (1970) *Nature 227*, 561-563.
3.   Hunter, N. (1999) *Trends Microbiol. 7*, 265-266.
4.   Prusiner, S. B. (1998) *Proc.Natl.Acad.Sci.U.S.A. 95*, 13363-13383.
5.   Wu, H. (1931) *Chin. J. Physiol. 5*, 321-344.
6.   Epstein, C. J., Goldberger, R. F. and Anfinsen, C.B. (1963) *Cold Spring Harbor Symp.Quant.Biol. 28*, 439.
7.   Levinthal, C. (1968) *J.Chim.Phys. 65*, 44.
8.   Leopold, P. E., Montal, M., Onuchic, J.N. (1992) *Proc.Natl.Acad.Sci.U.S.A. 89*, 8721-8725.
9.   Dobson, C. M., Evans, P.A., Radford, S.E. (1994) *Trends Biochem.Sci. 19*, 31-37.
10.   Dobson, C. M. (2003) *Nature 426*, 884-890.
11.   Selkoe, D. J. (2003) *Nature 426*, 900-904.
12.   Sunde, M., Blake, C.C.F. (1997) *Adv.Protein.Chem. 50*, 123-159.
13.   Sitia, R., Braakman, I. (2003) *Nature 426*, 891-894.
14.   Nissen, P., Hansen, J., Ban, Nenad, Moore, P.B., Steitz, T.A. (2000) *Science 289*, 920-930.
15.   Hartl, F. U., Hayer-Hartl, M. (2002) *Science 295*, 1852-1858.
16.   Walter, S., Buchner, J. (2002) *Angew.chem.Int.Ed. 41*, 1098-1113.
17.   Ruddon, R. W., Bedows, E. (1997) *J.Biol.Chem. 272*, 3125-3128.
18.   Horwich, A. (2002) *J.Clin.Invest. 110*, 1221-1232.
19.   Griffith, J. S. (1967) *Nature 215*, 1043-1044.
20.   Prusiner, S. B. (1982) *Science 216*, 136-144.
21.   Jarrett, J. T., Lansbury, P.T.Jr. (1993) *Cell 73*, 1055-1058.
22.   Serio, T. R., Cashikar, A.G., Kowal, A.S., Sawicki, G.J., Moslehi, J.J., Serpell, L., Arnsdorf, M.F., Lindquist, S.L. (2000) *Science 289*, 1317-1321.
23.   Lomakin, A., Teplow, D.B., Kirschner, D.A., Benedek, G.B. (1997) *Proc.Natl.Acad.Sci.U.S.A. 94*, 7942-7947.
24.   Ferrone, F. (1999) *Methods enzymol. 309*, 256-274.
25.   le Vine, H. (1999) *Methods enzymol. 309*, 274-284.
26.   Bloomfield, V. A. (2000) *Biopolymers 54*, 168-172.
27.   Harpaz, Y., Gerstein, M., Chothia, C. (1994) *Structure 2*, 641-649.
28.   Berova, N., Nakanishi, K., Woody, R.W. (Eds.). (2000) *Circular Dichroism: principles and applications Wiley-VCH*.
29.   Sreerama, N., Woody, R.W. (2000) *Anal.Biochem. 287*, 252-260.
30.   Surewicz, W. K., Mantsch, H. H., & Chapman, D. (1993) *Biochemistry 32*, 389-394.
31.   Valdes, R. J., Ackers, G.K. (1979) *Methods enzymol. 61*, 125-42.
32.   Fancy, D. A., Kodadek, T. (1999) *Proc.Natl.Acad.Sci.U.S.A. 96*, 6020-6024.
33.   Bitan, G., Lomakin, A., Teplow, D.B. (2001) *J.Biol.Chem. 276*, 35176-35184.
34.   Leavitt, S., Freire, E. (2001) *Current Opin.Struct.Biol. 11*, 560-566.
35.   McDonnell, J. M. (2001) *Current Opin.Chem.Biol. 5*, 572-577.
36.   Miyawaki, A., Tsien, R.Y. (2000) *Methods enzymol. 327*, 472-500.
37.   Heyduk, T. (2002) *Curr. Opin.Biotechnol. 13*, 292-296.
38.   Cavanagh, J., Fairbrother,W., Palmer,A.G.,Skelton, N.J. (1996) *Protein NMR spectroscopy principles and practice Academic press*.
39.   Stejskal, E. O., Tanner,J.E. (1965) *J.Chem.phys. 42*, 288-292.
40.   Yao, S. k., Howlett, G.J., Norton, R.S. (2000) *J.Biomol.NMR 16*, 109-119.
41.   Ferrage, F., Zoones, M., Warschawski, D., Popot, J., Bodenhausen, G. (2003) *J.Am.Chem.Soc. 125*, 2541-2545.
42.   Andrec, M., Prestegard, J.H. (1997) *J.Biomol.NMR 9*, 136-150.

43.   Buevich, A. V., Baum,J. (2002) *J.Am.Chem.Soc. 124*, 7156-7162.

44.   Grzesiek, S. (2003) *Notes on relxation and dynamics EMBO practical course on NMR, Heidelberg*, Germany.

45.   Bernado, P., Akerud, T., de la Torre, J., Akke, M., Pons, M. (2003) *J.Am.Chem.Soc. 125*, 916-923.

46.   Shuker, S. B., Hajduk, P.J., Meadows, R.P., Fesik, S.W. (1996) *Science 274*, 1531-1534.

47.   Tochtrop, G. P., Richter, K., Tang, C., Toner, J.J., Covey, D.F., Cistola, D.P. (2002) *Proc.Natl.Acad.Sci.U.S.A. 99*, 1847-1852.

48.   Diercks, T., Coles, M., Kessler, H. (2001) *Curr. Opin.Chem.Biol. 5*, 285-291.

49.   Hajduk, P. J., Meadows, R.P., Fesik, S.W. (1999) *Quart.rev.Biophysics 32*, 211-240.

2.0   Research objectives and thesis organization

The aim of the study is to understand the structural basis of protein aggregation and their interactions with molecular chaperones.

We used different systems to study the mechanism of protein aggregation, interactions of misfolded and aggregated proteins with molecular chaperones and the consequence of these interactions.

The molecular mechanism of aggregation was studied using (i) a polypeptide sequence which derived from yeast prion Sup35 (chapter 3), (ii) Alzheimer's disease causing amyloid peptide, Aβ(1-40) (Chapter 4) and (iii) a globular protein PI3-SH3 protein (Chapter 5).

The molecular mechanism of interactions of molecular chaperone with their substrates was investigated using (a) Sup35 and Hsp104 (Chapter 3) and (b) Aβ(1-40) and αB crystallin (Chapter 4).

Chapter 3:

Sup35 is a soluble protein which forms amyloid-like fibrils *in vitro*. The structural characterization of the aggregation pathway was studied in this research project, where the prion domain, comprising of amino acids 5 to 26 of the N-terminal domain of Sup35 (Sup35[5-26]) was selected. In addition, the aggregation of Sup35[5-26] is prevented upon interacting with the molecular chaperone, Hsp104. From solution state NMR and biochemical data, a more detailed picture of how the interaction between Sup35 and Hsp104 occurs at a molecular level is revealed.

We used 1D-NMR spectroscopy to monitor the aggregation kinetics, Real-time DOSY NMR spectroscopy to monitor the oligomerization and Real-time STD experiments to monitor the differential binding affinity of different oligomers. Secondary structural analysis were performed using ^{13}C HSQC and CD experiments. In addition, Equilibrium dialysis experiments were used to identify the distribution of oligomers.

Chapter 4:

Aβ(1-40) is an Alzheimer disease-causing amyloid peptide. Cytoplasmic and extra-cellular Aβ(1-40) are highly toxic to neuronal cells. During the interaction with a molecular

chaperone, αB crystallin, the fibril formation is drastically reduced, but at the same time, an increase in neurotoxicity is observed.

In this project, the chemical groups of Aβ(1-40) that are involved during fibril assembly and also during interaction with αB crystallin were identified. The reason behind increased neurotoxicity of Aβ(1-40) induced by αB crystallin was understood. An additional property of molecular chaperones that might be involved in substrate recognition and function was proposed.

We used 2D-STD TOCSY experiments to map the chemical groups of Aβ(1-40) that are involved during the interactions with αB crystallin. Diffusion experiment was used to analyse the size of the oligomers of Aβ(1-40). A simple NMR-based assay was designed to analyse redox proteins. Colorimetric assay were used to authenticate the NMR based redox assay. Isothermal titration calorimetry was used to understand the interaction between Aβ(1-40), αB crystallin and copper. Secondary structures were analysed using CD experiments.

Chapter 5:

Solid-state and solution-state NMR results were combined to address the fibril formation of a Src homologue protein (p58α subunit of bovine phosphatidyl-inositol-3'-kinase (PI3-SH3)). Modification of a single amino acid that modulates the quaternary arrangement of PI3-SH3 was observed. Key residues that may be involved in the process of PI3-SH3 protein association were identified. In addition, a key mechanism which the PI3-SH3 may used to form ordered aggregates was proposed.

Solid-state NMR experiments (Homonuclear, Heteronuclear correlation experiments, T-MREV, 1D-^{15}N-CP-MAS) were used to understand the PI3-SH3 protein in fibrillar conditions. Solution-state NMR experiments (3D-Isotope filtering experiments, standard 3D-^{13}C-TOCSY-^{15}N-HSQC) were used to understand the protein conformation in solution. Acid denaturation followed by HSQC spectra analysis and DOSY-NMR were used to probe the early event in fibril formation. Light scattering experiments were used to infer distribution of oligomeric population. Bis-ANS fluorescence assay was used to probe the stability of protein upon acid denaturation. Fibril morphology was analysed using electron microscopy.

Chapters 6 and 7 deal with the strategy for further research work that may be conducted in the area of fibrillogenesis and the chaperone- substrates interactions.

Chapter 6:

Molecular cloning protocol of Aβ(1-40) gene as a fusion system (Maltose-binding protein-Aβ(1-40) was constructed. The construction of the expression plasmid, its expression profile and proteolytic cleavage of fusion protein were studied.

This is useful in making ^{15}N, ^{13}C-labelled Aβ(1-40) for future work on solid and solution-state NMR studies of Aβ(1-40) .

Chapter 7:

Fluoro-derivatized N-methyl maleimide (F-NEM), which modifies sulfhydryl group of cysteines, was synthesized using organo-silicon chemistry. This small molecule may be useful to allow sensitive detection of protein-protein interactions via ^{19}F NMR studies. Upon modification, formation of two stereoisomeric forms of the F-NEM-derived cysteine was observed and the stereoisomers were isolated. Using this reagent, stereoselective property of the molecular chaperone Hsp104 was explored. This reagent may also be useful to explore the interactions of various membrane proteins and higher molecular aggregates.

Chapter 8:

Combining all the systems, a conclusion regarding the fibril formation and chaperones' action is derived.

Appendix

This section contains several C-source codes for the extraction of various NMR data and their analysis.

3.0   Structural investigation of Yeast prion system: Sup35 and Hsp104

3.1   Background

Translational termination is an important process in protein synthesis. In eukaryotes, following transport of the mRNA from the nucleus to the cytoplasm, ribosomes begin translating the mRNA through a complex initiation and elongation process (*1*). The decision to terminate the translation at the stop codon is made by a translational termination complex (*2*). It consists of the translational release factors (eRF1 and eRF3) and a putative surveillance complex (Fig. 3.1b). eRF1, which mimics t-RNA, recognizes the stop codon (Fig. 3.1a). Once the eRF1-eRF3 complex occupy the aminoacyl site (A site), hydrolysis of peptidyl-tRNA occurs. Simultaneous hydrolysis of GTP on eRF3 mediates the release of completed polypeptide and dissociates both uncharged tRNA and eRF1-eRF3 complex (*3*).
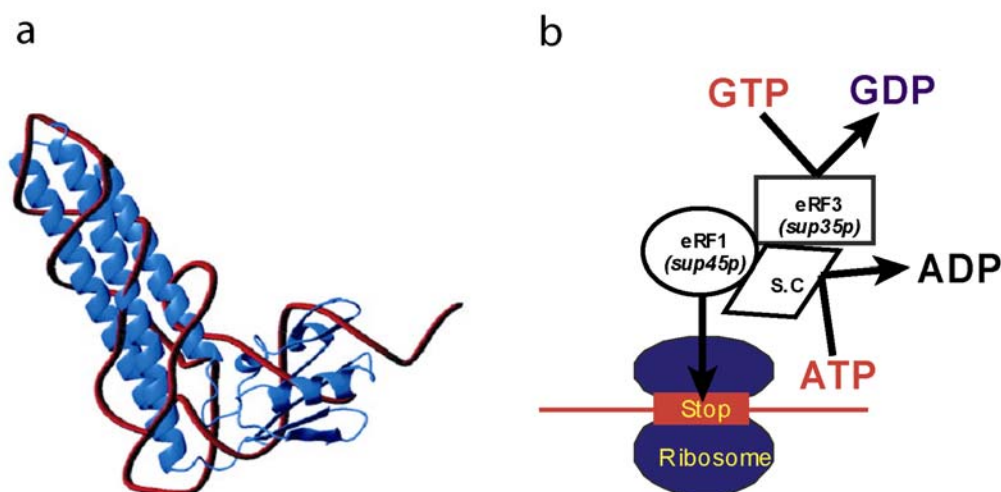
Figure 3.1. a) Superposition of eRF1(blue) and t-RNA(red) (Selmer, M., et al., Science, (1999), 286,5448)

b) Cartoon depicts the translational termination complex

In order to have efficient translational termination, the formation of eRF1- eRF3   complex should be efficient. In *Saccharomyces cerevisiae,* an eRF3 homologue is the Sup35 protein. It consists of 685 amino acids and is divided into three domains: N, M and C. The N domain, which consists of 123 amino acids, contains mostly asparagines and glutamines (*4*). This N-terminal section is known to form amyloid-like fibrils *in vitro* and is referred to the "prion domain" (*5*). It has also been shown from mutational analysis that occurrence of mutations within a core region of 22-amino acid region in prion domain (Sup35[5-26]) eradicate the prion state *in vivo (5)*. The other domains have been shown to play essential roles *in vivo*. The

highly-charged M region of Sup35, not only enhances the solubility of the protein, but also maintains the prion state inside the living system (*6*). The C terminal domain was shown to be interacting with Sup45, an eRF1 homologue in yeast system (*7*). In the amyloid state, Sup35 forms amyloidogenic aggregates. As a result, it fails to interact with its eRF1 homologue, Sup45, and hence unable to form the termination complex. Stop codons are therefore sometimes missed, producing protein with an extra segment. This phenotype is called [ψ^+] (Fig.3.2). Deletion of a heat shock protein gene, Hsp104, results in the elimination of the phenotype [ψ^+] (*8*). For mammalian prions, a similar mechanism is postulated, however, the so-called protein-X has not yet been identified (*9*).
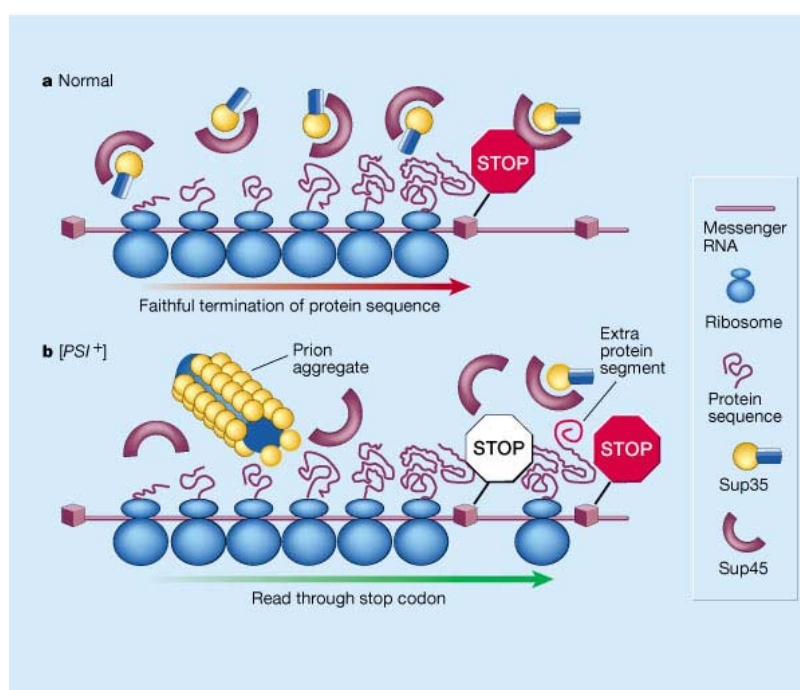
Figure 3.2. a) Sup35 protein interacts with its partners and renders efficient translational termination in cells with normal Sup35. b) In [PSI⁺] cells, the translational infidelity occurs. (Figure adapted from Partridge, L., Barton, N.H., Nature (2000), 407, 457)

Hsp104 is a member of the Clp/Hsp100 family of proteins that are heat-inducible molecular chaperones known to exist as higher order oligomers (*10*). Like other members of the Clp family of proteins, the presence of Walker-type ATP binding modulates the concentration dependent oligomerization of Hsp104 (*11*). The low resolution electron microscopy (EM) structure of Hsp104 shows that Hsp104 folds into hexameric complexes (*12*) (Fig. 3.3a). A recent crystal structure of its close homologue TClpB contains flexible loops that seem to be important for the activity of Clp family of proteins (*13*) (Fig. 3.3b).
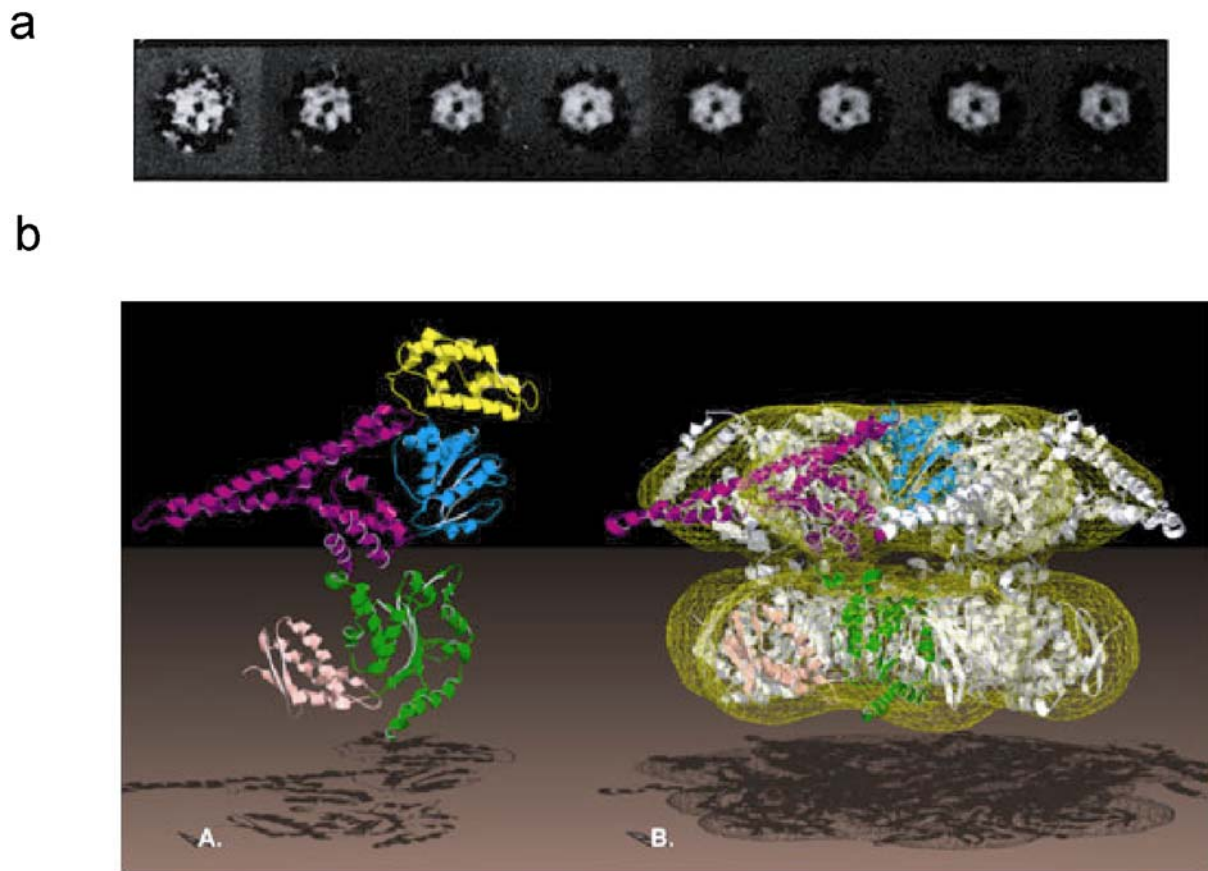
Figure 3.3. a) EM structure of Hsp104 reveals its hexameric nature. b) crystal structure of ClpB (left) and quaternary arrangement (right). The mobile loops that are important for the disaggregase activity are protruded from the upper part of the barrel.

It is known that Hsp104 is able to modulate the aggregation of other amyloid-like proteins as well (*14*). Two modes of action of Hsp104 have been portrayed so far to support the mechanistic aspect of Hsp104 in yeast prion maintenance. Lindquist and co-workers suggested that Hsp104 is necessary for the normal protein to adopt an aggregation competent state, which can be restored to the normal form by over-expression of Hsp104 alone (*15*) (Figure 3.4a). Hence, Hsp104 is believed to play a dual role, both in maintaining and rescuing the yeast prion at different concentrations. Alternatively, Ter-Avanesyan and co-workers proposed Hsp104 to act as a molecular scissor, which cleaves the fibrillar species into individual fragments and each of these oligomeric pieces can act as a seed to carry on their catalyzed polymerization (*16*) (Fig.3.4b).
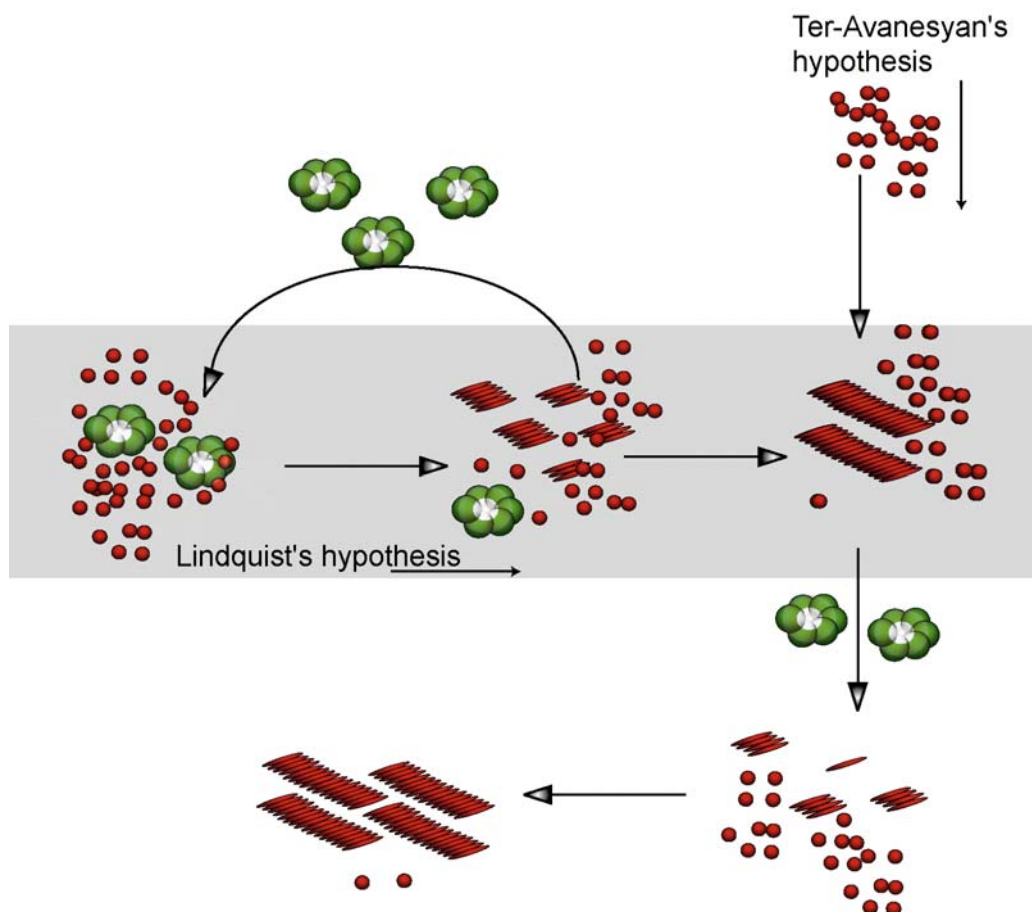
Figure 3.4. Modes of action of Hsp104. The Lindquist's hypothesis has been depicted within grey box. The Ter-Avanesyan's hypothesis has been portrayed in vertical slot.

The mechanistic aspect of the interaction of Sup35 with Hsp104 has not been addressed, though the CD experiments suggest that the Hsp104 interacts with Sup35 *in vitro* (*17*). Recent evidence from EM shows that the involvement of Hsp104 in disaggregation of Sup35[NM] (*18*).
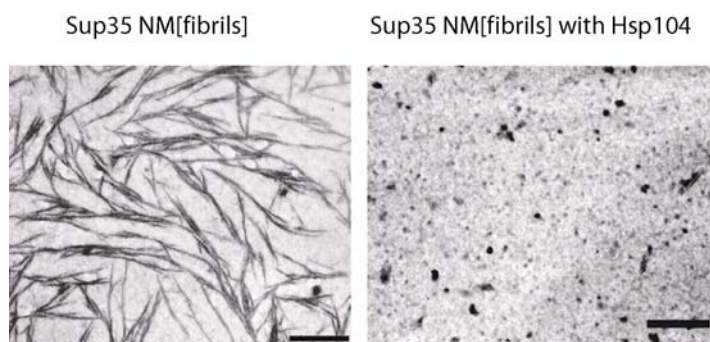
Figure 3.5. Sup35 [NM] fibrils without Hsp104 (left) and with Hsp104 (Shorter, J., Lindquist, S. Science, 2004, 304, 1793)

3.2   Aims of the project

1. Mapping the aggregation and disaggregation pathway of Sup35[5-26]

2. Structural and mechanistic understanding of the interaction of Hsp104 with soluble and fibrillar Sup35[5-26]

3. Understanding the driving force behind the disaggregation.

3.3   Material and Methods

In this project, a peptide Sup35[5-26], which corresponds to the 5-26 of Sup35, instead of the full length protein was used. The peptide was purchased from Biosource International (Camarillo, CA). The sequence of the peptide is $N^5GNNQ^{10}QNYQQ^{15}YSQNG^{20}NQQGN$. NMR experiments were performed on Bruker DMX 750 (750 MHz) and DMX 600 (600 MHz). The plasmid carrying the Hsp104 gene, was obtained from S. Walter, TU-Munich, Germany. Cibacron Blue F3GA (Affigel blue, Biorad, Germany), Resource Q (Pharmacia, Sweden), Superdex 200 (Pharmacia, Sweden) columns were used for the purification.

3.3.1   Buffers

Lysis buffer: 50mM Tris/HCl, pH 7.7, 2mM EDTA, 0.1M NaCl, protease inhibitor cocktail tablet (1tablet/50ml of buffer).

Buffer A1:   50mM Tris/HCl, pH 7.7, 2mM EDTA, 0.1M NaCl

Buffer A2:   50mM Tris/HCl, pH 7.7, 2mM EDTA, 0.3M NaCl

Buffer A3:   50mM Tris/HCl, pH 7.7, 2mM EDTA, 1M NaCl

Buffer B: 50mM Tris/HCl, pH 7.7, 2mM EDTA, 0.01% β-ME

Buffer C1:   50mM Tris/HCl, pH 8, 2mM EDTA, 0.1% β-ME

Buffer C2:   50mM Tris/HCl, pH 8, 2mM EDTA, 0.1% β-ME, 1M NaCl

Buffer D: 50mM Tris/HCl, pH 7.7, 1mM DTT

3.3.2 Expression and purification of Hsp104

Expression and purification was performed as described in the literature (*19*). In short, the purification protocol is described below:

A derivative of W303 yeast *Saccharomyces cerevisiae* strain, A544, carrying a plasmid encoding Hsp104 with Glycoprotein α-hormone promoter which selects His drop-out YNB medium, was used. The His drop-out mix was prepared as described (*20*). Large scale expression was performed in the fermenter under constant pressure (pO$_2$=60%) at pH 6.8. Eight litres of His medium without glucose was inoculated with 200ml of overnight culture (rich YPD medium). The medium was stirred at 300 rpm and fermentation was carried out at 30° C for 6 hours. Cells were harvested at centrifugation of 6000 rpm and re-suspended in lysis buffer.

Cells were cracked with cell disruptor at 2.4 kbar pressure and subjected to ultrasound. The cell debris were removed by means of centrifugation at 20,000 rpm for 45 min. The purification steps were performed at 4° C . The supernatant was filtered through 0.45 μM filter and loaded on affigel blue column (specific for nucleotide binding proteins) pre-equilibrated with Buffer A1. Buffers A2 and A3 were used for the washing and elution, respectively. The SDS-gel analysis was performed and the fractions were pooled and dialysed against 2 litre buffer B (Fig. 3.6a). The pooled fractions were subjected to ion-exchange chromatography on resource-Q column which was pre-equilibrated with buffer C1. Buffer C2 was used to elute the protein using 0-50% gradient. The fractions were analysed by SDS-PAGE and were pooled for the final purification (Fig. 3.6b).

The final purification step was performed using Superdex 200 gel filtration column. The pooled fractions from the second purification step was loaded onto the gel filtration column which was pre-equilibrated with the buffer D. The protein was eluted using buffer D and the fractions were analysed by SDS-PAGE . More than 95% of pure Hsp104 was obtained. (Fig. 3.6c).
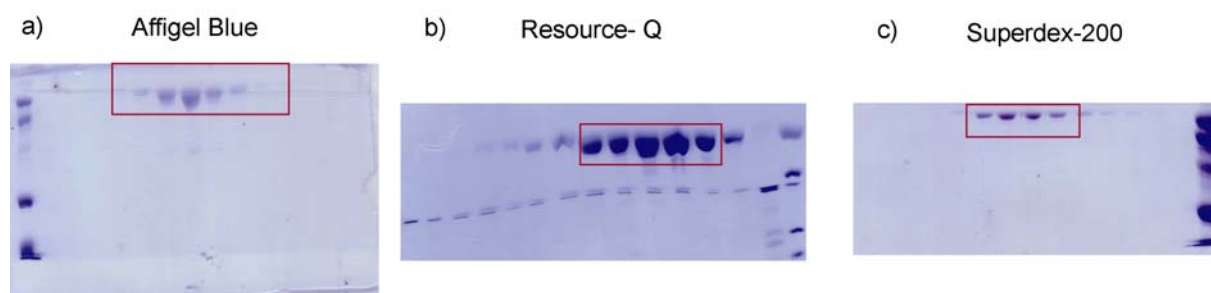
Figure 3.6 steps purification of Hsp104. Pooled fractions are highlighted

a) Dye affinity chromatography b) Anion exchange chromatography c) size exclusion chromatography

3.3.3 Real time DOSY (diffusion ordered spectroscopy) NMR experiments

The real time DOSY experiments were carried out by continuously monitoring the average molecular weight of Sup35[5-26], in order to determine the change in average molecular weight of Sup35[5-26] upon interaction with Hsp104 (*21, 22*). Catalytic amounts of Hsp104 were used, so the involvement of Hsp104 on the average molecular weight of free Sup35[5-26] was considered negligible. The buffer signal (trisodium glycerate, M_r = 161 Da) was used as internal reference to account for viscosity effects in the analysis of the DOSY NMR measurements. Experiments were carried out at constant temperature of either 27 °C or 12 °C depending upon the experiments. The average molecular weight of the oligomers can be estimated by encoding the diffusion of a molecule in a gradient echo. The measured diffusion constant *D* is related via the Stokes-Einstein relation

$$D = \frac{k_B T}{6\pi\eta \, F \, r_H}$$

to the hydrodynamic radius r_H of the molecule, and thus, to the molecular weight at a given viscosity η of the solution. k_B denote Boltzmann's constant, *T* the absolute temperature and *F* the dimensionless Perrin factor. Only molecules that are not diffusing along a given axis are detected. The decay of magnetization can be analyzed in an analytical way via

$$\frac{I}{I_0} = \exp\left\{-D\left(\Delta - \frac{\delta}{3}\right)q^2\right\}$$

where

$$q = \gamma\delta g$$

Δ refers to the separation of the gradient echo, δ to the duration of the gradient, γ to the magnetogyric ratio of the nucleus, and *g* to the strength of the gradient. The shape of a molecule has an impact on the diffusion constant. This is taken into account by the Perrin factor $F = f/f_0$ which defines the ratio of the friction coefficient of the molecule (*f*) to that of a hard sphere (*f₀*) with equivalent mass and partial specific volume (*23*). In the present study, the Perrin factor *F* has been always set to 1. This is in agreement with X-ray and electron microscopic data of model peptides. There it is found that polyglutamine peptides adopt a cylindrical β-helix with a diameter of 21.8 Å (*24*). In addition, it is reported that the hydrodynamic radii of native and unfolded peptides consisting of 22 amino acids are approximately identical (11.6-12.8 Å) (*25*) .

In the experiment, the duration of the gradients was set to 3.0 ms. A delay of 30.0 ms was employed to allow for diffusion between the dephasing and rephasing gradient. Gradient

strengths are always given in $(\%)^2$ with respect to maximum gradient strength ($100\% \equiv 35$ Gauss/cm). The experimental time for each pseudo-2D DOSY experiment amounted to approximately 2.3 h. But, the disaggregation time was approx. 48 hrs and hence, approx. 20 pseudo 2D (48/2.5 hrs) experiments were executed. The decay of magnetization during the measurement of 2.5 hours was also taken into account using piecewise discontinuity method, by recalculating DOSY intensities from the reference 1D experiments which are recorded in an interleaved fashion. As a function of time, the change in diffusion coefficient is monitored via the change in slope of the signal decay of experiments at different time points. The measured slope from the straight line (q^2 versus $\ln(I/I_0)$) plot was proportional to the diffusion coefficient,

$$\ln\left(\frac{I}{I_0}\right) = \left\{-Dkq^2\right\}.$$

Using the internal standard method, the average molecular weight of the oligomers at any time was calculated. The molecular weight is inversely proportional to the cubic power of the diffusion coefficient, i.e., $M_t \propto 1/D_t^3$ (26). D_t represents the diffusion coefficient at any time. The average molecular weight is then,

$$M_{pep}/M_{ref} = \left(D_{ref}^3/D_{pep}^3\right) = \left([Slope]_{ref}^3/[Slope]_{pep}^3\right)$$

The pseudo-2D files were evaluated and the individual slopes for reference and the peptide was computed using a self-written C source code (Ref. Appendix). Slopes were calculated using both analysis of integral and intensity values. Corresponding noise values were also included for each and every peak of the respective dataset.

3.3.4 Real time Saturation Transfer Difference (STD) NMR spectroscopy

Saturation transfer difference (STD) is based on magnetization transfer by protein signal saturation and its relayed effect to the ligand. Macromolecules have a large network of protons that are tightly coupled by dipole-dipole interactions. Saturation of a single protein resonance can result in a rapid spread of the saturation over the entire protein if spin diffusion within the protein is efficient (usually for protein >20 kDa). During the saturation period, progressive saturation is transferred from the protein to the ligand protons if the ligand binds to the target. The ligand protons nearest to the protein should be saturated to the highest degree and therefore, have the strongest signal in the STD spectrum. The ligand protons, that are further from the target surface will be saturated to a lower degree and their STD intensities will be weaker. Therefore, the degree of saturation of individual ligand protons reflect their

proximities to the protein surface and can be used as a method to describe the target-ligand interactions (Fig. 3.7) (*27*).
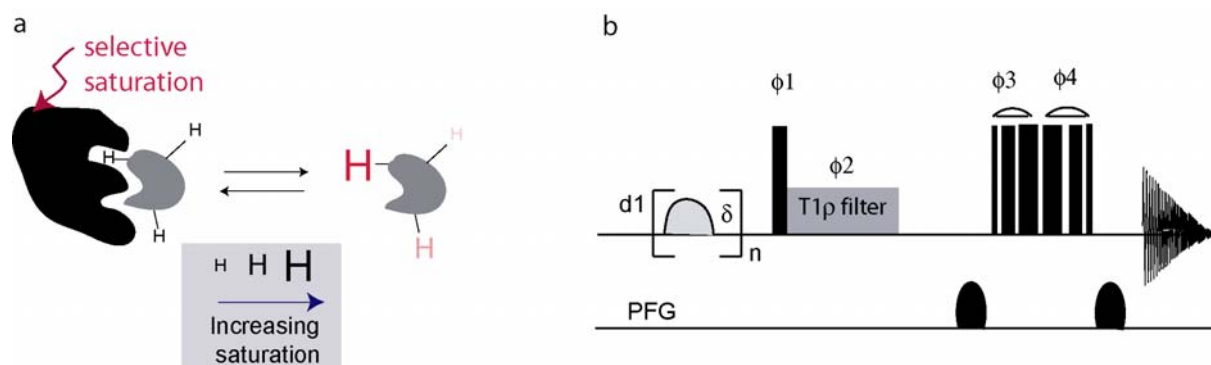
Figure 3.7 a) Principle of Saturation transfer difference experiment. b) In the pulse programme, a train of gausian shaped pulses are used for the saturation. The phase cycles are used for the difference spectrum.

STD was successfully used in the past to screen compound mixtures for binding to a receptor protein(*28*). Furthermore, it was used to characterize interactions between membrane channel proteins and neurotoxic peptides (*29*). This way it is possible to identify the chemical groups involved in ligand binding. The STD amplification factor, α is defined as

$$\alpha = \frac{I_0 - I_{sat}}{I_0} * c$$

where I_0 and I_{sat} correspond to the intensity of the signal during off-resonance and on-resonance irradiation, respectively. c refers to the relative concentration of the ligand with respect to the enzyme.

Assuming a ligand L and protein E is in equilibrium with the free (L and E for ligand and protein, respectively) and bound (L' and E' for bound ligand and protein, respectively).

$$L + E \leftrightarrow L'E'$$
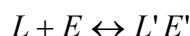
Also, the protein protons can be subdivided into two classes, E1(E1') and E2(E2'), where E2 and E2' stand for protons that experience RF irradiation directly (aromatic/selected CH_3) E1 and E1' stand for all the remaining protons in the protein in its free and bound form, respectively. Some of these E1 and E1' may experience considerable saturation indirectly through spin diffusion due to their spatial proximity to the irradiated E2 and E2' protons or weakly saturation if the spin diffusion is not efficient. Thus in the model, the saturation

originates from the E2 and E2' and spreads to the other protons through cross-correlated and exchange (through free and bound form of E). The observable magnetization in an intermolecular NOE experiment, in which the E2 and E2' protons are instantaneously saturated by RF pulses is given by

$$I(t) = I_0 + [1 - \exp\{-R+K\}t](R+K)^{-1}Q$$

where t is time period for which the protons remain saturated. This expression is a solution of the homogenous set of differential equations obtained from standard equations of motions for the magnetizations coupled by dipolar and chemical exchange processes after setting the magnetizations of saturated receptor protons to zero (E2 and E2'). K is the kinetic matrix that contains equilibrium constant terms. R is a generalized relaxation matrix composed of rate matrices for the free and bound states .

$$R = \begin{bmatrix} R_F & 0 \\ 0 & R_B \end{bmatrix}, \text{Where } R_F = \begin{bmatrix} R_L & 0 \\ 0 & R_E \end{bmatrix} \text{ and } R_B = \begin{bmatrix} R_{L'} & R_{C'} \\ R_{C'}^T & R_{E'} \end{bmatrix}$$

in which, R_F and R_B are the relaxation rate for the free and bound form. R_L and R_E are the relaxation rate of free ligand and protein, respectively. $R_{L'}$ and $R_{E'}$ are the relaxation rate for the complexed form of the ligand and protein, respectively. The absence of cross-correlated relaxation in free states of protein and ligand denoted by the zero off-diagonal elements of R_F matrices whereas, in the bound for $R_{c'}$ denotes the intermolecular dipolar cross correlated relaxation in the complex and $R_{C'}^T$ is its transpose. All the relaxation matrices are included in the $R_{L'}$ and $R_{E'}$ matrices, which include the diagonal elements associated with the intermolecular dipolar cross relaxation between the ligand and the enzyme, as well as the complete relaxation matrix elements for the intramolecular relaxation within the bound forms of the ligand and the enzyme. If the protein and ligand are in fast exchange on both relaxation and chemical shift time scale then the observed signal intensity I_{sat} is given by the following matrix (*30*).

$$\begin{vmatrix} (I_L + I_L^{'}) \\ (I_{E1} + I_{E1'}) \end{vmatrix} \approx \begin{vmatrix} (I_{L0} + I_{L'0}^{'}) \\ (I_{E10} + I_{E1'0}) \end{vmatrix} + (1 - \langle R \rangle t)t \times \begin{vmatrix} (R_{L'E2'}{'}I_{E2'0}) \\ (R_{E1E2}I_{E20} + R_{E1'E2'}I_{E2'0}) \end{vmatrix} + H(\langle R \rangle, t)$$

The initial portion of the ligand signal in the STD NMR experiment (irrespective of the exchange term) is sensitive to the direct transfer of saturation from the saturated protein proton to the ligand proton in the complex, as reflected by the $R_{L'E2'}I_{E2'0}$ term. For longer times, the STD NMR intensity spectrum displays additional effects from the indirect

saturation transfer from the E1' set of protons in the bound state, as well as exchange mediated effects ($R_{E1E2}I_{E20}$ term). For accurate epitope mapping, the knowledge of the longitudinal relaxation rate may be helpful, as it has been shown to have an influence in the transfer efficiency (*31*) .

Experimentally, a train of 300 square pulses of 12 msec duration (with 1.0 msec gap between the pulses) was applied in order to saturate the ^{1}H(methyl) resonances of Hsp104. Even (on-resonance irradiation at –0.02 ppm) and odd scans (off-resonance irradiation at +30.0 ppm) were recorded subsequently and subtracted from each other by the incrementation of receiver phase. 512 scans were accumulated at each time point, resulting in a total experimental time of 45 min per time point. STD experiments was always adjusted in the first place in a reference experiment on a sample containing no Hsp104 in order to minimize artifacts. The amount of the relative concentration of Sup35[5-26] was always taken into account in the evaluation by recording 1D reference spectra in parallel with the STD experiments.
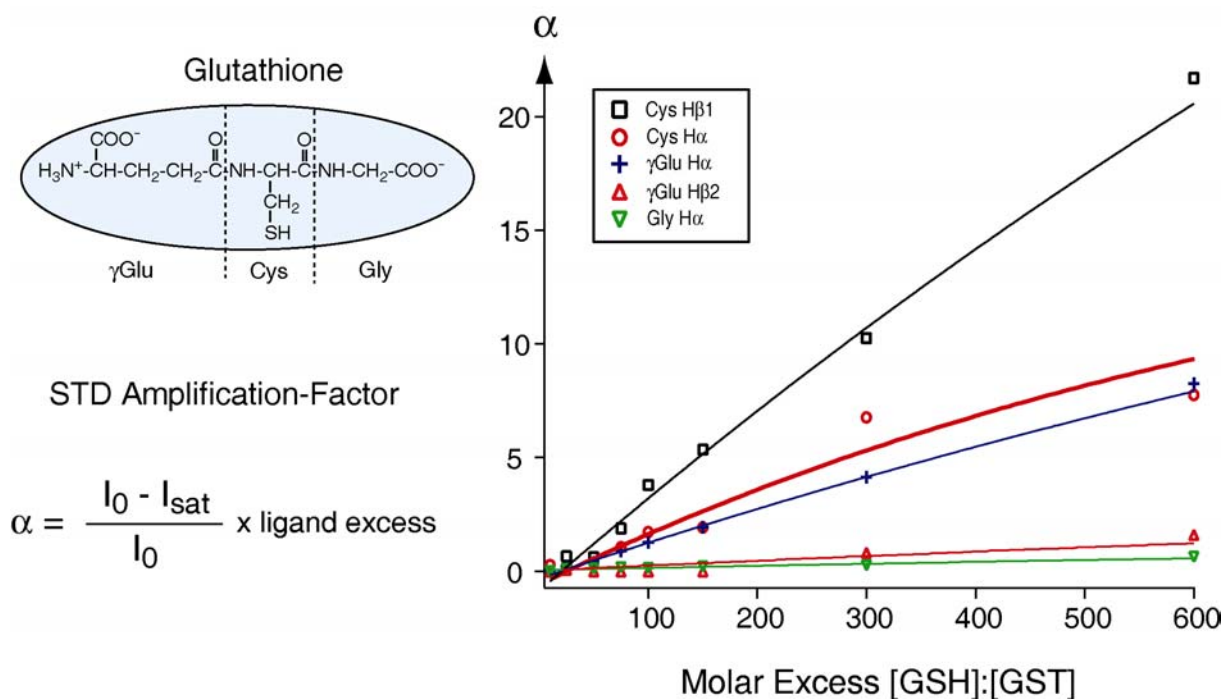
Figure 3.8. STD amplification factor as a function of concentration for Glutathione and Glutathione-S-transferase system.

STD experiments were first calibrated using reduced-Glutathione (GSH) and Glutathione-S-transferase (GST) as a model system. The build-up of the STD amplification factor was defined as the ratio of the resonance intensities between on and off resonance irradiation of the enzyme times the molar excess of the ligand, is exponentially growing with the molar

ratio of ligand to enzyme concentration. Increased STD build-up rates correspond to group epitopes which are closer in space with respect to the enzyme binding site. In the setup system, the Hβ resonances of Cys show the fastest build-up which was in agreement with interaction models for GSH and GST (Fig. 3.8).

In standard STD experiments, the magnitude of the amplification factor is directly correlated to the concentration of the ligand (at a given enzyme concentration). Here, we exploit the aggregation behaviour to obtain the concentration dependence of the ligand in an indirect way. If the ligand consists of multiple inter-converting species with different affinities, as we have shown for Sup35[5-26] via DOSY NMR experiments (*vide supra*), a change in the relative population of these species during the time course of the experiment will result in a corresponding change in STD signals. Due to fast chemical exchange, the different conformers exhibit the same isotropic chemical shift. A decrease of the population of one species is associated with an increase of the second. This corresponds to two STD curves which exhibit a relative minimum after superposition. No change in the STD signal is expected for a ligand that always shows the same binding affinity. The real time STD spectra were quantitatively analysed using my compiled C source code. (Appendix) Noise ratio was also taken into account. Error propagation were also included.

3.3.5   Heteronuclear (^{1}H,^{13}C) single quantum coherence (HSQC) NMR spectroscopy

All NMR spectra were acquired at 27°C using a Bruker DMX 750 NMR equipped with a triple-resonance probe and triple-axis self-shielded gradient coils. Natural abundance ^{1}H{^{13}C}HSQC experiments were performed with gradient selection and sensitivity enhancement employing water flip-back pulses to minimize saturation of exchanging protons. A typical experiment comprised 512t_1 increments with 1024 transients of 2048 points each, covering a spectral width of 15205 and 8012 Hz in f_1 and f_2, respectively. ^{1}H Chemical shift calibration was done using Tetramethylsilane (TMS) as a standard. ^{13}C chemical shift calibration was achieved using the hydroxymethyl resonance in 2-amino-2-hydroxymethyl-1,3-propanediol (Tris buffer) (SDBS databank, National Institute of Advanced Industrial Science and Technology, Japan). Secondary structural analysis of the ^{13}C chemical shifts was accomplished by comparison of the experimental chemical shifts published in the BioMagResBank , University of Wisconsin, USA.

3.3.6 Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded on Jasco J-715 spectropolarimeter with a PTC 343 peltier unit. The experiments were carried out in quartz cuvettes of 0.1 cm path length. Far-UV spectra were recorded within the bandwidth of 185-250 nm. All spectra were buffer-corrected. In order to approximately reproduce the NMR conditions, a final concentration of 300 μM of Sup35[5-26] was employed. To study the influence of Hsp104, the peptide sample was solubilized in a Hsp104 solution using a molar ratio of 1:25 of [Hsp104]:[Sup35[5-26]], yielding again a peptide concentration of 300 μM. After four days, the mixture solution was filtered using centricon centrifugal filter units (Millipore GmbH, Germany) with a MWCO of 3.5 kDa in order to eliminate Hsp104. The filtrate was subjected to CD spectroscopy, in order to investigate the secondary structural propensity of the released Sup35[5-26] monomer. CD spectral analysis was performed using the software CDSSTR

3.3.7 Equilibrium dialysis experiments
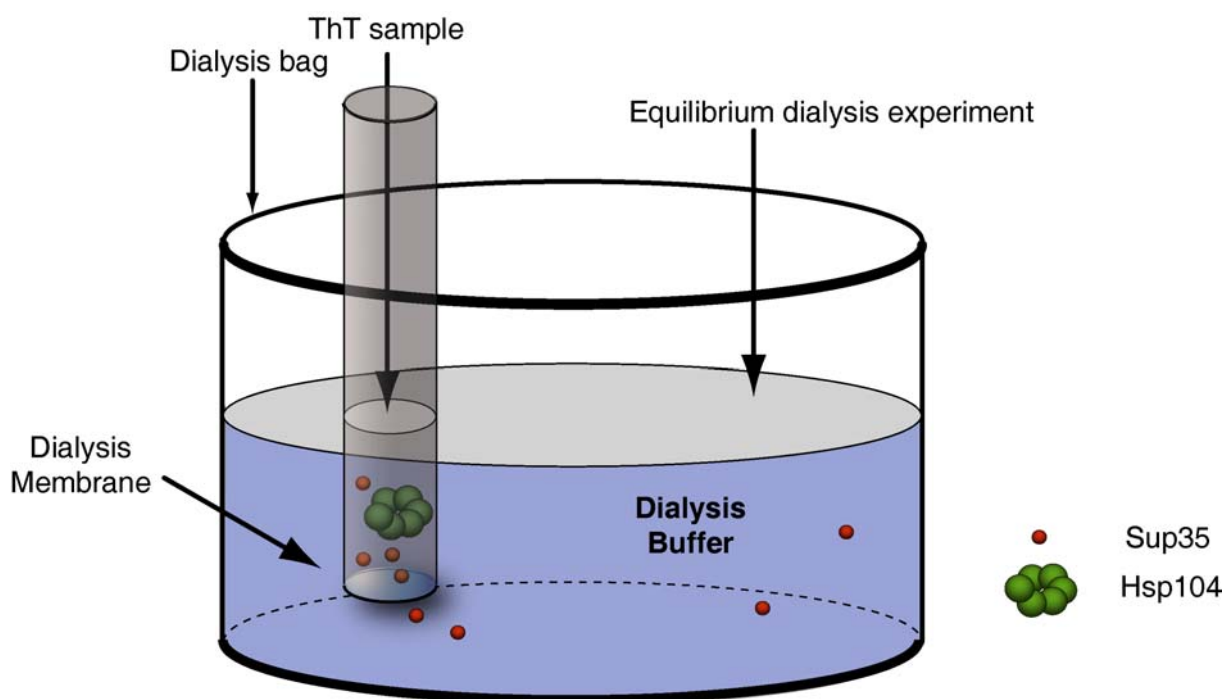
Figure 3.9. Schematic picture of combined Thioflavin-T and Equilibrium dialysis experiment

The equilibrium dialysis experiment was set up using ultra performance float-a-lyzer dialysis tubes (KMF Laborchemie Handels GmbH, Buisdorf, Germany) of sample volume 500 μl with varying MWCO membranes (3.5 kDa, 10 kDa and 25 kDa) depending upon the

experimental conditions. Dialysis bags were filled either with freshly solubilized peptide or with pre-formed aggregates. A volume of 10 ml of dialysis buffer was kept outside the dialysis bag in all cases. Dialysis buffer solution for all the experiments were continuously stirred under the same conditions to allow for rapid equilibration. The amount of fibrils formed within the dialysis bag was quantified by ThT fluorescence (Fig.3.9). The amount of peptide diffusing out of the chamber was quantified by the intensity of the tyrosine fluorescence. HPLC purification was employed prior to quantification of the eluted peptide using a C-18 reverse phase column, in order to suppress artefacts originating from ATP/ADP. At regular time intervals, 120 μl of dialysis buffer was withdrawn and stored separately for quantification. Quantification was done in all cases by injecting 100 μl of the dialysis buffer samples into the HPLC PU-1580 system (Jasco, Germany) and detecting the tyrosine fluorescence intensity of Sup35[5-26] using a FP 1520-S fluorescence detector (Jasco, Germany). Therefore, tyrosine fluorescence excitation and emission wavelength of 275 nm and 303 nm, respectively, were employed. Calibration of the fluorescence intensities was achieved upon injection of 100 μl of buffer containing 1 μM of the peptide. Tyrosine fluorescence intensities were normalized to 100%, corresponding to complete equilibration between dialysis solution and the reservoir. In order to study the influence of ATP on disaggregation process, 10mM ATP was added in the dialysis bag as well as in the dialysis buffer. 50 mM phosphate, pH 7.4 was used as a buffer throughout the experiments.

3.3.8 Thioflavin-T (ThT) assay

ThT based fluorometric analyses were performed in a FluoroMax I (Spex, Edison, USA) fluorescence spectrophotometer. During the measurements the excitation wavelength was set to 446 nm with a spectral bandwidth of 5 nm. The emission spectra were obtained by scanning from 460 nm to 540 nm. Aggregation was monitored by taking samples of 2 μl out of the dialysis bag at regular time intervals. Prior to the fluorometric analysis, 50 μl of the 500 μM ThT stock solution was mixed with the Sup35[5-26] peptide yielding a molar mixing ratio of 10:1 of ThT with respect to Sup35[5-26] concentration.

3.4 Results and Discussion

3.4.1 Aggregation of Sup35[5-26] as monitored by 1D-^{1}H NMR spectra

Upon dissolution of freeze-dried Sup35[5-26] in buffer, the peptide aggregated quickly, as shown by the rapid decay of observable ^{1}H NMR resonances (Fig 3.10a, Blue). Only signals originating from soluble peptide molecules were visible in the spectrum. At a peptide concentration of 1.5 mM and T = 27 °C, the characteristic time τ_{agg}, after which 50 % of the peptide became insoluble, was found to be approximately 80 min. At 12 °C, aggregation was too fast to be monitored with this experimental setup. A possible explanation for this temperature-dependent behaviour is that the peptide consists of an ensemble of species with differing aggregation probabilities. At higher temperature, this distribution was shifted to the forms that have a smaller tendency to aggregate. This idea was supported by the observation that aggregated peptide could be re-solubilized by heating the sample to 50 °C (Fig. 3.10b). Also, no detectable conformational changes occurred prior to protein aggregation.
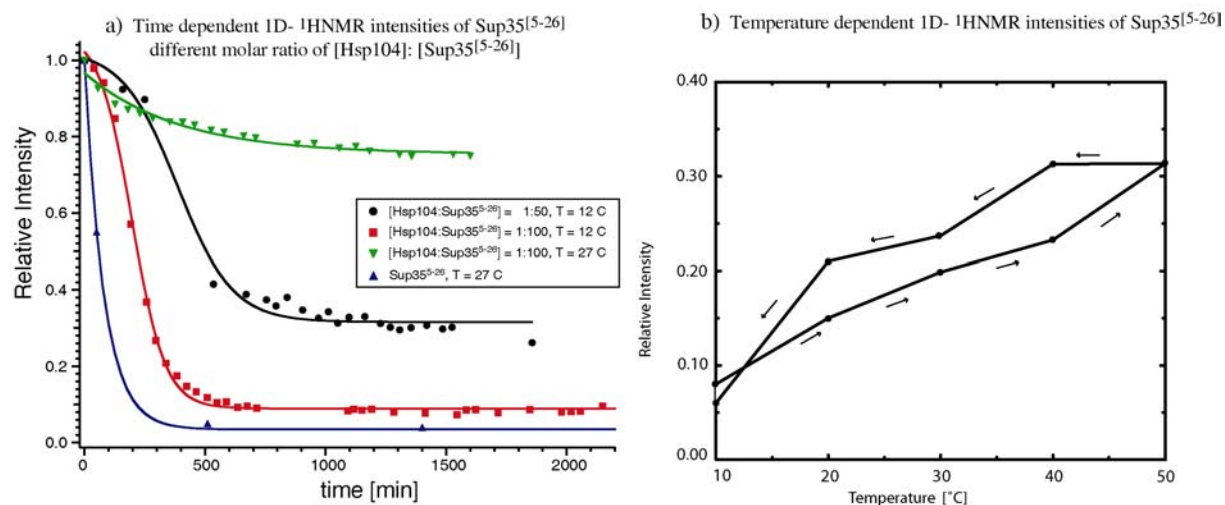
Figure 3.10. 1D proton NMR intensities as a function of time (Left) and Temperature (right)

Addition of Hsp104 slowed down aggregation (Fig. 3.10a, Red). Furthermore, aggregation was no longer quantitative in the presence of Hsp104. At a molar ratio of Hsp104 to Sup35[5-26] of 1:50 (monomer : monomer, T = 12 °C), ~35 % of the peptide remained in solution after incubating the sample for 800 min (Fig. 3.10a, Black). τ_{agg} was determined to be in the range of 400 min. NMR resonances of protons stemming from Hsp104 were not observable in these experiments due to the low concentration and the large molecular weight of the hexameric

Hsp104 complex. At 27 °C and a molar ratio of Hsp104 to Sup35[5-26] of 1:25, the peptide aggregation was drastically reduced (Fig. 3.10a, Green).

3.4.2   Hsp104 induces conformational changes in Sup35[5-26]-^{1}H –1D NMR analysis
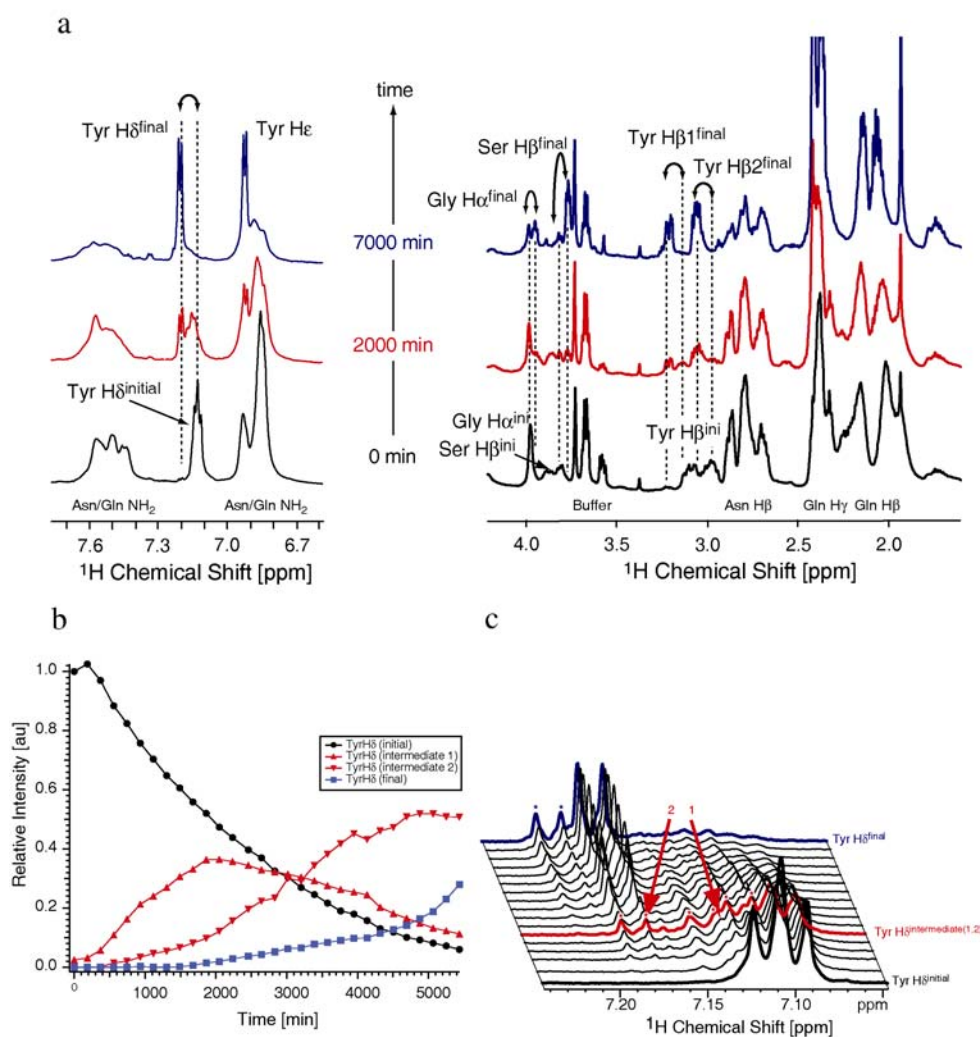
Figure 3.11. a) At a molar ratio of [Hsp104]:[Sup35[5-26]]=1:25, conformational change in proton NMR observed. b) Population of different conformers as a function of time. c) Tyrosine aromatic region clearly displayed conformational change.

To obtain a more detailed view of the interaction between Hsp104 and Sup35[5-26], 1D ^{1}H NMR spectra were recorded at 27 °C and a molar ratio of Hsp104 to Sup35[5-26] of 1:25.

Under these conditions, the peptide apparently can no longer aggregate (Fig. 3.10a, Green). The intensities of non-exchangeable protons were constant throughout the experiment. After 300 min, however, new signals appeared in the spectrum that cannot be found in the reference spectrum recorded in the absence of Hsp104 or in the experiments recorded at lower temperature (12 °C) in the presence of Hsp104 (Fig. 3.11).

MALDI-MS was carried out to exclude the possibility that the peptide was degraded or chemically modified. The mass of the peptide before and after the NMR experiment was found to be unchanged (Fig. 3.12 and Fig. 3.13). The new signals – highlighted in Fig. 3.11 by dashed lines – must therefore reflect a change in the conformation of Sup35[5-26]. The spectral changes were illustrated in Fig. 3.11b, which showed a detailed view of the Tyr aromatic region. The decay of the initial resonance line (black) was accompanied by the rise of two intermediate species (red). At later times, a fourth conformer (blue) accumulated. The absolute intensities of the four different species are shown as a function of time in Fig. 3.11a.
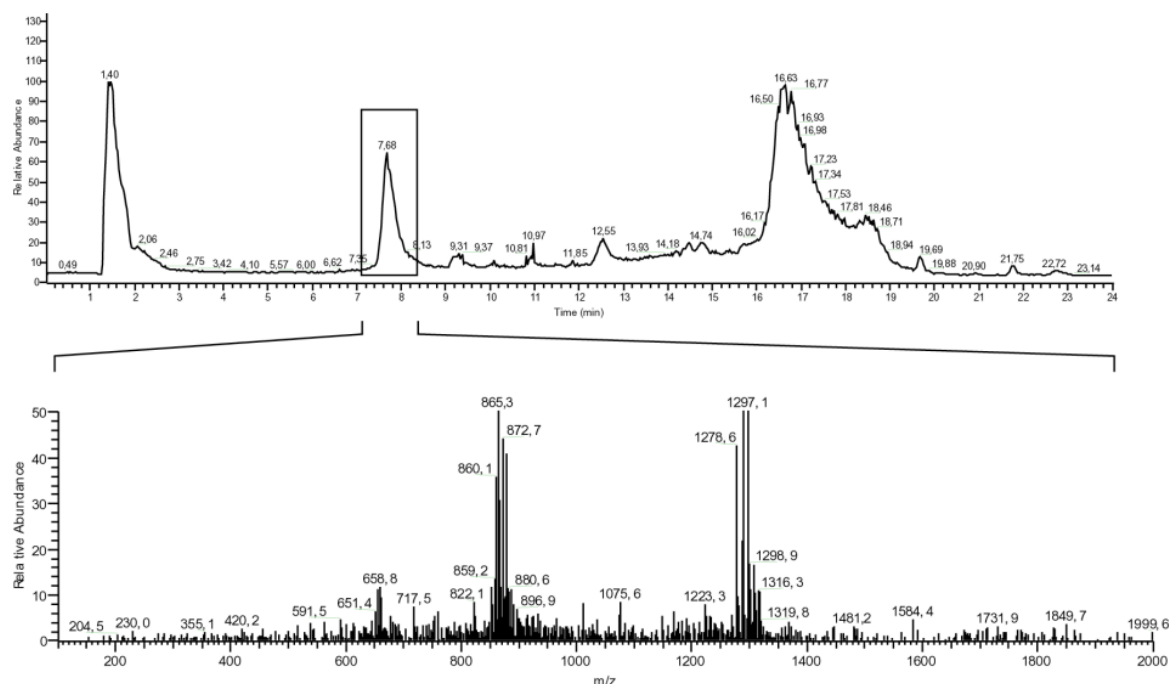
Figure 3.12. HPLC-MS trace for the monomeric SUP35[5-26] confirmed that the peptide did not degrade or modify during the course of the interaction with HSP104. The mass represent Z+ and Z(2+) states
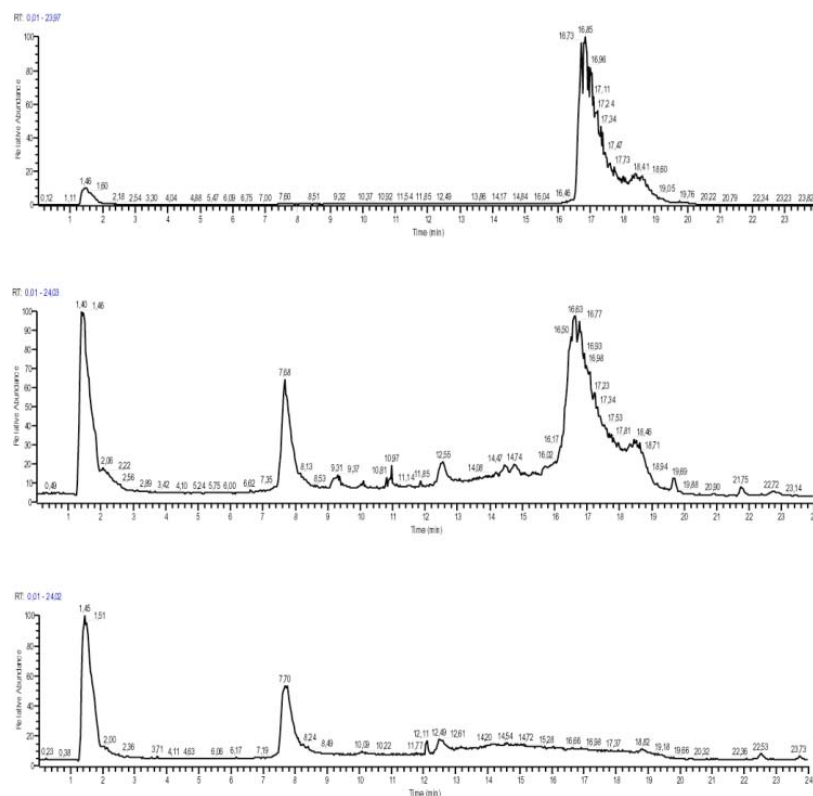
Figure 3.13. HPLC trace for the Sup35 alone (bottom), Hsp104(top) and sup35:HSP104(middle)

Sup35[5-26] contained two tyrosines. The scalar coupling between Hδ and Hε caused a splitting of the signal into a doublet. For clarity, the intensities of the doublets for intermediate 2 and the final conformer were scaled down by a factor of two, to take into account the overlap of two resonance lines. The decay of the initial species occurred with a time constant of 1600 min. Concomitantly, intermediate-1 became populated and reached its maximum concentration after ~2000 min. After 5000 min, the prevailing species of Sup35[5-26] was intermediate-2. At this point, the final conformer was only populated with ca. 25 %. 100 % population of this final Sup35[5-26] conformer was found only after 7 days.

This change in proton environment may indicate either major or minor conformational change induced by Hsp104 as a function of time. Notably, this conformational change was not observed without Hsp104 presence.

3.4.3 Various oligomeric species of Sup35[5-26] observed from DOSY NMR Analysis

The differences in the aggregation behaviour at 12 °C and at 27 °C (Fig.3.10), and the conformational changes observed in the presence of Hsp104 (Fig. 3.11) indicated that Sup35[5-26] assembled into variety of conformers with differing aggregation properties and affinities for Hsp104. To further characterize this ensemble, the average molecular weight of Sup35[5-26] was determined with DOSY NMR as described in material and methods .
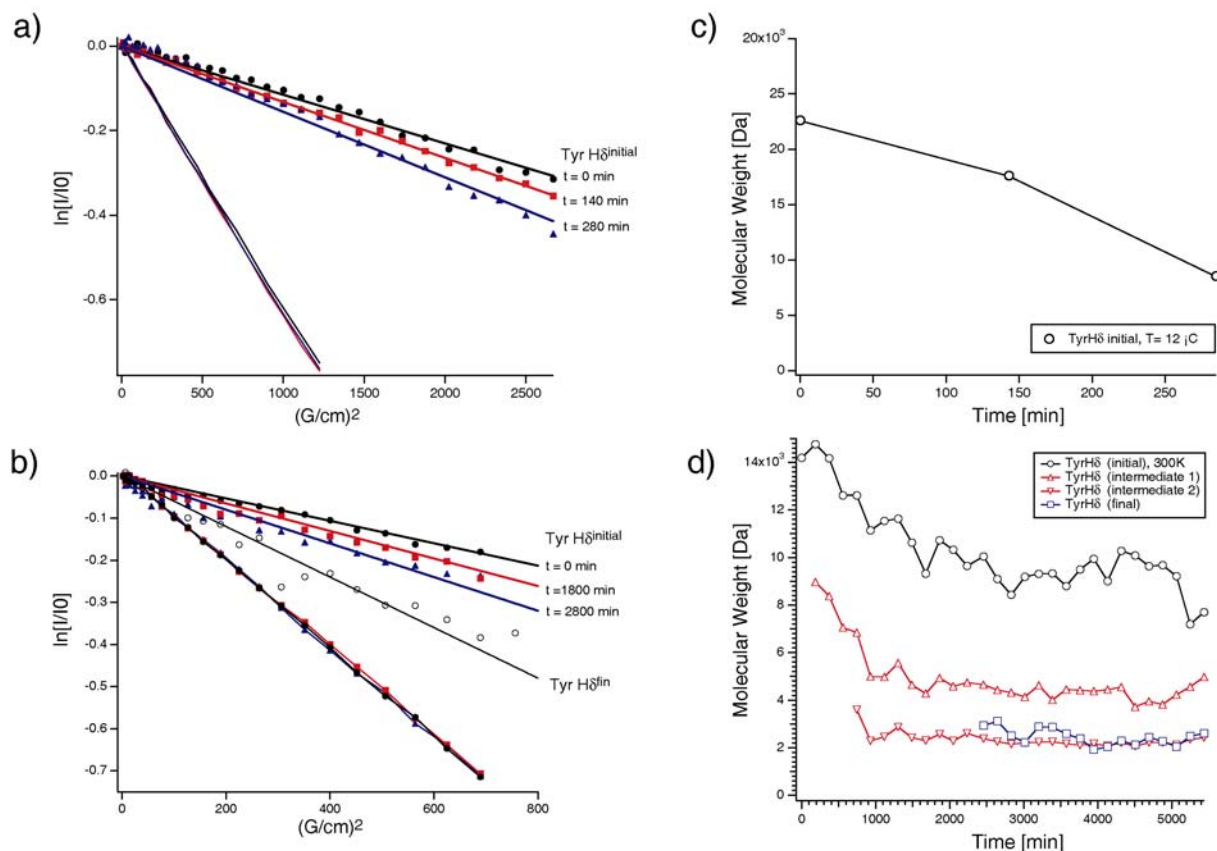
Figure 3.14. NMR diffusion data for Sup35[5-26] . The linearized semilog plot (a,b) and calculated average molecular mass (c,d). Every point in c,d is a pseudo 2D experiment. DOSY data in presence of Hsp104 at a molar ratio of protein: peptide, 1:50 at T=12°C(a,c) and 1:25 at T=27°C(b,d)

Figures 3.14 a-d show the time dependence of the diffusion constants for Sup35[5-26] at 12 °C (a,c) and at 27 °C (b,d), both in the presence of Hsp104. It is apparent that the average molecular weight of the Sup35[5-26] complex changed due to a shift in equilibrium between different oligomeric states. At 12 °C (Figures 3.14a, c), the molecular weight decreased from approx. 22 kDa (t = 0 min; I = I_0) to 10 kDa (t = 280 min; I = 0.8-0.4 I_0). Under these conditions, ~65 % of the peptide molecules precipitated (cf. Fig. 3.10). Apparently, the initial distribution of conformers were centred around octameric species. Upon longer incubation, the majority of the molecules formed insoluble aggregates, whereas a fraction dissociates into smaller species that stayed in solution.

At a molar ratio of 1:25 for [Hsp104]:[Sup35[5-26]] at T = 27 °C (cf. Fig. 3.14b, d), different oligomeric behaviour of the peptide was observed. Initially, the average molecular weight of Sup35[5-26] was ~14 kDa, corresponding to a population of a predominantly hexameric species. With longer incubation, this species dissociated into smaller oligomeric states, and the

64

apparent molecular weight decreased (~8 kDa). Importantly, the chemical shift of the Tyr Hδ proton remained the same, indicating that the environment of this proton did not change when the tetrameric species gets dissociated. From the spectra displayed in Figure 3.11, slow conversion of Sup35[5-26] into new species was observed, indicated by changes in the chemical shifts. Using signals specific for the new species, their molecular weight could be determined in the same DOSY experiment. Intermediate-1 was mainly tetrameric at the beginning of the experiment. Upon incubation with Hsp104, the equilibrium was shifted to smaller oligomeric states and Sup35[5-26] was slowly converted into a dimeric species. Intermediate-2 as well as the final conformer were monomeric throughout the experiment (Fig 3.14d).

These results suggest that at 27 °C, the Sup35[5-26] peptide adopted a range of oligomeric states that was centred around hexameric species. Upon interaction with Hsp104 these hexamers slowly dissociated into structurally closely related tetramers. Subsequently, these tetramers dissociate into monomers in a two step process, which was accompanied by characteristic changes in the 1D ^{1}H NMR spectrum.

From the above experiments, it is clear that, at 12°C the seed for the polymerization was rapid formation of octameric Sup35[5-26] upon dissolution. Though aggregation cannot be prevented in absence of Hsp104, it can be controlled by shifting the octamers to hexamer by means of shifting the temperature to 27°C. Now, it is important to know what is the driving force for the disaggregation mechanism of Hsp104. To address this issue differentially binding affinity of the Sup35[5-26] oligomers was measured.

3.4.4 Binding of Sup35[5-26] to Hsp104 observed by 1D-STD experiments

In order to identify the chemical groups of Sup35 that interact with Hsp104, saturation transfer difference (STD) experiments were carried out. The time dependence of the STD amplification factor for Sup35[5-26] in the presence of Hsp104 at a molar ratio of 50:1 at 12 °C is shown in Figure 3.15.

No STD signal was observed in the beginning, indicating that the initial ensemble of Sup35[5-26] species showed no interaction with Hsp104. Taking into account the results of the DOSY experiment, this observation suggests that octameric Sup35[5-26] did not bind to Hsp104. With further incubation, however, STD signals were increasing monotonously, suggesting that species were formed that bind to Hsp104. Side-chain resonances of Asn and Gln, as well as Tyr showed fastest STD build-up.

a)

b)
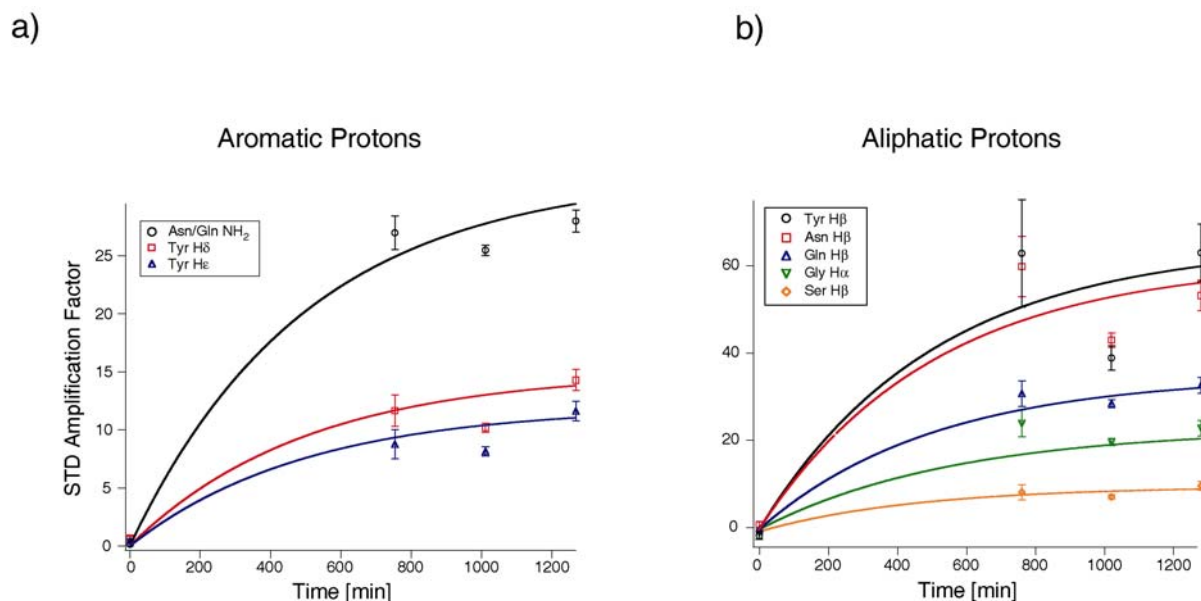
Figure 3.15. Time dependence of the STD signal for the Sup35[5-26] in the presence of Hsp104 for the molar ratio of [Hsp104]:[Sup35[5-26]]=1:50 at T=12°C

This may be expected, since mutation studies show that these residues are essential for prion propagation and interaction with Hsp104. On average, the time constant of the process monitored in the STD experiment was similar to the time constant observed for the decay of magnetization shown in Figure 3.7 (ca. 400 min). Since no new ^{1}H resonances could be observed under these conditions, the decrease in signal intensity was therefore purely due to aggregation. These results suggested that the molecular chaperone Hsp104 modulates a shift of equilibrium between different conformers of Sup35[5-26]. Initially, the Sup35[5-26] species that interact with Hsp104 were not populated and therefore no STD signal were observed. Binding of Hsp104 to these molecules, however, could shift the distribution towards these conformers by the law of mass action. Their concentrations increase and hence STD signal observed.

At 27 °C, Sup35[5-26] behaved strikingly different. Figure-3.16 displayed the STD amplification factors of the tyrosine Hδ resonance lines for the initial, the intermediate and the final conformers of Sup35[5-26] (as indicated in Figure 3.11). Other resonance lines showed qualitatively the same behaviour. The tyrosine signal was selected in the analysis since it did not overlap with other ^{1}H resonances.
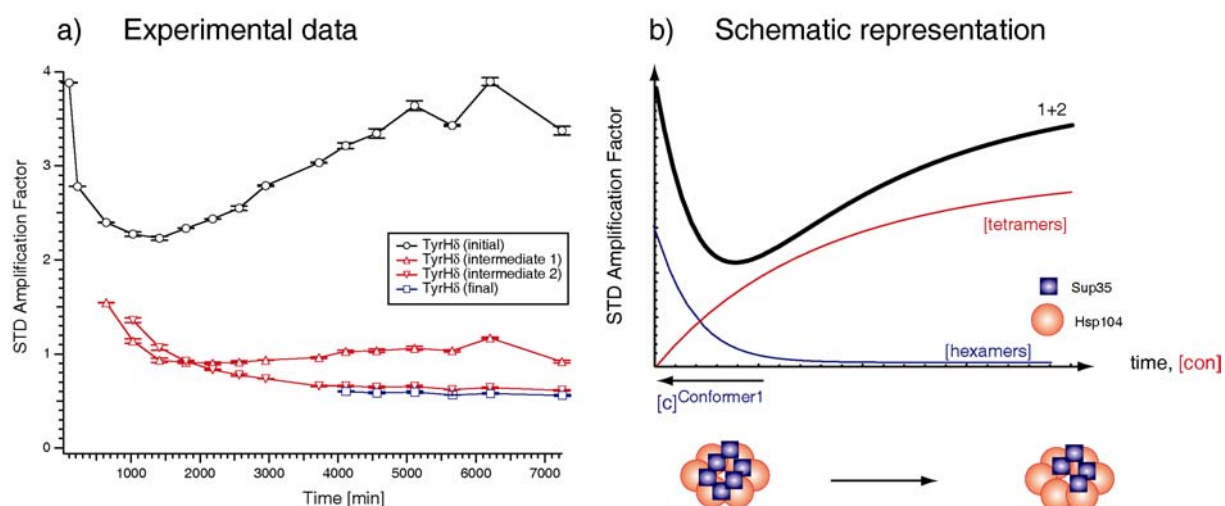
Figure 3.16. a) Time dependence of STD signal for Sup35[5-26]:Hsp104 in a molar ratio 25:1, at T=27°C. b)Schematic representation of influence of binding affinity on STD amplification factor at different oligomeric populations

 The initial conformer (black circles) showed the strongest interaction with Hsp104, whereas the interaction of the intermediate conformers (intermediates 1 and 2, red triangles) were reduced. The final conformer (blue squares) only showed very weak binding to the chaperone. Error bars were based on the signal-to-noise ratio of the respective data sets. Generally, the error bars for the initial conformer were increased in the course of the experiment, whereas the error bars for the final conformer were decreased, due to the change in populations of the two species. Taking again into account the DOSY results (Figure 3.14), it can be concluded that a hexamer as well as a tetramer of Sup35[5-26] can interact with Hsp104 at this temperature (black circles). Superposition of one exponentially decaying and one increasing STD curve led to the observed amplification factor with an absolute minimum after ca. 1300 min.

A dip-type curve would also be expected for the STD amplification factor of intermediate 1 (red triangles up). The population of this species was increasing in the beginning, and then decreased after 2000 min (Figure 3.10). At the same time, the equilibrium for the average molecular weight was shifted from a tetrameric to a dimeric species, as monitored by DOSY experiments. However, the STD amplification factor (Figure 3.16) for this species is only monotonously decreasing, and then remains constant at a low absolute intensity. This was in agreement with a model in which a refolded, tetrameric Sup35[5-26] exhibits still some

interaction with Hsp104, whereas the dimeric conformer shows a decreased binding affinity to Hsp104.

For the second intermediate conformer of Sup35[5-26] (intermediate 2), a decrease in the STD amplification factor was detected. Simultaneously, small changes in the apparent molecular weight for this conformer were observed which correspond to a shift of equilibrium from a dimeric to monomeric species of Sup35[5-26]. However, the initial points in the DOSY experiments for intermediate-2 were difficult to access for a reliable interpretation due to the low signal to noise ratio for intermediate 2 at the beginning of the experiment.

Similarly, no change in molecular weight or the STD amplification factor for the resonance lines could be observed for the final conformer of Sup35[5-26] (blue squares). The STD amplification factor was very low throughout the experiment, indicating that there was no interaction between the release conformer of Sup35[5-26] and Hsp104.

The above experimental interpretation were confirmed by performing the same STD experiment, but keeping a constant binding affinity oligomers all the time. A ratio of Hsp104: Sup35[5-26] =1:100, at 27°C were used. No change in binding affinity was observed throughout the time of the experiment as the system was stable and non-dynamic.
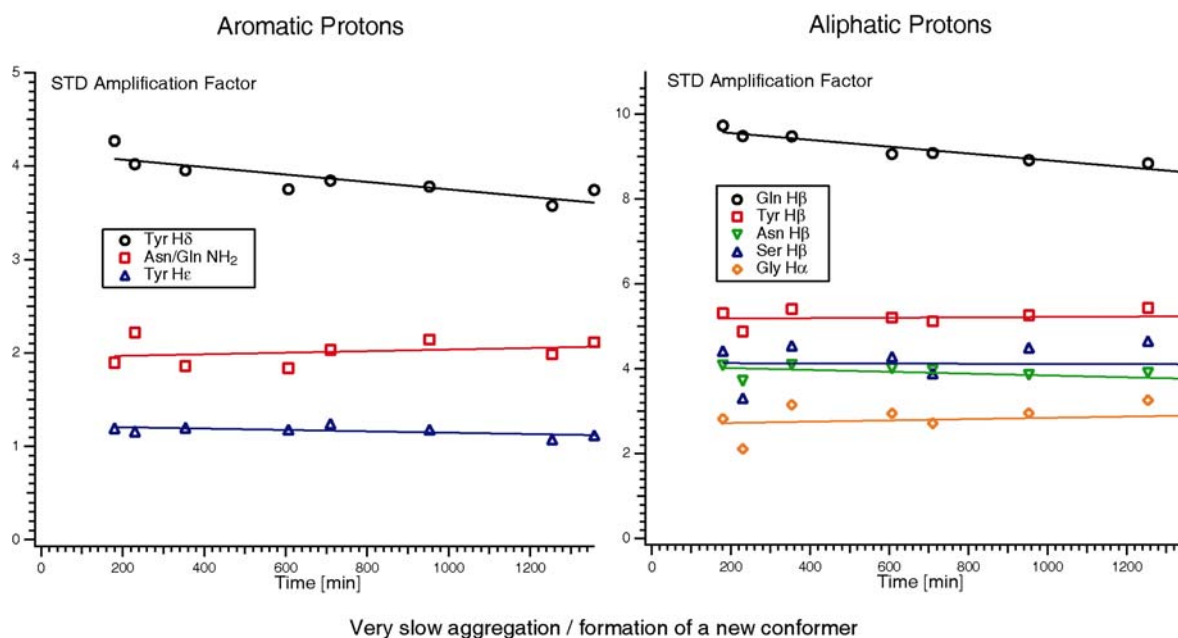
Figure 3.17. STD profile for [Hsp104]:[Sup35[5-26]] at 1:100 at T=27°C . As the system is stable and no appearance of differentially binding oligomers appearing, the STD amplification profile as a function of time is constant

Combining binding affinity and oligomeric distribution, I conclude that the driving force of the disaggregation process is the formation of new oligomeric species that are not interacting with Hsp104. After dialysing out the new monomeric species, it was clear that the new conformation can no longer aggregate. Also, the catalytic behaviour of Hsp104 towards the oligomers was set-in, because the new oligomeric species showed minimum interaction with Hsp104 (Fig. 3.17).

3.4.5 Interactions between fibrillar Sup35[5-26] and Hsp104 as studied by equilibrium dialysis experiments
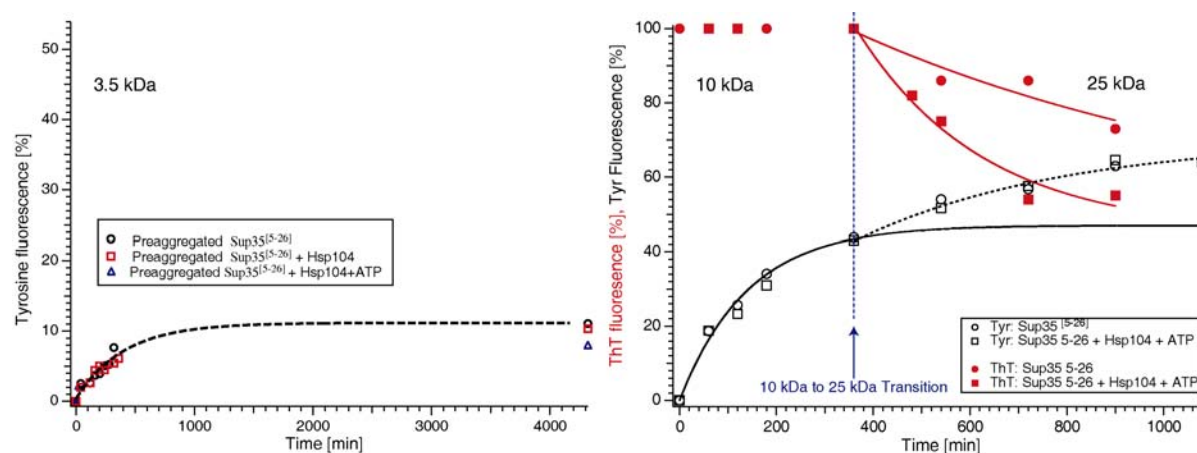
Figure 3.18. The equilibrium dialysis experiment by monitoring monomer release (Left). The equilibrium and Thioflavin-T binding assay by monitoring octamers and tetramers release (Right)

Temperature dependent ^{1}H intensities in NMR spectroscopy studies suggested that fibrils are always in equilibrium with soluble oligomers (Fig. 3.10). Also diffusion studies suggested the existence of various oligomeric Sup35[5-26] species. From the diffusion NMR, No information about the distribution of oligomers can be obtained. Only the average molecular weight is available in these experiments, Therefore, equilibrium dialysis experiments were performed for freshly solubilized Sup35[5-26] as well as for fibrillar Sup35[5-26]

First, the disaggregation of Sup35[5-26] fibrils was monitored to find out whether Hsp104 breaks down fibrils into monomers or not. Pre-aggregated Sup35[5-26] was injected into the membrane chamber. Using a 3.5kDa MWCO membrane, we find that only 9% of Sup35[5-26] monomers diffused out of the membrane irrespective of the presence of Hsp104 (Fig. 3.18a). Using a 10 kDa MWCO membrane which allows oligomeric Sup35[5-26] up to tetramers to diffuse out of the dialysis bag, a value of 40% of the Sup35[5-26] outside of the dialysis bag

was obtained after equilibration (Fig. 3.18b). At the same time, ThT fluorescence was not affected (Fig. 3.18b) as a function of time. Upon changing the membrane cut off value to 25 kDa, which would now also allow octameric Sup35[5-26] to move out of the membrane, ThT fluorescence values are rapidly decreased (Fig.3b). This decrease was enhanced in the presence of Hsp104 and ATP (50%), whereas a value of 75% was obtained in the absence of Hsp104. Irrespective of the presence of Hsp104, 60% of the Sup35[5-26] oligomers diffused through the dialysis membrane (Fig. 3.18b).

Aggregation kinetics of solubilized Sup35[5-26] using equilibrium dialysis is not easy to analyze since this characterized as competition between aggregation and dilution processes. The former is driven by fibril assembly and the latter is governed by diffusion of the Sup35[5-26] out of the dialysis bag. From the results on pre-aggregated Sup35[5-26] (Fig. 3.18a) it is clear that the monomers of Sup35[5-26] were not in equilibrium once the fibrils formed. From DOSY and STD experiments, interactions of Hsp104 with higher order soluble oligomers ((Sup35[5-26])$_{8-12}$) were not observed. Due to molecular weight restrictions, interactions between of Hsp104 and protofibrillar or fibrillar Sup35[5-26] cannot be addressed by NMR. But, ThT fluorescence (ThT) is sensitive to higher order oligomers ((Sup35[5-26])$_{>20}$) and fibrils. The differential effect that was observed in ThT fluorescence upon changing the MWCO from 10 kDa to 25 kDa was not observed in tyrosine fluorescence quantification (Fig. 3.18b). This result suggests that the involvement of Hsp104 upon fibrils and higher order oligomers ((Sup35[5-26])$_{>20}$) may not be specific, but, the Brownian forces exerted by the movement of Hsp104 upon protofibrils and fibrils could be enormous. Hence, these non-specific interactions may disrupt the protofibrils to fibril reorganization and hence disrupt higher order aggregates into lower order aggregates that are higher in molecular sizes than octamers but lower than the oligomers that can be detected by ThT binding. It has also been postulated that molecular chaperones, the family of Clp proteins, use Brownian-like mechanical forces to disaggregate the pre-aggregated materials (*32*). But, these forces can be effective only to higher order aggregates. Re-solubilization or refolding of fibrils might rely on a different mechanism for which specific molecular chaperones would have been needed (*13, 32*).

3.4.6 Interactions between soluble Sup35[5-26] and Hsp104 studied by equilibrium dialysis
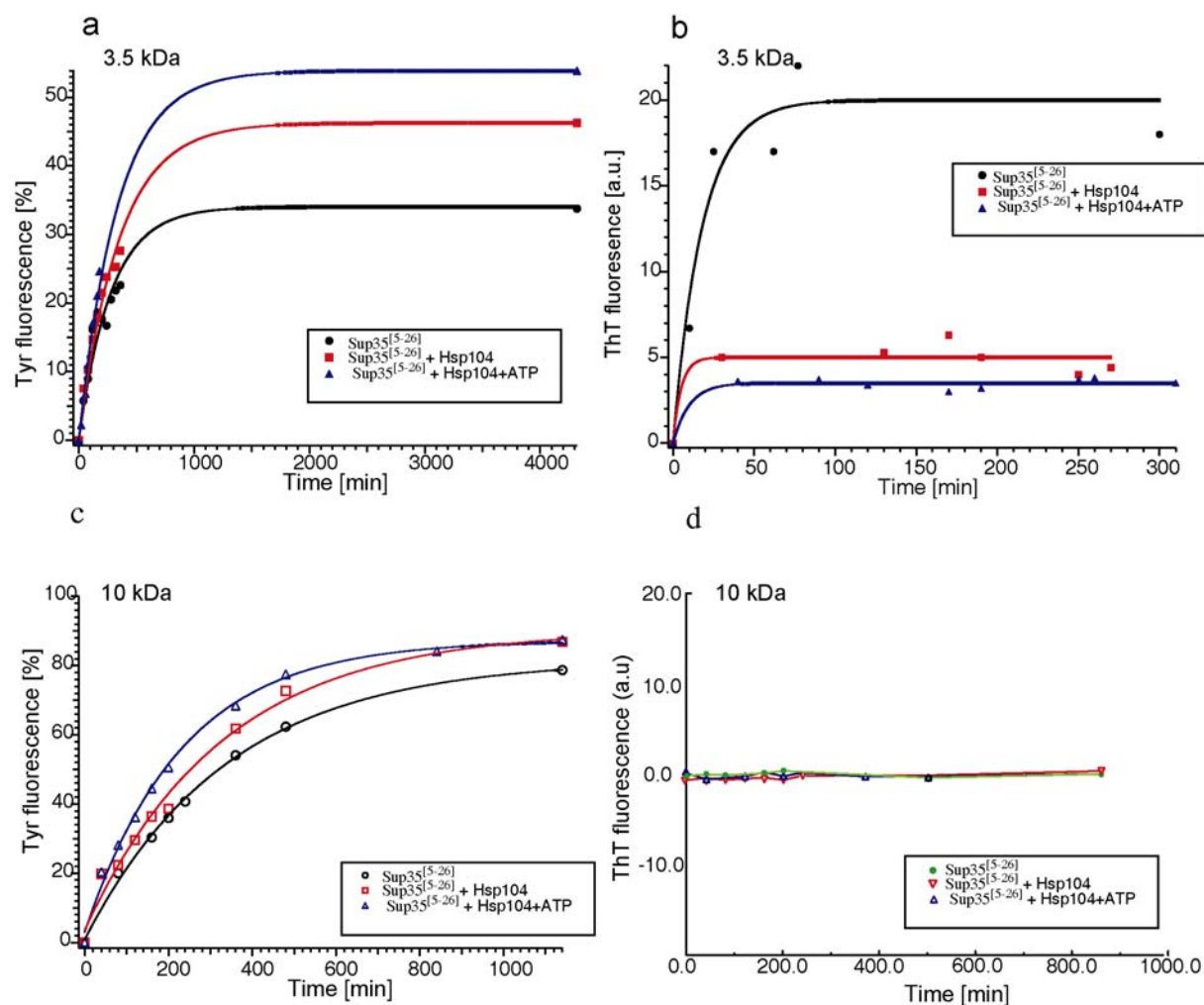
Figure 3.19. Equilibrium dialysis experiments (a,c) and ThT experiments (b,d) monitoring monomer release (a,b) and tetramer release (c,d) for the freshly solubilized Sup35:Hsp104 at a molar ratio 1:25

In order to obtain information about the distribution of oligomers equilibrium dialysis experiments on freshly solubilized material were performed. Hsp104 and Sup35[5-26] at a molar ratio of 1:25 were incubated in dialysis bags with different MWCO. Using a MWCO of 3.5 kDa, allowing the release of monomeric Sup35[5-26], a strong monotonous increase in ThT fluorescence was observed in the absence of Hsp104. This increase was less pronounced in the presence of Hsp104 (Fig. 3.19b). Tyrosine fluorescence analysis confirms that an increased amount of monomer is released for samples incubated with Hsp104 (Fig. 3.19a). At

the end of the experiments, 30% of Sup35[5-26] diffused through the membrane in the absence of Hsp104, whereas in the presence of Hsp104, 50% and 44% of Sup35[5-26] could escape the dialysis bag in the absence and presence of ATP, respectively (Fig.3.19a).

The experiments have been repeated using a 10 kDa MWCO membranes, corresponding to the size of a tetrameric Sup35[5-26]. At the end of this experiment, approximately 80% of Sup35[5-26] diffused out of the membrane in the absence of Hsp104 and complete equilibration was observed for samples containing Hsp104 irrespective of the presence of ATP (Fig.3.19c). No amyloidogenic intermediates were observed in the ThT fluorescence assay using the 10 kDa MWCO membrane irrespective of the presence of Hsp104 and ATP (Fig.3.19d).

The above experiments clearly indicate that the low-oligomeric weight species comprising three to seven molecules of Sup35[5-26] ((Sup35[5-26])$_{3-7}$) are critical for the aggregation.

3.4.7 Structural investigation of the interactions between Hsp104 and Sup35[5-26]

In order to follow structural changes of Sup35[5-26] upon addition of Hsp104, CD spectroscopy analysis were carried out. CD analysis shows that the initial oligomers had a higher percentage of helices. Presence of Hsp104 did not alter the population of helices significantly (Table 3.1). Convolution of CD data Sup35[5-26] of before and after incubation with Hsp104 showed very similar profile, But CD curves are markedly different. CD curve for aggregated Sup35 indicates β-sheet structure (Fig. 3.20a). These results supported previous CD-secondary structural analysis on full length Sup35 and Hsp104 that the mixture of Sup35 and Hsp104 differed significantly from individual spectra (*17*).

Table 3.1 CD Analysis of Sup35[5-26]

Samples	Helices	Sheets	Turn	Others
Solubilized Sup35[5-26] without Hsp104	39%	28%	16 %	17%
Aggregated Sup35[5-26] without Hsp104	0%	48 %	23%	29%
[Sup35[5-26]]:[Hsp104](Molar ratio=1:25) after dialysed out the peptide	37 %	31%	16%	16%

After aggregation, the CD spectrum yields a higher percentage of β-sheets (Fig. 3.20a). To understand the nature of Sup35[5-26] oligomers and the impact of the interaction of Hsp104 upon secondary structures, structural investigations were carried out using NMR experiments.
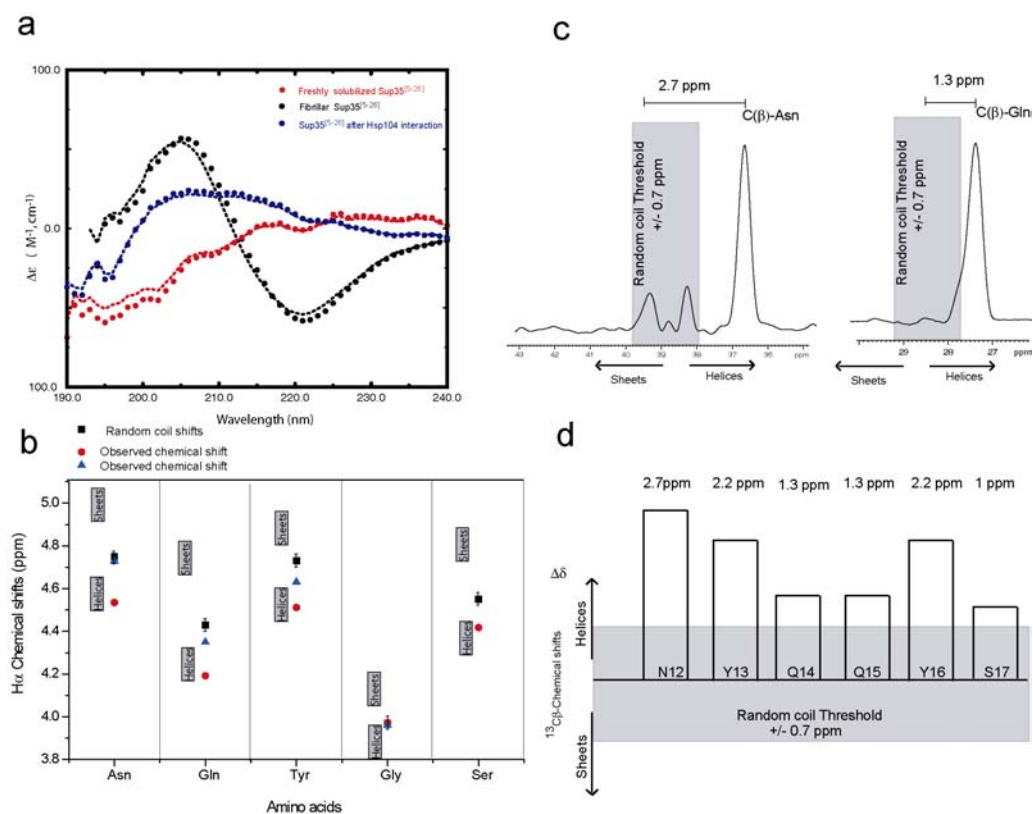
Figure 3.20. a) Structural analysis using CD spectroscopy. b)Secondary structure analysis using deviation of Hα chemical shift from random coil for respective aminoacids. c) Dispersion of C-β of Asparagines and Glutamines. d) a sequence stretch in Sup35[5-26] adopts α- helical structure

The dispersion of the H^N NMR resonance frequencies indicated that Sup35[5-26] did not adopt a random coil conformation (Fig. 3.10). At the same time, the Hα chemical shifts were up field shifted compared to the water resonance frequency. A down field shift would be indicative for β-sheet structures. A dispersion in Hα Chemical shifts of about 0.2 ppm for Asn, Gln and Tyr is observed (Fig. 3.20a). The experimental H^{α} chemical shifts were then compared to the Hα Chemical shift values of 2613 proteins from the BMRB protein database, (BioMagResBank , University of Wisconsin, USA). We observe 0.15 ppm upfield shift compared to a random coil chemical shift value for Hα resonances of Asn, Gln, Tyr and Ser.

The observed deviations in Hα chemical shift values indicates the presence of helical conformation within Sup35[5-26]. All the above findings indicate that the observed oligomers are structured and are not composed of β-sheets. In order to confirm the above findings, ^{1}H{^{13}C}HSQC spectra at ^{13}C natural abundance were recorded. For the immediately solubilized Sup35[5-26], the asparagine C chemical shifts show a dispersion of about 2.7 ppm, which clearly indicates the presence of structured regions within Sup35[5-26] (Fig. 3.20b). In addition, the Cβ-carbon chemical shift values of Asn, Gln, Tyr, Ser for residues N^{12}YQQYS^{17} in Sup*35*[5-26], which could be unambiguously assigned by combining NOESY and TOCSY spectral information, deviate from the random coil shifts threshold (*33*) by 2.0 ppm, 0.6 ppm, 1.5 ppm and 0.3 ppm, respectively and indicate an α-helical propensity for these residues (Fig.3.20c). Interestingly, upon interaction with Hsp104, the released low-molecular weight Sup35[5-26] oligomers did not display significant changes in carbon chemical shifts (data not shown), though subtle changes in structure were observed estimated from aromatic proton chemical shifts. The conservation of carbon chemical shifts indicated that no major structural changes occur during the dissociation of higher order oligomers (hexamer) into smaller ones (monomer) induced by Hsp104.

From the results, it can be concluded that the critical oligomers were rich in α-helices and Hsp104 either stabilizes the helical conformation or destabilizes the fibril assembly where backbone interactions overcome the side chain and extended to form sheets. All the above results prove that the structural re-organization happened at a later stage of fibril assembly and occur only in higher order oligomers.

From all the experiments, the state of Hsp104 was unchanged. It existed as hexameric complex at all the conditions employed. This had been proven from glutaraldehyde cross-linking experiment (Stefan Walter, TUM) and dynamic light scattering experiments (DLS) (.Niko Neumeier and Sevil Weinkauf, TUM). Hsp104 could be cross linked at 15 μM as well as 60 μM, which the concentration that used for the NMR experiments (Fig. 3.21a). DLS experiments showed that, at both the temperature 12 °C and 27 °C, the solution was mono-disperse. The average hydrodynamic radius was ca.7.5 nm (Fig. 3.21b,c).
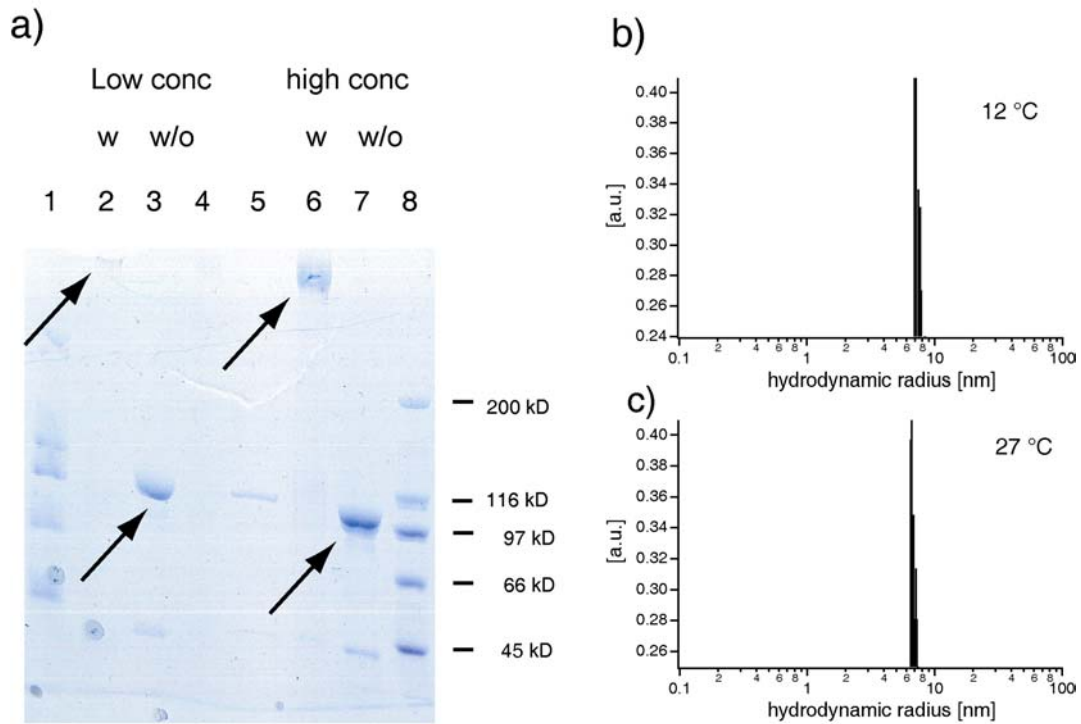
Figure 3.21 a) Cross-linking using glutaraldehyde and SDS-PAGE showed that the Hsp104 can be cross-linked to higher molecular weight oligomers both at low and high concentrations. b)Dynamic light scattering experiments at 12 °C. c)DLS at 27°C

3.5   Concluding remarks
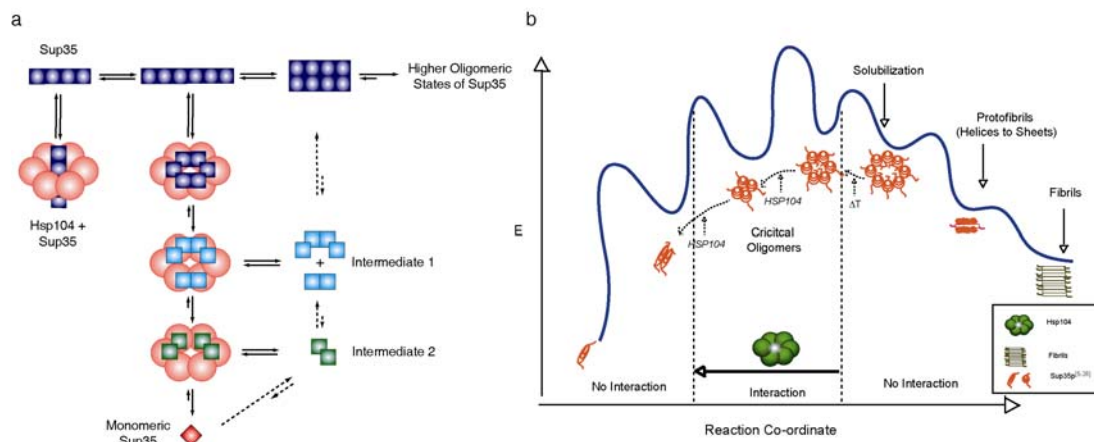
Figure 3.22. a) Interaction model for Hsp104 and Sup35; b)Schematic energy level diagram showed that the Hsp104 helped prion seeds to transit the critical energy level barrier.

We could show that the low-molecular weight oligomers (Sup35$^{[5-26]}$)$_{3-7}$) are critical for fibril formation and should be important for the prion propagation. Keeping these oligomers out of the equilibrium pathway would prevent aggregation. A model of different equilibria between monomeric and oligomeric species Sup35 and their interaction with Hsp104 is depicted in Fig 3.22a. The proposed model of interaction between Sup35 and Hsp104 also has possible implications on the assembly process of amyloid fibrils. It is plausible to assume that monomeric Sup35 associates with small oligomers and eventually larger species that may serve as nuclei for amyloid fibril formation or aggregation. In the absence of Hsp104, these intermediates are short-lived and cannot be detected by NMR. However, in the presence of the chaperone, they are stabilized and therefore amenable for structural investigation. In a model described by Lindquist, Hsp104 interacts specifically with oligomeric Sup35 (*15*). This model is in agreement with our observations. Hsp104 specifically recognizes critical oligomers (3 < n > 7), dissociates them into soluble monomers and prevents aggregation. These critical oligomers adopt helical conformation and the conformational reorganization into β-sheets did not occur prior to aggregation. Without changing the secondary structures of the critical oligomers, Hsp104 modulates the activation energy barrier for the transition between oligomeric intermediates (Fig. 3.22b). The specificity of molecular chaperones towards these proteins is one of the prime factors in manipulating the aggregation pathway. Using such specificities, designing "chemical chaperones" that can slow down, arrest or revert the disease progression is an emerging strategy for the prevention of aggregation dependent diseases (*34*).

3.6   References

1.   Moldave, K. (1985) *Annu.Rev.Biochem. 54*, 1109-1149.
2.   Ruiz-Echevarria, M. J., Czaplinski, K., Peltz, S.W. (1996) *Trends Biochem.Sci. 21*, 433-438.
3.   Bertram, G., Innes, S., Minella, O., Richardson, J.P., Stansfield, I. (2001) *Microbiol. 147*, 255-269.
4.   Patino, M. M., Liu, J.-J., Glover, J. R., and Lindquist, S. (1996) *Science 273*, 622-626.
5.   Depace, A. H., Santoso, A., Hillner, P., Weissman, J.S. (1998) *Cell 93*, 1241-1252.
6.   Liu, J., Sondheimer, N., Lindquist, S.L. (2002) *Proc Natl Acad Sci U S A. 99*, 16446-16453.
7.   Stansfield, I., Jones, K.M., Kushnirov, V.V., Dagkesamanskaya, A.R., Poznyakovski, A.I., Paushkin, S.V., Nierras, C.R., Cox, B.S., Ter-Avanesyan, M.D., Tuite, M.F. (1995) *Embo J 14*, 4365-4373.
8.   Krobitsch, S., Lindquist, S. (2000) *Proc Natl Acad Sci U S A. 97*, 1589-1594.
9.   Kaneko, K., Zulianello, L., Scott, M., Cooper, C.M., Wallace, A.C., James, T.L., Cohen, F.E., Prusiner, S.B. (1997) *Proc Natl Acad Sci U S A. 94*, 10069-10074.
10.   Schirmer, E. C., Glover, J.R., Singer, M.A., Lindquist, S.L. (1996) *Trends Biochem.Sci. 21*, 289-296.
11.   Schirmer, E. C., Queitsch, C., Kowal, A. S., Parsell, D. A., and Lindquist, S. (1998) *J. Biol. Chem. 273*, 15546-15552.
12.   Parsell, D. A., Kowal, A.S., Lindquist, S. (1994) *J. Biol. Chem. 269*, 4480-4487.
13.   Lee, S., Sowa, M.E., Watanabe, Y.H., Sigler, P.B., Chiu, W., Yoshida, M., Tsai, F.T. (2003) *Cell 115*, 229-240.
14.   Parsell, D. A., Kowal, A.S., Singer, M.A., Lindquist, S.L. (1994) *Nature 372*, 475-478.
15.   Chernoff, Y. O., Lindquist, S.L., Ono, B., Inge-Vechtomov, S.G., Liebman, S.W. (1995) *Science 268*, 880-884.
16.   Paushkin, S. V., Kushnirov, V.V., Smirnov, V.N,, Ter-Avanesyan, M.D. (1996) *Embo J 15*, 3127-3134.
17.   Schirmer, E. C., Lindquist, S.L. (1997) *Proc Natl Acad Sci U S A. 94*, 13932-13937.
18.   Shorter, J., Lindquist, S. (2004) *Science 304*, 1793-1797.
19.   Schirmer, E. C., Lindquist, S.L. (1998) *Methods Enzymol., eds. Lorimer, G.H., Baldwin, T (Academic, Sandiego) 290*, 430-444.
20.   Burke, D., Dawson,D., Stearns, T. (2000) *Methods in yeast genetics*, Cold Spring Harbor Laboratory Press.
21.   Stejkal, E. O., Tanner, J.E. (1965) *J.Chem.Phys. 42*, 288.
22.   Wu, D., Chen, A., Johnson, C.S. (1995) *J.Magn.Reson.A. 115*, 260-264.
23.   Cantor, C. R., Schimmel, P.R. (1980) *Biophysical chemistry, Part II: Techniques for the study of biological structure and function, pp.530-590, W.H.Freeman, Newyork, NY.*
24.   Perutz, M. F., Finch,J.T., Berriman,J., Lesk, A. (2002) *Proc Natl Acad Sci U S A. 99*, 5591-5595.
25.   Wilkins, D., K., Grimshaw, S,B., Receveur, V., Dobson, C.M., Jones, J.A., Smith, L,J. (1999) *Biochemistry 38*, 16424-16431.
26.   Yao, S., Howlett, G.J., Norton, R.S. (2000) *J.Biomol. NMR 16*, 109-119.
27.   Mayer, M., Meyer, B. (2001) *J.Am.Chem.Soc. 123*, 6108-6117.
28.   Mayer, M., Meyer, B. (1999) *Angew.Chem.Int.Ed. 38*, 1784-1788.
29.   Takahashi, H., Nakanishi, T., Kami,K., Arata, Y., Shimada, I. (2000) *Nature Struct Biol 7*, 220-223.
30.   Jayalakshmi, V., krishna, N, R. (2002) *J.Magn.Reson 115*, 106-118.

31.　　　Yan, J., Kline,A.D., Mo,H., Shapiro, M.J.Zartler, E.R. (2003) *J.Magn.Reson 163*, 270-276.

32.　　Diamant, S., Ben-Zvi, A. P., Bukau, B., and Goloubinoff, P. (2000) *J. Biol. Chem. 275*, 21107-21113.

33.　　Wishart, D. S., Sykes, B.D. (1994) *J.Biomol. NMR 4*, 171-180.

34.　　Cohen, F. E., Kelly, J.W. (2003) *Nature 426*, 905-908.

4.0   Study of interactions between β-Amyloid and αB-crystallin

4.1   Background

Alzheimer's disease (AD) is the most abundant age-related, devastating neurodegenerative disease. This disease is characterized by the degeneration and death of neurons in the brain regions that are concerned with learning and memory processes (*1*) (Fig. 4.1 a,b). In widely accepted amyloid cascade hypothesis, the burden of secreted and intracellular accumulation of β-amyloid peptides (Aβ) of varied lengths (39-43 amino acids) is considered a main event (*2, 3*) and the Aβ peptides are considered as neurotoxic elements (*2, 3*) (Fig. 4.1 c). Aβ peptides are fragments of a 110-130 kDa type I integral membrane glycoprotein, the amyloid precursor protein (APP), which is cleaved by a class of endoproteases termed as secretases (*4*) (Fig. 4.1 d).
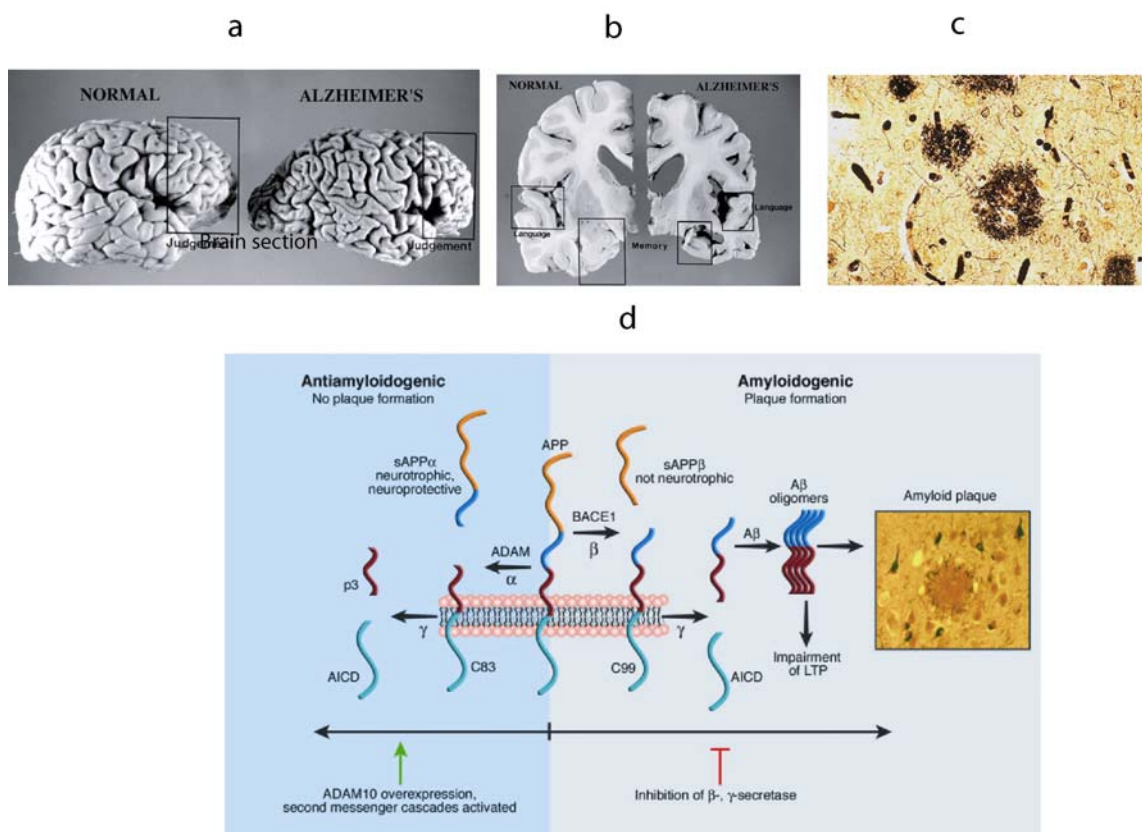
Figure 4.1. a) Comparison of normal brain and Alzheimer's brain; b) Comparison of normal brain and Alzheimer brain's slice c) Accumulation of amyloid peptides on brain tissue d) Processing of secretases cause Alzheimer's deposits (picture adapted from Lichtenthaler, S.F., Haass, C., J .Clin. Invest. (2004),113, 1384).

These secretases are classified as α, β and γ secretases. The α-secretases (consisting of ADAM10, ADAM17 enzymes) are metallo-proteases which cleave the APP both intracellularly (*5*) and at the cell surface (*6*), thus leaving membrane bound C83 and soluble sAPPα fragments. Simultaneously, intra-membrane γ secretases (consisting of presenilins, nicastrin and other complexes (*4*)) cleave the C83 fragment and produce soluble p3 fragment (Fig. 4.1d). The α-and γ-secretase processing of APP is known as anti-amyloidogenic pathway. Alternately, the β secretase (known as BACE, an aspartyl protease (*7*)) cleaves the APP and leaves the membrane bound C99 fragment. Simultaneously, γ-secretases process the membrane bound C99 and produce Aβ peptides which are in turn cleared by the neural endo-peptidase similar to neprilysin enzyme (*8*). This simultaneous processing of β-and γ-secretase on APP is known as amyloidogenic pathway. The production, processing and clearance of Aβ peptides are parts of the normal constitutive APP metabolism. Profound changes in the production and/or clearance pathway cause the disease (*9*) (Fig. 4.1d).

Abnormal protein deposition is a shared characteristic of other late onset neurodegenerative diseases, such as Parkinson's, Huntington's, and the Prion diseases. There are increasing evidences that the mechanism of this aggregation may be similar in each of these diseases (*10, 11*). At the same time, it is found that probably not the fibrillar, but a protofibrillar state, which is in fast equilibrium with a so-called Low Molecular Weight (LMW) state, is responsible for the neurotoxicity (*12, 13*). These protofibrils are known to be causative agents for selective neuronal depletion (*14, 15*). Large fibrillar aggregates are believed to act as physical barriers to transport and other essential neuronal functions (*16*). So far, the reasons for the neuronal degeneration have not been identified. It is speculated that Aβ(1-40) forms a pore in the membrane that leads to an unregulated flux of Ca^{2+} (*17*). Furthermore, free radical formation (*18*), mitochondrial dysfunction (*19*) and capsase activation (*20*) have also been postulated as reasons for neurotoxicity (Fig. 4.2). However, the causes of neurotoxicity and the initial events in the cascade of neuronal cell death are still unclear.
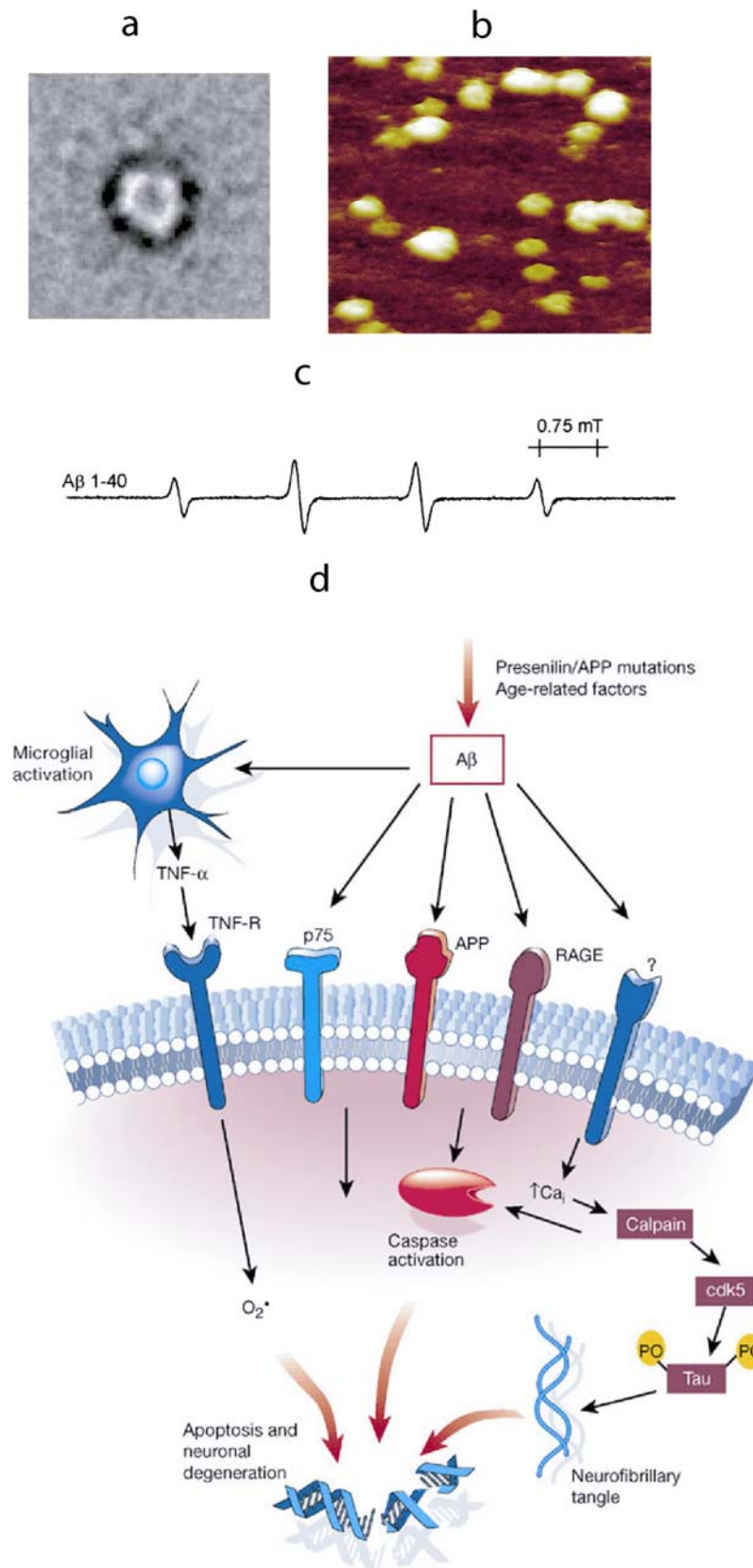
Figure 4.2. a) "Amyloid pore" as observed by electron microscopy (Lashuel, H.A.., et al., Nature (2002), 418, 291); b) Channel formation by amyloid as observed by AFM (Lin, H., et al., FASEB J (2001), 15, 2433) c) Free radical generation by Amyloid as observed by ESR (Dikalov, S.I., et al., J.Biol.Chem.(1999), 274,9392) d) Cellular pathways of amyloid protein neurotoxicity in Alzheimer's disease (Yuan, J. and Yankner, B.A.., Nature (2000), 407, 802).

Using transgenic *Caenorhabditis elegans model*, the intracellular factors that contribute to neurotoxicity and metabolism of Aβ have been studied recently (*21*). The results identify molecular chaperones as binding partners of Aβ (*21*). Importantly, in the brain of AD patients, high expression of molecular chaperones has also been observed (*22*). In particular, a major lenticular, non-tissue specific, small heat shock protein (sHsp), αB-crystallin, is found to co-precipitate together with Aβ(1-40) (*22, 23*). The sHsp family is both large and diverse, with members found in virtually all organisms. Many of the sHsps have conserved C-terminal α-crystallin domains (*24*) and they are involved in assembly of oligomers. The recent crystal structure of one of the small heat shock proteins, wheat sHsp16.9, shows that the dodecameric oligomers of sHsp16.9 are stabilized by α-crystallin domain that lie at the interface of sHsp16.9 (*25*) (Fig. 4.3).
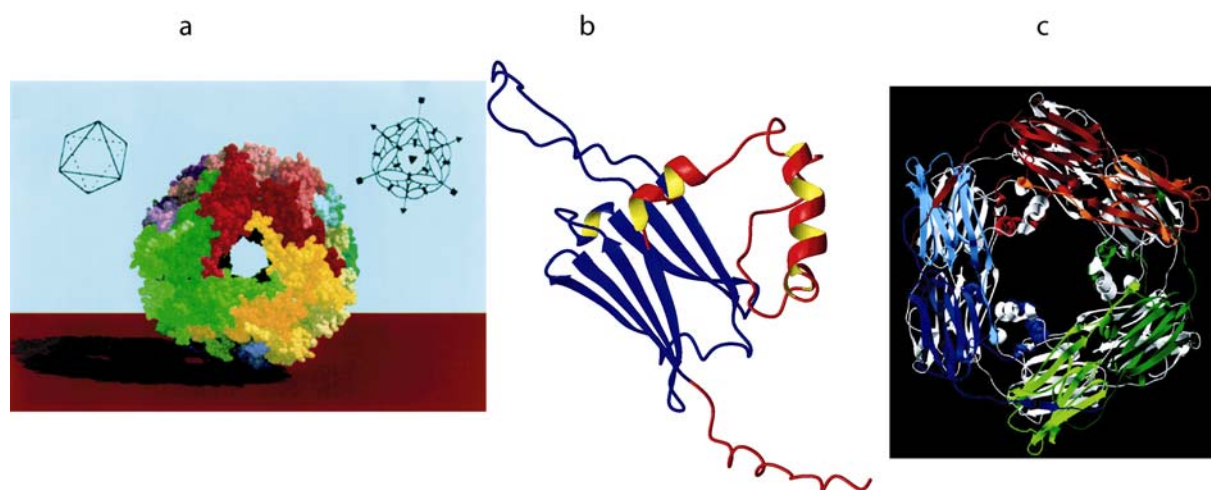
Figure 4.3. Dodecameric structure of sHsps (a) formed by the unique α -crystallin fold (colored blue in (b)) . Dodecamer is arranged as two disks, each one of them consists of 12 subunits arranged into two hexameric rings. c) One of the two hexameric rings (van Montfort, R. M., et al., Nature Struct.Biol.(2001), 8, 1025)

The link between Alzheimer's disease and αB-crystallin is even more striking taking into account that nearly all the cases of AD patients with trisomy-21 (Down's syndrome) display abnormal expression of αB-crystallin (*26*). Notably, all Down's syndrome (DS) patients show AD pathology above 40 years of age (*27*). The observation of frequent equatorial supramolecular cataracts in lenses from AD patients as compared to normal patients, co-localized with amyloid peptide, Aβ (1-40) (*28*) established the connection between crystallins and Aβ (1-40).

αB-crystallin also plays a major role in several neurological (*29, 30*) and neuro-muscular diseases (*31*). The molecular mechanism of the interactions between αB-crystallin and Aβ(1-40) is not understood. *In vitro*, αB-crystallin has been shown to interact with amyloid peptides and reported to inhibit fibril formation (*32, 33*) (Fig. 4.4). At the same time, pre-incubation of Aβ (1-40) with αB-crystallin yields increased neurotoxicity in spite of the fact that fibril formation is reduced(*34*) (Fig. 4.4).
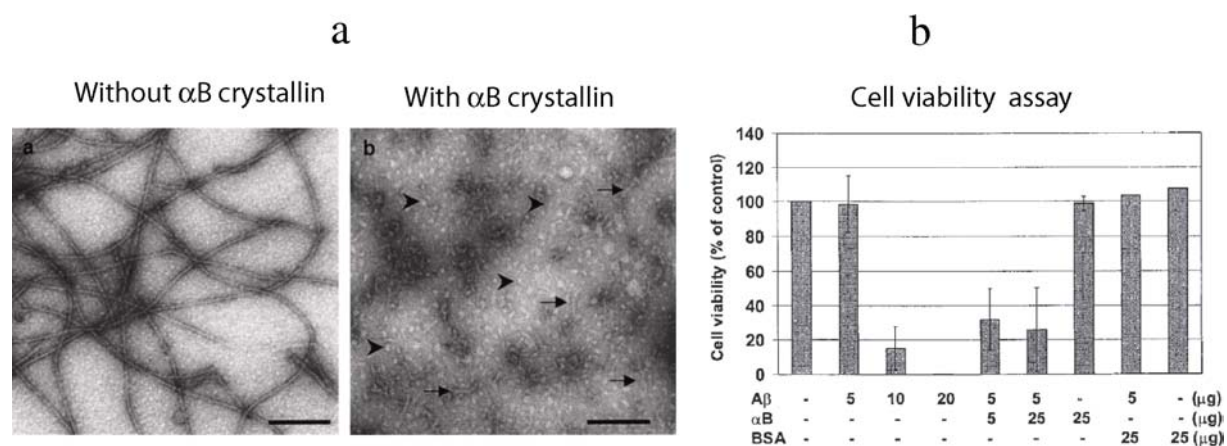
Figure 4.4. a) Disruption of amyloid fibrils upon incubation with aB-crystallin; b) Increase in neurotoxicity to the cultured neuronal cells upon incubation with αB–crystallin as observed by cell viability assay. (Stege, G.J., et al., Biochem.Biophys.Res.Commun.(1999), 262, 152)

4.2   Aims of the project

1. To study the Aβ(1-40) aggregation process
2. To characterize the chemical groups that are involved in Aβ(1-40)·Aβ(1-40) and Aβ(1-40)·αB-crystallin interactions
3. To understand how αB-crystallin can increase neurotoxicity of Aβ(1-40)

4.3   Material and Methods

Aβ(1-40) was purchased from Biosource International, Camrillo, CA, USA. αB-crystallin was obtained from our collaborator Dr. Wilbert Boelens, University of Nijmegen, the Netherlands.

4.3.1 Stabilization of oligomers upon change in anionic strength

Different Aβ(1-40) samples were prepared by employing different salt conditions. To generate more soluble and stable Aβ(1-40) solution, the lyophilized Aβ(1-40) was dissolved in 50 mM phosphate buffer containing 100 mM sodium sulphate, pH 6.9. To prepare higher order oligomers of Aβ(1-40), 100 mM sodium chloride was used instead of 100 mM of sodium sulphate in the solubilization buffer. Fibrils of Aβ(1-40) were prepared by incubating Aβ(1-40) in Hexafluoro- isopropanol (HFIP) (1mg/ml) (Courtesy: Ms. Zhongjing Chen, FMP-Berlin). After 2 hours of incubation, HFIP was evaporated using a slow stream of dry nitrogen gas. A thin layer of Aβ(1-40) resulted. The Aβ(1-40) layer was dissolved in 10mM phosphate buffer, pH 7.0 and incubated for 4 days at 25 °C. Ordered amyloid fibrils were observed by electron microscopy (shown below).

Aβ(1-40) fibrils as observed by EM
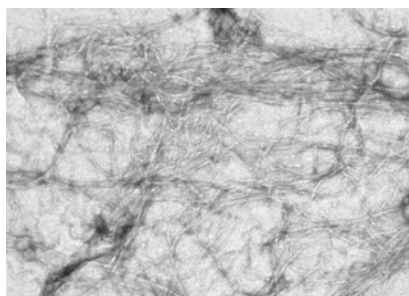
4.3.2 Saturation Transfer Difference (STD) 2D- correlation experiment
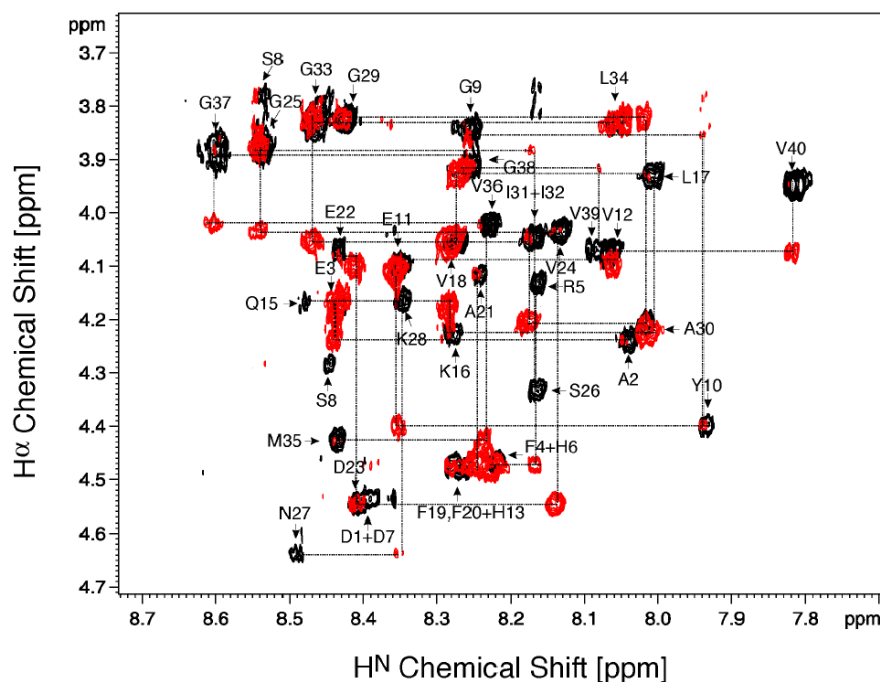
Figure 4.5. Sequential H^{N}-H^{α} assignment using NOESY (red) -TOCSY(black) spectra

Table 4.1 Chemical shift table for Aβ(1-40) –H^N-H^α region

Assignment	H^α - Chemical Shifts	H^N - Chemical Shifts
A2	4.246	8.045
E3	4.181	8.443
F4	4.473	8.232
R5	4.140	8.169
H6	4.470	8.225
D7	4.541	8.397
S8	4.289	8.453
G9	3.851	8.260
Y10	4.406	7.941
E11	4.106	8.358
V12	4.082	8.072
H13	4.483	8.262
H14	4.543	8.367
Q15	4.176	8.488
K16	4.234	8.285
L17	3.937	8.010
V18	4.062	8.282
F19	4.483	8.287
F20	4.483	8.273
A21	4.123	8.248
E22	4.076	8.439
D23	4.552	8.415
V24	4.042	8.138
G25	3.885	8.545
S26	4.339	8.171
N27	4.646	8.502
K28	4.170	8.354
G29	3.832	8.431
A30	4.213	8.022
I31	4.052	8.173
I32	4.054	8.180
G33	3.835	8.472
L34	3.839	8.071
M35	4.433	8.446
M35ox	4.487	8.625
V36	4.031	8.240
G37	3.893	8.605
G38	3.920	8.259
V39	4.074	8.102
V40	3.956	7.815

The STD NMR experiments were performed on 500 μM solution of Aβ (1-40) together with 20 μM of the αB-crystallin, yielding a molar ratio of 25:1 of [Aβ (1-40)]:[αB-crystallin]. STD experiments were also performed on different oligomeric species of soluble amyloid species that are in fast exchange with aggregated state. STD-TOCSY spectra were recorded using a mixing time of 60 ms and 16 scans per t_1 increment. A total of 512 t_1 increments were collected with a pre-saturation period of 2.7 s. A train of 300 Gaussian shaped pulses of 12.5 ms (with 1.0 ms gap between the pulses) were employed on odd scans at –380 Hz (on resonance) and on even scans at +40000 Hz (off resonance) in order to saturate ^{1}H (methyl) resonances of protein. Subtraction was done by incrementing the receiver phase for every alternate scans. Identification of the entire spin systems was done using TOCSY and NOESY experiments and most of the resonances were assigned unambiguously (Fig. 4.5). The chemical shift values of H^N and H^α are tabulated (Table 4.1). STD experiments were adjusted in the first place with two reference samples, one containing a 1 mM solution of an amyloid fibril binding peptide, LPFFD (Fig 4.6a) and another containing 1.5mM of a Sup35p[5-26] peptide (Fig. 4.6b) (see also Chapter 3). 1D- STD experiment on Sup35 sample containing no protein were measured (Fig 4.6b).
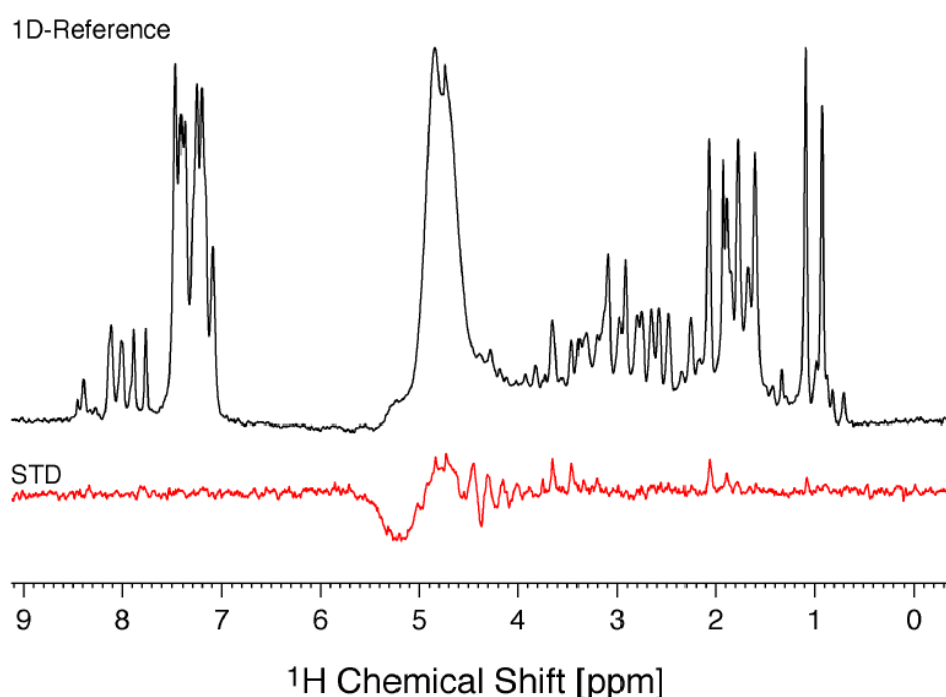
Figure 4. 6a. Calibration of STD experiment using peptide LPFFD

86
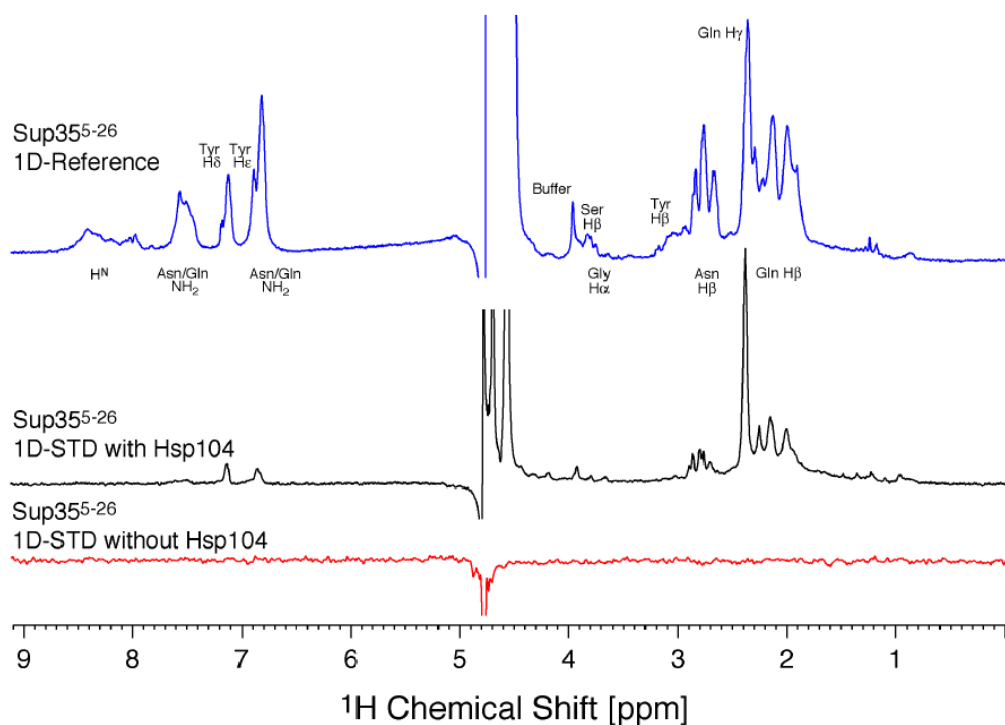
Figure 4.6b. Calibration of STD experiment using peptide LPFFD

4.3.3 Diffusion Ordered Spectroscopy (DOSY)

Pulsed-field gradient (PFG) NMR experiments using a water-sLED pulse sequence were used to determine the relative molecular size of oligomers of the amyloid peptide in presence and absence of αB-crystallin (*35, 36*). In the experiment, the duration of the gradient was set to 2.0 ms. A delay of 30.0 ms was used to allow diffusion between the dephasing and rephasing gradient, and 32 points were recorded in the indirect dimension. Sine-shaped gradients and a gradient ramp with a maximum gradient strength of 35 G/cm was used. Gradient strengths are always given in $\%^2$ with respect to the maximum gradient strength (100% = 35 G/cm). Acetone (M_r = 58 Da) and *β-* Mercapto-ethanol (M_r = 156 Da) were used as internal references to calculate the average molecular weight of the amyloid oligomers. The data analysis was performed using the procedure mentioned in section 3.3.2

4.3.4 NMR assay for the Redox active proteins

In order to assess the redox potential of Aβ(1-40) and αB-crystallin, the relative auto-oxidation of reduced glutathione (GSH) was determined by NMR. Auto-oxidation is mostly due to solubilized oxygen in the sample buffer, but could be altered in presence of added substances. Monomeric and dimeric form of GSH were differentiated by their different ${}^1H^N$

87

chemical shifts and were unambiguously identified by diffusion NMR experiments (DOSY). The ratio of the dimer to the monomer was obtained after recording 1D NMR experiments of the GSH mixtures at different time intervals. As a reference, the auto-oxidation rate of monomeric GSH to dimeric GSSG was determined. In addition, the auto-oxidation of GSH in the presence of 50 µM ascorbic acid was also monitored. 1 mM solution of GSH was incubated with 50 µM of Aβ(1-40) (oligomers / fibril) and/or 10 µM of αB-crystallin. Samples were prepared at the same time and solubilized in the same buffer. NMR tubes were capped and sealed uniformly with parafilm paper.

4.3.5 Colorimetric redox assay

A simple colorimetric redox assay was performed by taking a resazurin based viable cell counting reagent, *Uptiblue* (Uptima, distributed by KMF Laborchemie Handels GmbH, Leipzig, Germany) as an indicator. 10 µl of *uptiblue* reagent was transferred into a 96-well plate. Afterwards, solutions of Aβ(1-40) with increasing concentrations, ranging from 80 µM to 400 µM, were added to the wells yielding a final volume of 0.1 mL. For αB-crystallin, the concentrations were increased from 8 µM to 40 µM, in steps of 8 µM. After the incubation for 48 hours, the samples were centrifuged and 50 µl of the supernatant was transferred into another set of 96-well plate. The colorimetric change was observed directly by scanning the samples. The experimental results were verified using dichloro-indophenol (Sigma) as a redox indicator.

4.3.6 Secondary structural analysis by circular dichroism spectroscopy

Circular dichroism (CD) spectra were measured using a Jasco J720 spectropolarimeter at ambient temperature. The path length of the cell was 1 mm. Spectra of peptides in solution were corrected by subtracting the buffer base line.

In order to study the change in the secondary structure of the amyloid peptide upon incubation with αB-crystallin, 200 µM of amyloid peptide was pre-incubated at room temperature with αB-crystallin (final concentration 8 µM) in a cuvette for several days (ratio 1: 25 of αB-crystallin to amyloid). CD spectra were recorded at different time points. The solutions were not filtered before the measurements. A reference experiment was performed at each time point, employing the same concentration of Aβ(1-40) in the absence of the αB-crystallin. CD experiments were also carried out for different samples of Aβ(1-40) that were dissolved in various anionic buffers (50 mM of Cl⁻/F⁻/ SO_4^{2-} / CrO_4^{2-}) and different NaCl concentration (50

mM, 100 mM, 0 mM). Here, the solutions were filtered through 0.2 μm membrane filters before measurements. The spectra were analysed using software CONTIN and CDSSTR (*37*).

4.3.7   Isothermal Titration Calorimetry (ITC)

Experiments were carried out on a high precision VP-ITC titration calorimetric system (Microcal Inc., MA). All titrations were performed by injecting 30 times 10 μl of titrant with a time delay of 300 s. Buffer titrations were performed in all the experiments in order to calculate the heat of dilution. 50 mM phosphate, 100mM sodium sulphate, pH 6.9 was used as a buffer. The titration of amyloid with αB-crystallin was performed by taking 1.39 ml of the 160 μM amyloid solution into the cell, and 300 μl of a 120 μM αB-crystallin solution in the syringe. The experiment was repeated by titrating copper (molar ratio of 2:1, [copper]:[αB crystallin]) to the αB-crystallin solution. Addition of copper to the amyloid was achieved by adding 800 μM of copper into the 160 μM of Aβ(1-40) solution. To study the effect of copper on αB-crystallin, the copper concentration was kept constant (150 μM) in the cell and was titrated against a 300 μl of 120 μM solution of αB-crystallin in the syringe. 300 μl of 480 μM of EDTA was used to remove the copper from the reaction mixture containing Aβ(1-40), αB-crystallin and copper at the concentration of 136 μM, 24 μM and 48 μM, respectively.

4.4   Results and Discussion

4.4.1   Understanding the Aβ(1-40).Aβ(1-40) interactions

Using buffer solutions containing 100 mM of either F^- or Cl^- or NO_3^- or ClO_4^- or SO_4^{2-} salts, a change in solubility of Aβ(1-40) was observed. The population of non-structured and structured region of solubilized Aβ(1-40) was followed by monitoring the ratio of the normalized intensity of random coil (I_{198}) and structured region (I_{220}) of the CD spectrum. We find that Aβ(1-40) solubilized in the sulphate buffer has increased the population of unstructured Aβ(1-40). In addition, the solubilization was proportional to the strength of the anion added in the order of its hardness $F^- < Cl^- < NO_3^- < ClO_4^- < SO_4^{2-}$. Moreover, a shift from random coil to structured region was observed upon decreasing concentration of chloride in the solubilization buffer (Fig. 4.7b). This effect was clearly visible at higher concentration of Aβ(1-40) (Fig. 4.7c) . (In Fig. 4.7b and 4.7c, the samples were not filtered prior to the CD experiments.) Though CD spectra are indicative of secondary structural analysis, the exact nature of anions upon aggregation of Aβ(1-40) can only be obtained by NMR experiments.
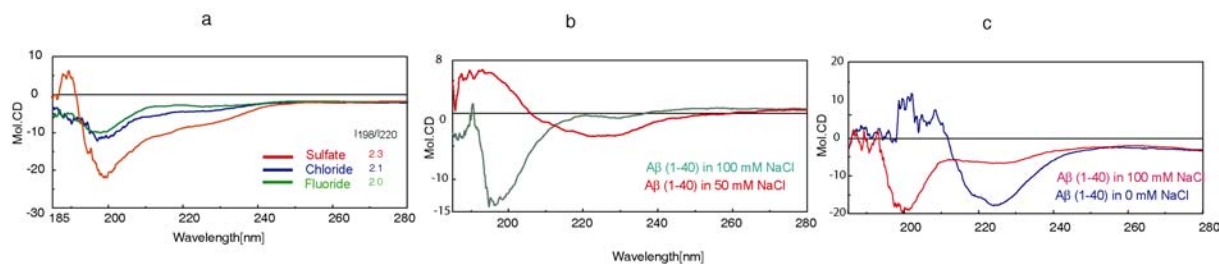
Figure 4.7. a) CD spectra of Aβ(1-40) solubilized in different anionic conditions; b) CD spectra of Aβ(1-40), solubilized at different concentration of NaCl, Concentration of Aβ(1-40)= 200 mM; c) CD spectra of Aβ(1-40) in presence and absence of salt, Concentration of Aβ(1-40) = 500 μM.

The effects observed in CD experiments were also observed by NMR spectroscopy. Comparing the intensity in a 1D-^{1}H NMR experiment, a gain of 25% intensity was observed in the sample dissolved in 50mM of sodium sulphate as compared to the sample that was dissolved in 50mM sodium chloride (Fig. 4.8a). The error in balancing can be considered to be on the order of 5%.

Also, reducing the salt concentration from 100 mM to 0 mM of sodium chloride drastically reduced the amyloid concentration in the solution by about 34% as judged from the peak intensities in 1D-^{1}H NMR spectroscopy (Fig. 4.8b). Interestingly, using chloride in the buffer, a broad peak at ~0 ppm was observed that was found to be increased upon decreasing the salt concentration. This resonance was not observed for the Aβ(1-40) sample dissolved in sulphate buffer (Fig. 4.8b). DOSY experiments were used to calculate the average molecular weight of Aβ(1-40). We find a molecular weight of approximately 20.8 kDa (in sulphate buffer), assuming a spherically shaped Aβ(1-40) molecule (Fig. 4.9). This molecular weight was larger than what was expected for a monomeric Aβ(1-40) molecule (M=4.33 kDa). The deviation could be explained either by assuming Aβ(1-40) associated into low-order oligomers or by assuming that Aβ(1-40) was in exchange with its fibrillar form, so that the average molecular weight appears to be larger.
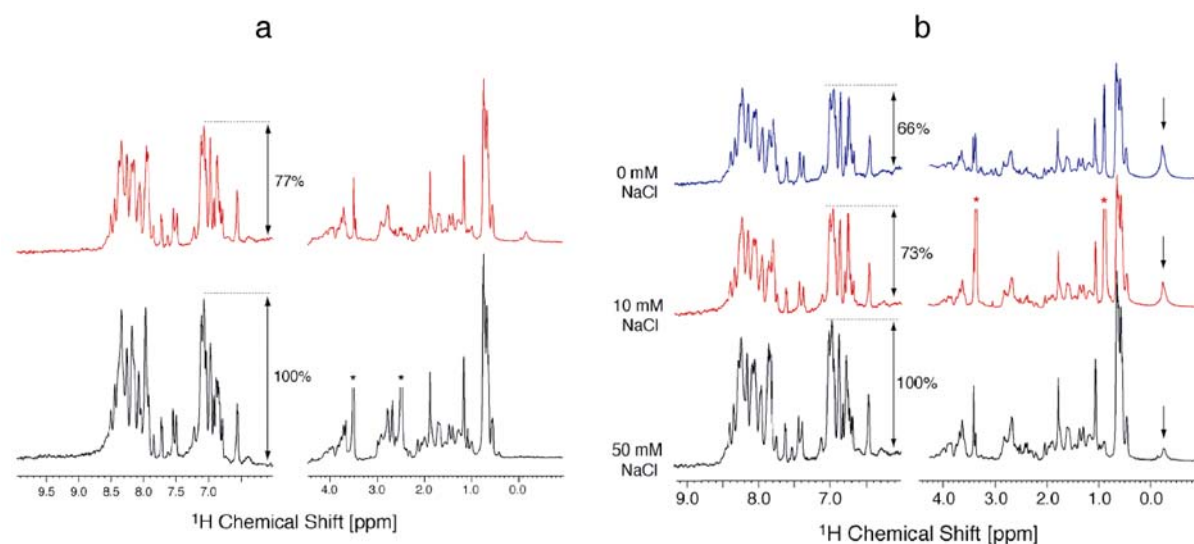
Figure 4.8. a) Comparison of 1D NMR spectra of Amyloid peptide in chloride buffer (red) and sulphate buffer (black); b) Decreasing the concentration of NaCl reduced the intensity of soluble Amyloid and increased the intensity of broad peak at -0.2 ppm

For the chloride buffer, the molecular weight was fitted to 24 kDa (Fig. 4.9). The broad resonance peak at ~ 0 ppm, which is populated in chloride buffer (Fig.4.8) can be fitted to a molecular weight of >100 kDa. We show below that this resonance is related to Aβ(1-40) and we assign this molecule to a critical oligomeric species of Aβ(1-40).

STD experiments were performed to confirm that the broad peak is indeed related to Aβ(1-40). By saturating the oligomeric or fibrillar species, the NMR peak intensity of soluble Aβ(1-40) was affected. This clearly indicates that the broad peak belonged to the higher order oligomers of Aβ(1-40) that are in exchange with soluble Aβ(1-40). The above conclusion was made, after performing additional STD experiments, where the STD intensity was reduced when resonance saturation was not exactly "on-resonance" to the broad oligomeric peak (Fig. 4.10). This supported the observed fact that the STD intensity would depend upon the employed salt conditions, (salts proved to influence the population of oligomeric states (see above)) as in case of Cl⁻, higher STD intensities were observed as compared to the STD experiments on SO_4^- (Fig. 4.11)
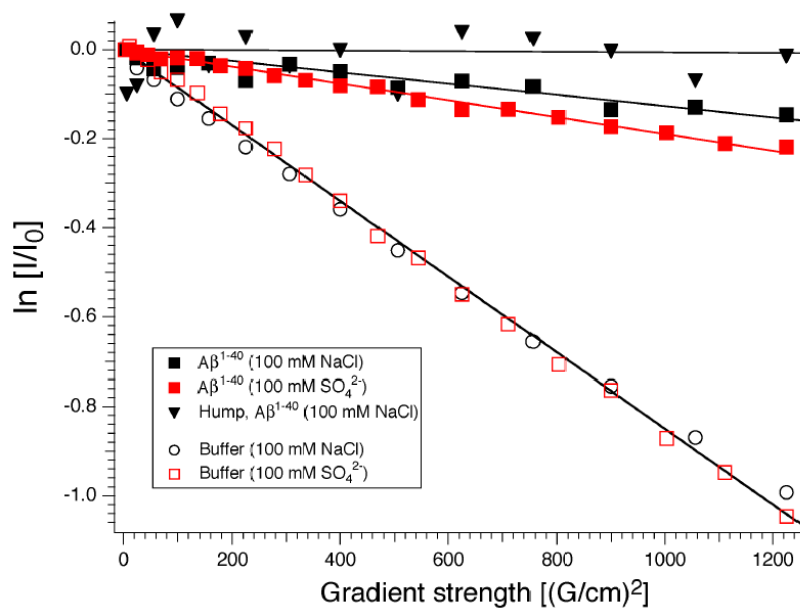
Figure 4. 9. DOSY NMR for Amyloid peptide dissolved in chloride and sulphate.
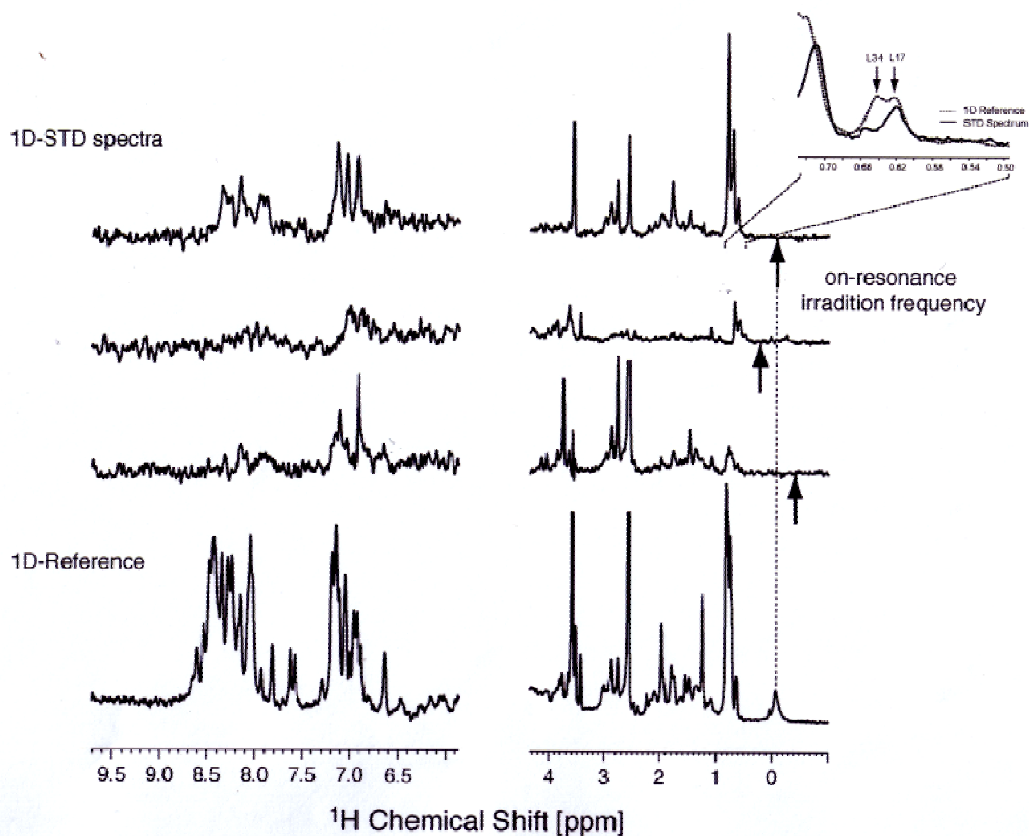
Figure 4. 10. a) STD intensity of soluble amyloid species increased upon irradiating the broad hump.

Also, the interaction appeared to be specific as we observed higher STD intensity for Leu17 compared to Leu34 resonance (Fig.4.10, inset).
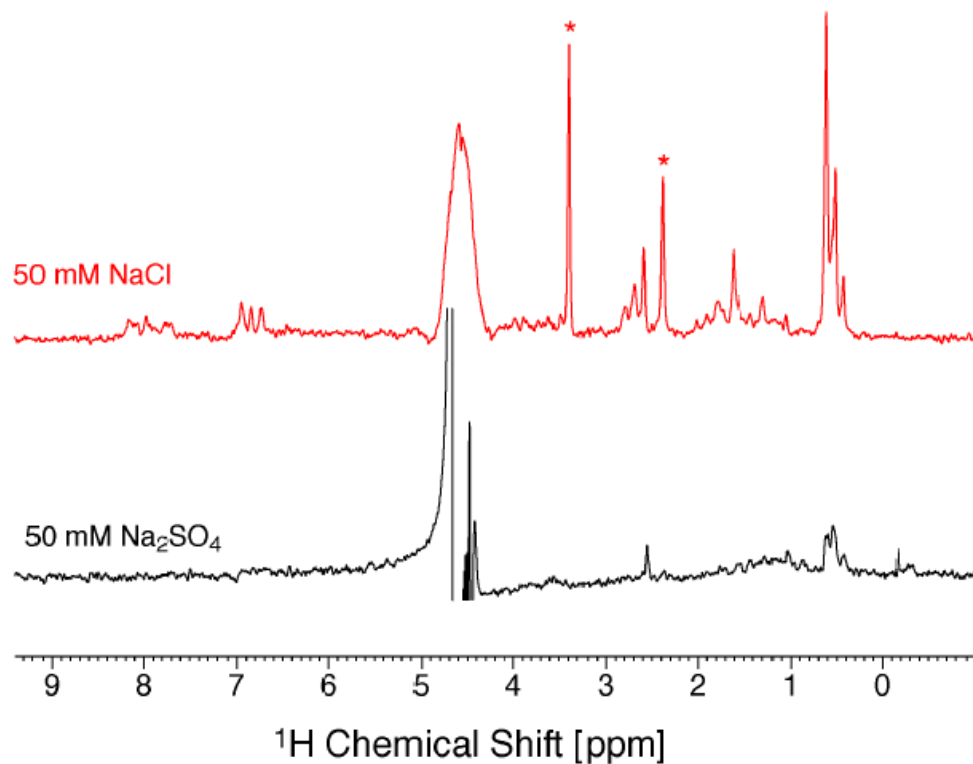
Figure 4. 11. STD intensity of amyloid increased when the amyloid peptide dissolved in sodium chloride

In order to identify the chemical groups involved in Aβ(1-40)·Aβ(1-40) contacts, STD-TOCSY experiments were carried out.

All residues that contributed to the Aβ(1-40)·Aβ(1-40) contacts are highlighted in the primary sequence:

H$_2$N-D^1-EAFRHDSG**Y**^{10}-EVHH**QKLVFF**^{20}-**AEDV**GSNKGA^{30}-**II**GLMVGGVV^{40}- -COOH

Except for the Y10, V18, I31/I32, and the buffer resonances, no STD intensity was detected for the cross peaks (Fig. 4.12 and Fig.4.13).
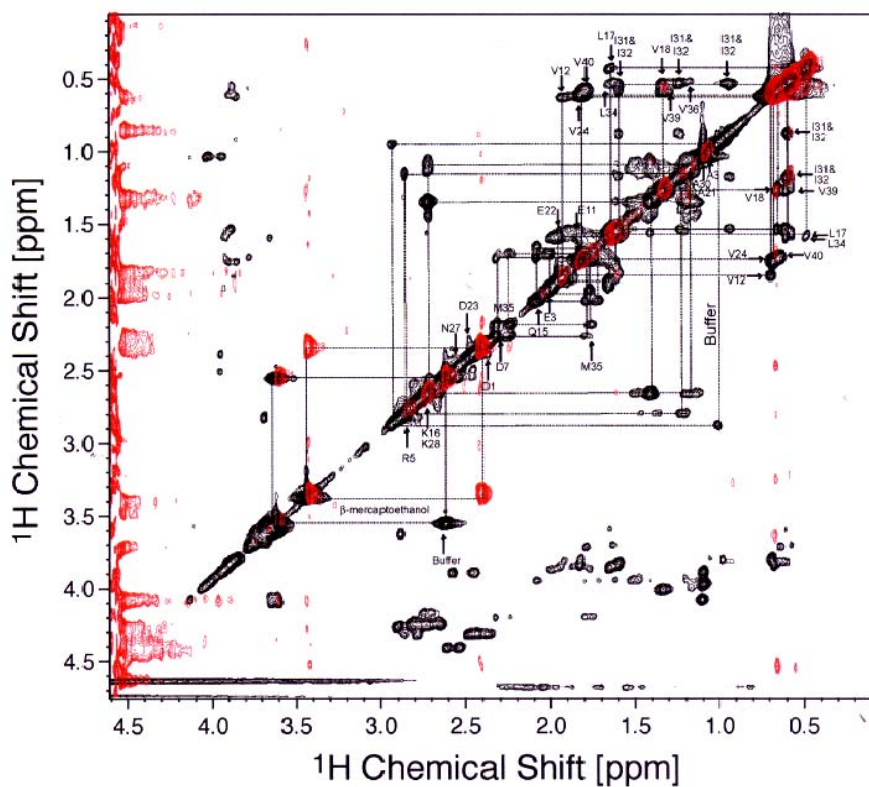
Figure 4.12. Aliphatic region of STD-TOCSY spectrum of soluble Aβ(1-40) (Red) superimposed with Reference spectrum (Black)
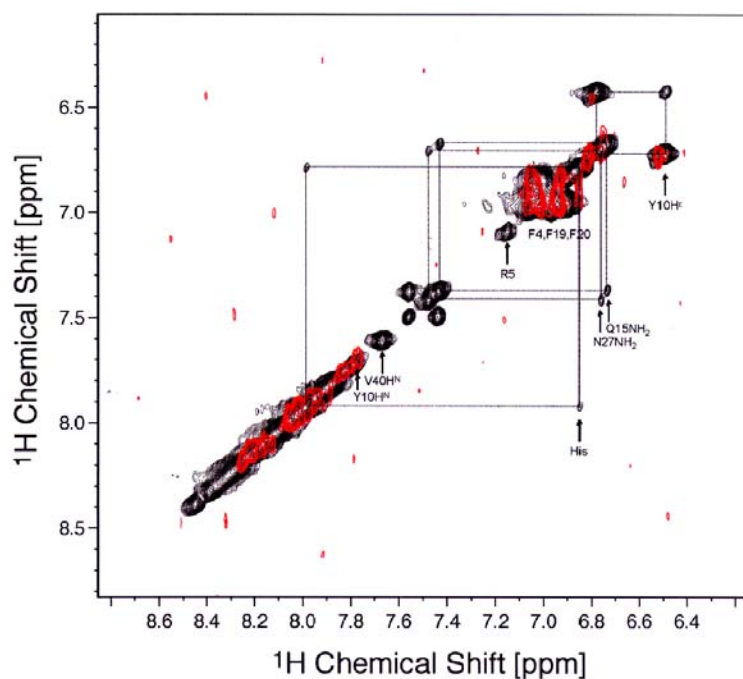
Figure 4.13. Aromatic region of STD TOCSY spectrum of soluble Aβ(1-40) (Red) superimposed with Reference spectrum (Black)

Hence, unambiguous assignment of all chemical groups was difficult to obtain. But, careful analysis of STD-TOCSY spectra revealed that the region around the hydrophobic core (residues 15-24) contributes more to the Aβ(1-40)·Aβ(1-40) contacts (Fig. 4.12). The assumption that hydrophobic interactions govern Aβ(1-40) aggregation is further supported by temperature dependence experiments. The intensities of Aβ(1-40) in 1D-^{1}H NMR resonances were decreased at elevated temperatures (27° C), an effect which was found to be reversible (Fig. 4.14). This result can be explained by the fact that the increased temperature enhances adsorption (*38*). As hydrophobic interactions are also governed by the adsorptive forces, Aβ(1-40) association is favoured at elevated temperature.
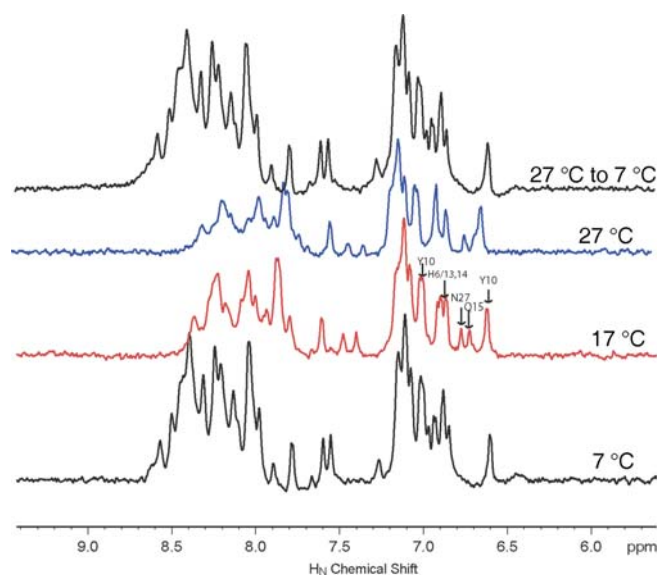
Figure 4.14. Temperature-dependent studies revealed that the fibrils were formed by the contributions from hydrophobic core in the amyloid peptide

It was shown already that the anions induced structures into the previously unstructured proteins (*39*). But, from our experiments, it became clear that the anions also influence the aggregation pathway. How can we explain this? In the absence of salt, the CD spectrum showed a transition from unstructured to the beta sheet conformations (Fig. 4.7c). At the same time no significant change in the H^N H^α chemical shift pattern in the presence and absence of salt was observed (Fig. 4.15 a). The only explanation can be that the population of higher order oligomeric states are not detectable by solution-state NMR due to molecular weight restrictions. TOCSY spectra recorded on Aβ(1-40) in presence and absence of NaCl, showed chemical shifts perturbations for Q15,N27, H14 (Fig. 4.15). Interestingly, again the

same stretch of amino acids, H13, H14,Q15,V24,S26,N27 were affected when Aβ(1-40) was dissolved in sulphate buffer (Fig. 4.15b).
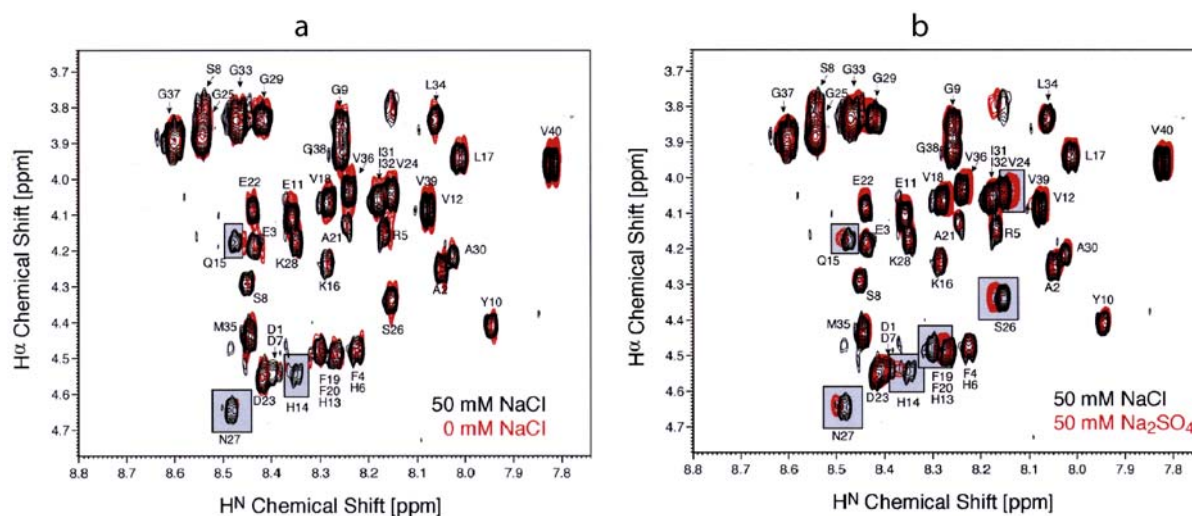
Figure 4.15. a) Finger-print region in TOCSY spectra of amyloid with and without chloride buffer; b) Finger-print region in TOCSY spectra of amyloid with chloride and with sulphate buffer

From the structural biology point of view, the structure and assembly of Aβ(1-40) govern the nature of interaction with its partners. From all the available structures, it is clear that the secondary structure of Aβ(1-40) is highly sensitive to buffer conditions. In 60% TFE-water mixture, Aβ peptides form two helices involving Ala[2]-His[6] and Glu[11]-Ala[42] and adopt helix-turn-helix conformation (*40*). In 40% TFE- water mixture, considerably shorter α-helices extending from Gln[15]-Asp[23]   and Ile[31]-Met[35] were determined (*41*). In SDS solvation condition, formation of two α-helices ranging from Ala[2]-Glu[11] and His[13]-Asn[27] that form a helix-loop-helix was observed (*42*). Recent structure from solid-state NMR of Aβ(1-40) fibrils indicated the existence of two beta-strands, Val[12]-Val[24] and Ala[30]-Val[40] and a loop region, G[25]-G[29] (*43*). Comparing the Aβ(1-40) structures in all these conditions, it is very clear that the central hydrophobic core regions, Gln[15]-Asp[23] and Ile[31]-Met[35] formed helices under all solution conditions employed. Also, in both membrane-mimic condition and in fibrillar state the presence of a mobile hinge comprising of Ser[26]-Gly[29] was revealed. From the above model, it is evident that the C-terminus of the peptide folds back onto the hydrophobic core (Gln[15]-Gly[25]). Taking above facts into account, in the presence of anions, the loop region is affected. Hence, it can be concluded that the anions bound to the positively-charged side chains of the Aβ(1-40), modulate the aggregation pathway.

4.4.2   Interactions of αB-crystallin and Aβ(1-40) and consequences

In order to understand the interactions between the two partners, it is essential to know their localization within the cell compartments. Aβ(1-40)-producing APP is abundant in neurons, astrocytes and microglia. The extra-cellular Aβ(1-40) secretion is abundant in neurons rather than in astrocytes, but intracellular Aβ (1-40) secretion is greater in astrocytes (*44*). Apart from lenticular tissues, the cytosolic αB-crystallin is localized mainly in astrocytes and microglia (*45*). Notably, increasing evidences suggest that the reactive astrocytes rather than neurons seem to play a critical role in neuro-degeneration (*19, 46, 47*). The intracellular localization of αB-crystallin and the Aβ(1-40) is vital for the biological function. The first evidence came from the study where all the AD patients consistently showed localization of both αB-crystallin and Aβ(1-40) within the same intracellular compartment of eye lenses and eventually develop supra-nuclear cataracts (*28*).
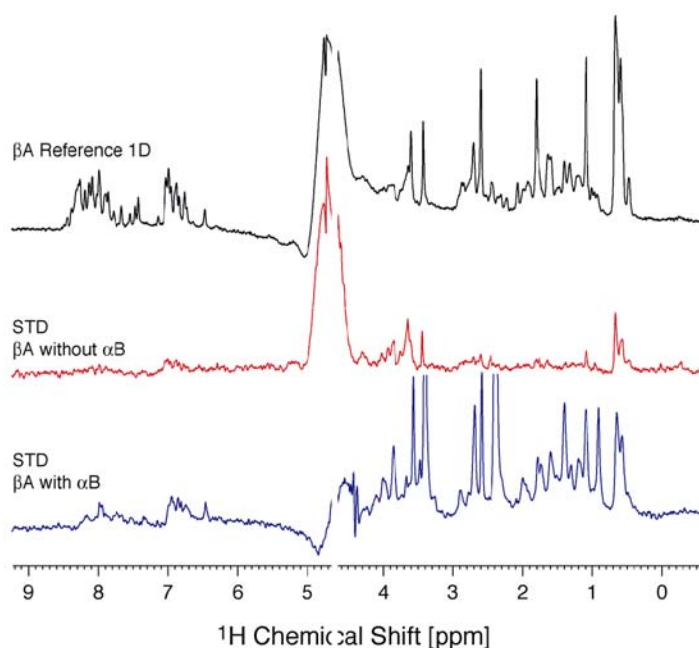
Figure 4.16. 1D STD experiment on Amyloid peptide with and without αB−crystallin in chloride buffer

 What can we say about the structural basis of the interactions of Aβ(1-40) with αB-crystallin? The 1D – STD NMR of Aβ (1-40) together with αB-crystallin, showed a larger STD attenuation compared to Aβ(1-40) without αB-crystallin (Fig. 4.16). Also, the Aβ(1-40) sample which was dissolved in chloride buffer showed higher STD intensity than the one dissolved in sulphate buffer, though the interacting pattern was similar. In order to identify the

chemical groups that are involved in interactions between Aβ(1-40) and αB-crystallin, STD-TOCSY experiments were performed. The contributions from Aβ(1-40)·Aβ(1-40) interactions were subtracted prior to analysis. This way, the exact contributions from the chemical groups can be mapped. Our results show that Aβ(1-40) uses the same stretch of hydrophobic sequence that is involved in fibrillar assembly, to interact with αB-crystallin (Fig.4.17).
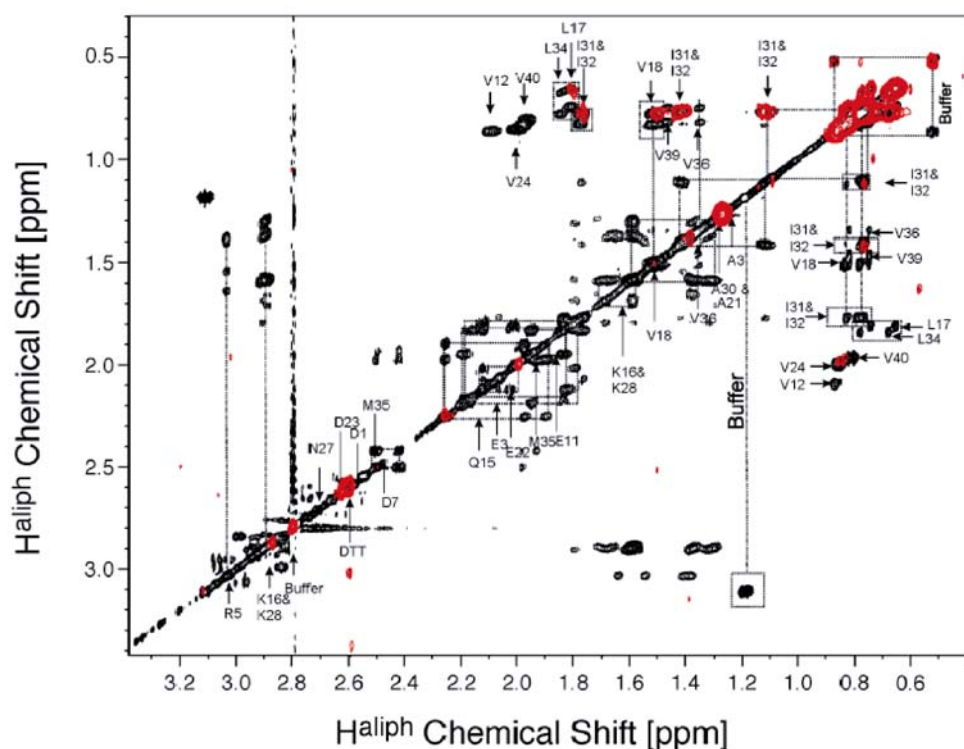
Figure 4.17. STD TOCSY spectrum of Amyloid with αB–crystallin in sulphate buffer

The interacting amino acids are highlighted in the primary sequence of Aβ(1-40).

D[1]-EAFRHDSGY[10]-EVHH**QKLVFF**[20]-**AEDV**GSNKGA[30]-**II**GLMVGGVV[40]
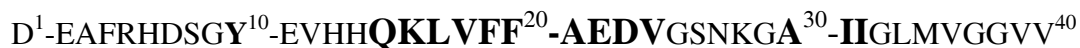
Ambiguities persisted in assigning the resonances that showed STD signals (ref. Section 4.4.1). Therefore, to confirm that the interactions were indeed mediated through hydrophobic residues, Isothermal titration calorimetry experiments were performed by titrating αB-crystallin against Aβ(1-40) solution. The observed positive binding enthalpy in ITC experiments supported the STD-TOCSY results that the interaction between Aβ(1-40) and αB-crystallin is mainly governed by hydrophobic residues (Fig. 4.18, blue).
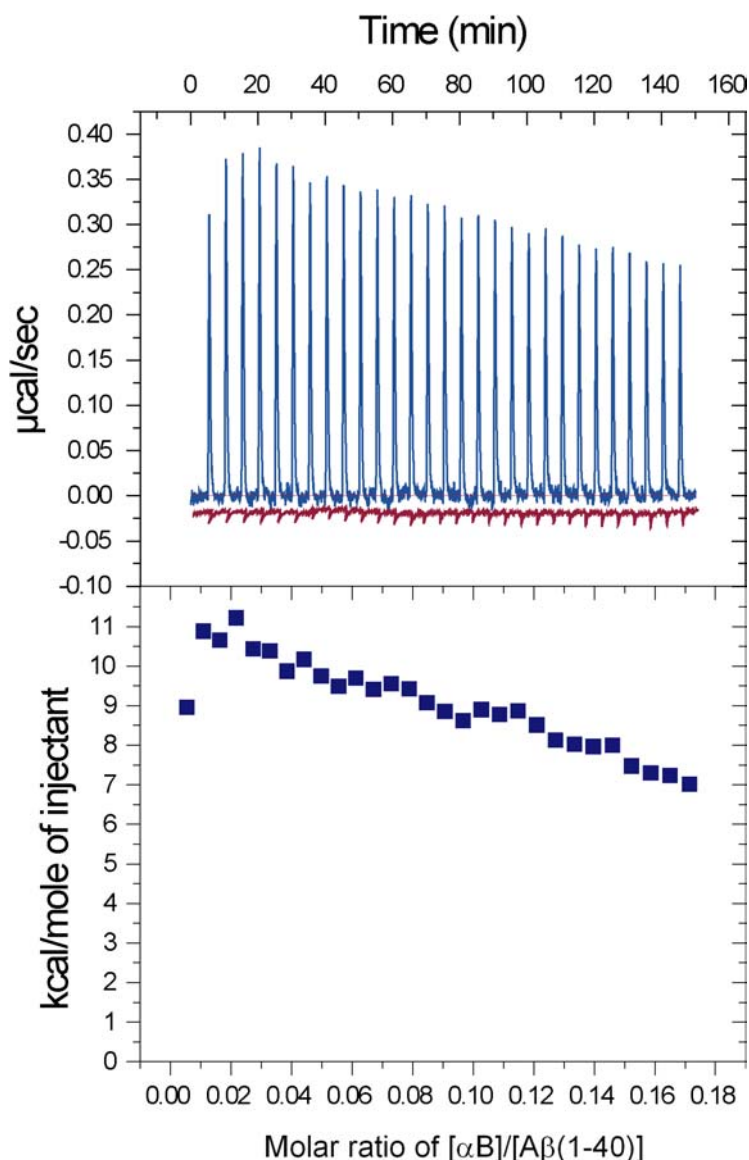
Figure 4.18. Isothermal Titration calorimetry titration of αB–crystallin against Aβ(1-40) (blue curve). The heat of dilution is represented in red .

ITC measures directly the energy associated with a chemical reaction triggered during interactions of a protein with its substrates. The heat absorbed or released during the titration can be directly correlated to the amount of the bound substrates (*48*). Experiments were performed at NMR concentrations. Hence catalytic amounts of αB-crystallin at a molar ratio of [Aβ(1-40)] : [αB] = 25:1, were used. The observed positive binding enthalpy indicates that the interactions were indeed mediated through hydrophobic amino acids. A correlation between positive binding enthalpy and the hydrophobicity has been published and this relationship can be modulated by temperature (*38*). Increasing the temperature causes an increase in the adsorption isotherm and subsequent water desorption. The temperature, free energy, entropy and partition coefficient are interrelated. Temperature causes increase in

99

entropy and partition coefficient of hydrophobic amino acids. This combined effects led to the negative free energy and the reaction was favourable. In the ITC experiments, the positive entropy was observed even at lower temperature (7° C). Therefore, taking into consideration the significant endothermic profile (ΔH>0) attained from the ITC experiments and site specific NMR binding results, we confirm that the binding of αB-crystallin and Aβ(1-40) is mediated through the hydrophobic core of Aβ(1-40).

In order to investigate the consequences of the interactions, DOSY NMR experiments were carried out. It was observed that oligomeric states of Sup35p[5-26] were modulated by Hsp104, a molecular chaperone (Chapter 3). But, in case of Aβ (1-40) the diffusion constant did not change in presence of αB-crystallin as a function of time (Fig. 4.19).
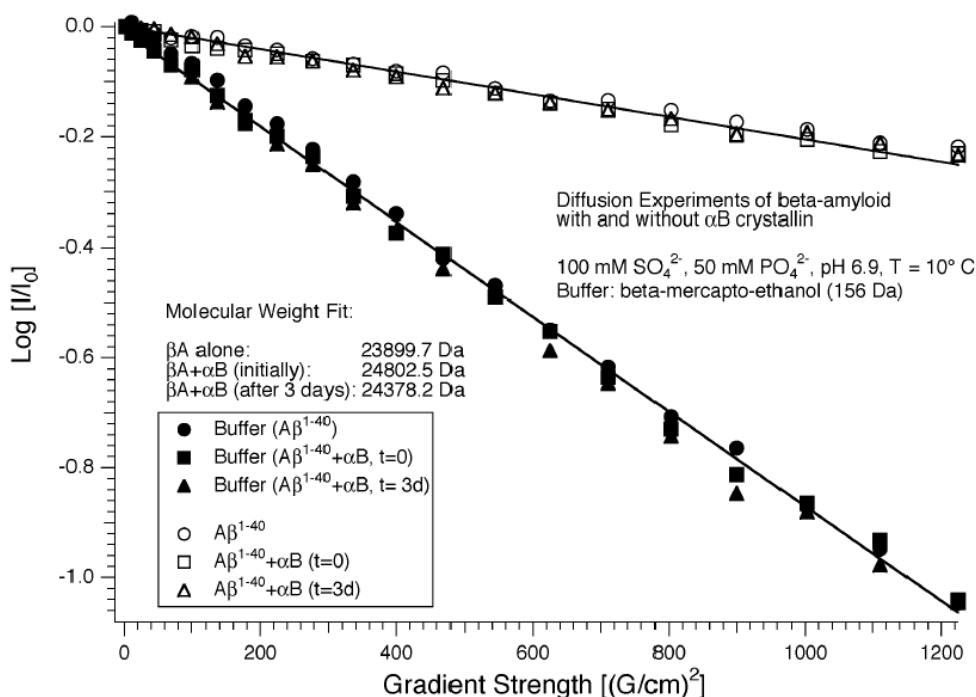
Figure 4.19. DOSY NMR studies on αB–crystallin with amyloid peptide as a function of time

To better understand the consequence of the interactions of αB-crystallin with Aβ(1-40), change in the secondary structure of Aβ(1-40) as a function of time was monitored using far-UV CD spectroscopy. From the previous CD analysis using various salt conditions, it is clear that the Aβ(1-40) adopts a helical conformation in aqueous solution at physiological pH value. The helical content was increased as the population of protofibrilar species increased (Fig. 4.7a). Upon incubation of Aβ(1-40) with αB-crystallin, a decrease in the helical

100

propensity was observed for Aβ(1-40) (Fig. 4.20). At the same time, the CD spectrum of Aβ(1-40) or αB-crystallin did not change significantly as a function of time. Drastic changes in ThT fluorescence upon incubation of Aβ(1-40) with αB-crystallin have been reported (*34*). This can only be understood if we assume that a conformational change is induced for the fibrillar or the protofibrilar state of Aβ(1-40), but not for the soluble state.
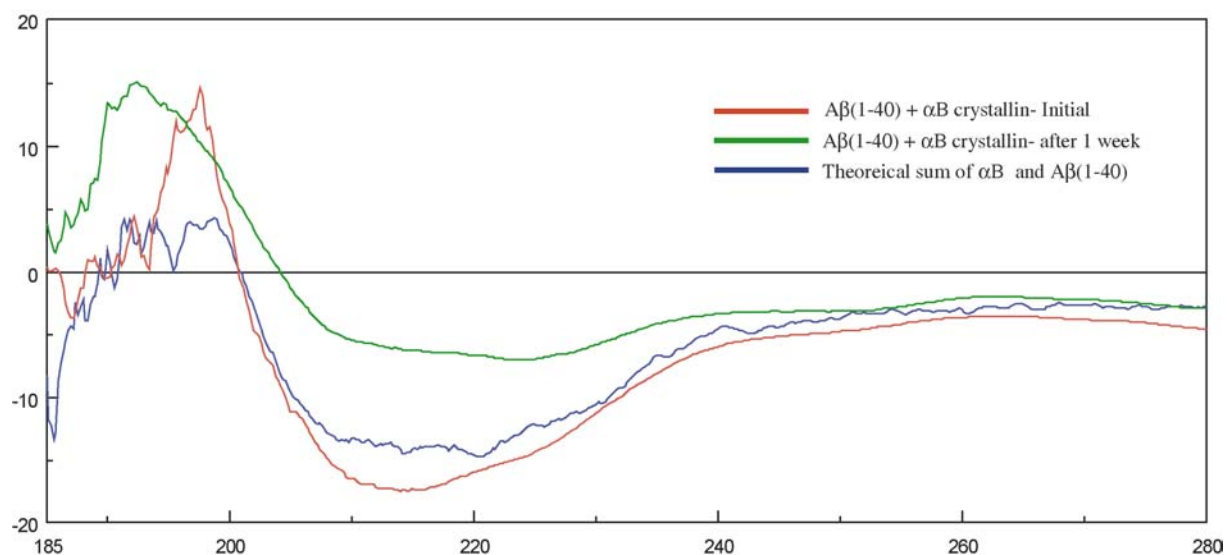
Figure 4.20. Change in secondary structures at different time intervals as monitored by far-UV CD spectra of samples containing a molar ratio of 25:1 mixture of [Aβ(1-40)] and [αB-crystallin]

Monitoring the NMR resonances in a 1D-^{1}H NMR spectrum over a period of several days, we observed that the rate of the auto-oxidation of β-mercapto ethanol (β-ME), (used initially as an internal standard for the DOSY-NMR experiments) was increased by a factor of two at low temperature (7° C) and a factor of five at higher temperature (25° C) in the presence of αB-crystallin (Fig. 4.21). This led us to the investigation of the role of the redox potential of αB-crystallin and its implications during the interactions with Aβ(1-40).

101
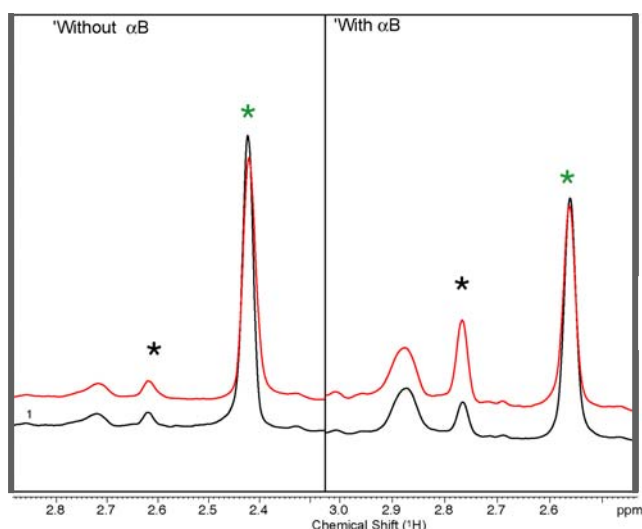
Figure 4.21. Mercapto ethanol oxidation as observed by NMR for the samples in presence and absence of aB-crystallin. The green asterisks represent resonances of non-oxidized β-ME, black represent oxidized form of β-ME. The final data (red spectra) were recorded after 20 hours and 30 hours of incubation, for the mercapto ethanol sample in presence and absence of aB-crystallin, respectively

Taking advantage of molecular oxygen under aerobic conditions, a NMR-based assay was developed to compare the redox activity of proteins. The auto-oxidation of monomeric glutathione, GSH, to the dimeric form, GSSG, was monitored by solution-state NMR spectroscopy. This could be a useful assay for other proton-driven electron transfer reactions. This simple method allows us to qualitatively determine the redox nature of various proteins with respect to the redox potential of GSH. If the protein possesses reducing activity compared to GSH then, GSSG forms very slowly. In order to test the system, we used ascorbic acid as a model system. We observed a retardation of the auto-oxidation of GSH in the presence of this reducing agent (Fig. 4.22).
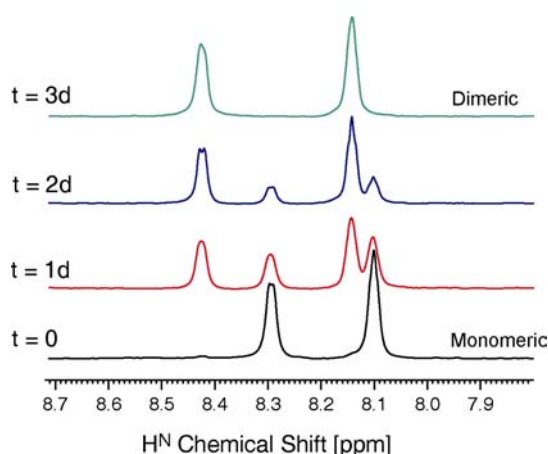
Figure 4.22. Monitoring auto-oxidation of GSH in presence of foreign substance as a function of time

The addition of catalytic amounts of αB-crystallin significantly influenced the redox activity of the GSH (Fig. 4.23). Similar measurements were performed on Aβ(1-40) alone, and it was found that the redox nature of GSH was not altered. This could be interpreted as, the redox property of Aβ (1-40), can be close to GSH or it does not possess any redox activity. However, a mixed sample of Aβ (1-40) and αB-crystallin displayed an effect on the redox nature of GSH. As Aβ (1-40) alone did not affect GSH auto-oxidation, the same curve that was observed for αB-crystallin alone should be expected for the mixture. The observation of an intermediate effect can only be explained if cooperative nature of the complex of Aβ (1-40) and αB-crystallin is taken into account (Fig. 4.23).
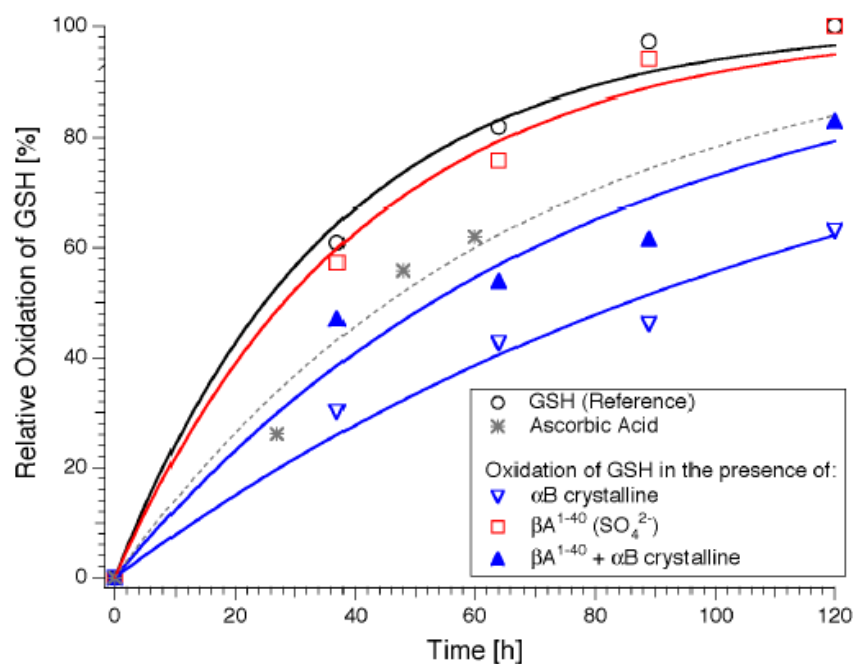
Figure 4.23. NMR based redox assay for the Amyloid(1-40) and αB–crystallin and the mixture.

Now, it is inevitable to confirm whether Aβ(1-40) possesses any redox activity or not. Previous reports show that Aβ(1-40) can reduce metal ions (*49*). The trapped free electrons that are generated from Aβ(1-40) can also be observed via ESR spectroscopy (*50*). We therefore, wanted to study the redox properties of various oligomers of Aβ(1-40) in detail. Various oligomers of Aβ(1-40) were generated by varying the hardness of the anions as described in the Section 4.3 (*51*). An *in vitro* assay was designed to monitor the redox activity, which is based upon redox mediated colour change. The *Uptiblue* reagent was used

for this purpose. *Uptiblue* is a resazurin based sensitive dye, originally designed to sense the electron transport of mitochondria in a living system. *Uptiblue* changes colour from blue to red upon reduction. The advantage of the system is that the end products are very stable and that *Uptiblue* provides the possibility to compare the redox activity of various proteins.
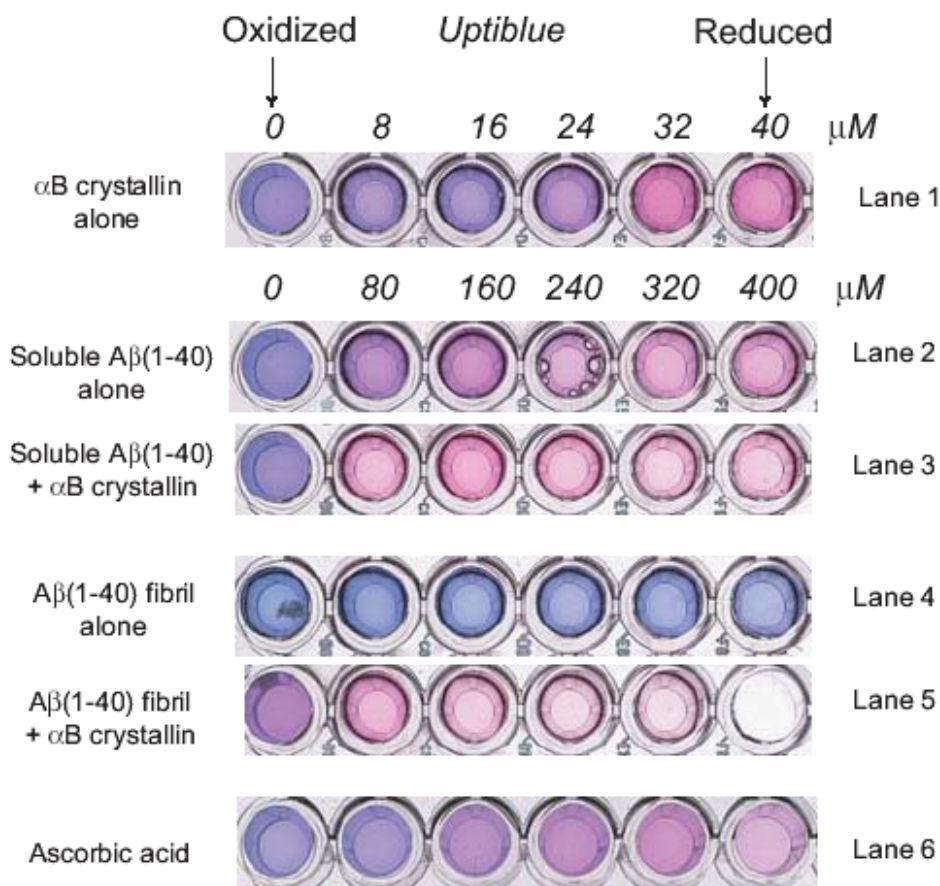
Figure 4.24. Colorimetric assay for monitoring the redox activity of αB-crystallin and Aβ(1-40)

The colorimetric assay (Fig. 4.24) confirmed that αB-crystallin induced a fast reduction of *Uptiblue* (lane 1). This behaviour is in agreement with the GSH auto-oxidation data from the NMR experiments (Fig. 4.23). Aβ(1-40) reduced *Uptiblue* only at much higher concentrations (concentration ranging from 0 to 400 μM) (lane 2). This proved that Aβ(1-40) also possessed a redox activity. In the presence of αB-crystallin (at a molar ratio of 1:10 with respect to Aβ (1-40)), the redox potential turned out to be even more positive (lane 3), since the red shift induced by the reaction, occurred at lower concentrations of Aβ(1-40) and αB-crystallin. Interestingly, fibrillar Aβ (1-40) seemed to have a redox potential larger than 380

mV, since it did not reduce *Uptiblue* even at very high concentrations (lane 4). This situation was changed, once αB-crystallin was added (lane 5). The colorimetric effect was similar to the case in which non-fibrillar Aβ (1-40) was mixed with αB-crystallin, indicating that αB-crystallin interacts with the fibrillar form of Aβ(1-40) (Fig. 4.24).

This observation explains the results obtained from CD spectra that the observed changes in secondary structural elements could be due to the higher ordered aggregates that are not observable by NMR spectroscopy (Fig. 4.20). The redox experiments suggest that αB-crystallin possesses redox activity. This redox activity is well reflected in the NMR spectra of Aβ(1-40) which was incubated with αB-crystallin for 3 days. Careful inspection of the correlation NMR spectra indicates that αB-crystallin populated the oxidized form of Met-35 in Aβ(1-40) (Fig. 4.25).
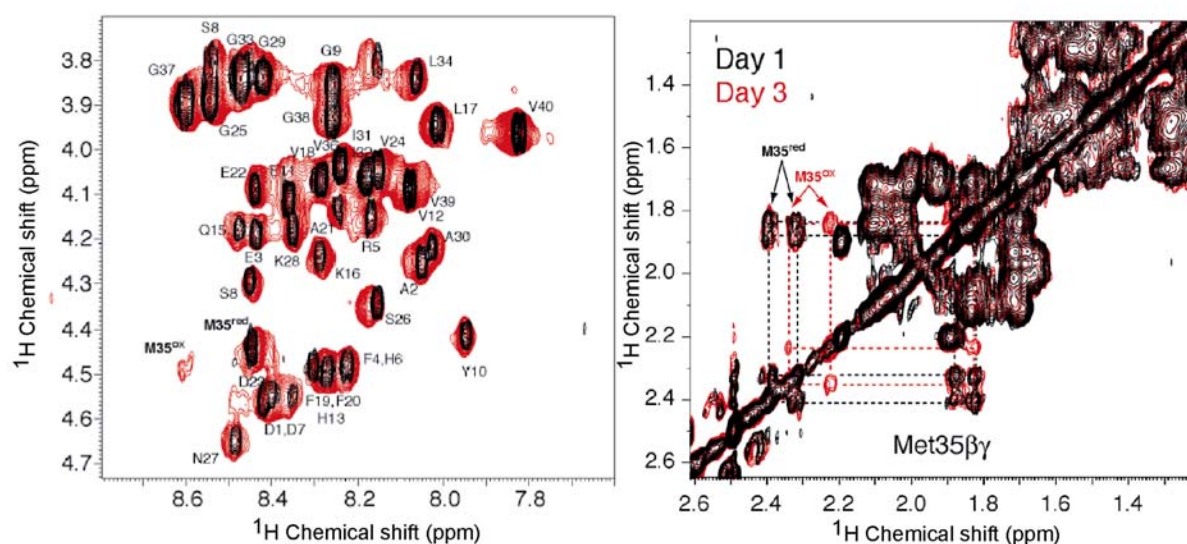
Figure 4.25. Oxidation of Met-35 residue in amyloid - consequence of the interactions between amyloid and crystallin

The assignment of the chemical shifts was corroborated by a recent study which had been carried out by Zagorski and co-workers (*52*). Interestingly, our STD correlation experiment does not demonstrate any interaction between αB-crystallin and Met35 of Aβ(1-40) (Fig.4.17). This observation indicates that αB-crystallin changes the redox nature of Aβ(1-40). The increased difference in redox potential of Aβ(1-40) with respect to molecular oxygen, facilitates the auto-oxidation of Met-35 of Aβ(1-40).

From *Uptiblue* experiments, we could show that αB-crystallin possesses redox activity. The importance of redox chemistry originates from the transition metals, copper and iron. Many neurodegenerative diseases including Alzheimer's disease implicate the importance of copper connections (*53*). Familial amyotrophic lateral sclerosis (ALS) is a motor neuron disease that occurs mainly due to the inactivation of the radical scavenging superoxide dismutase-1 protein (SOD1), which transports copper across mitochondria (*54*). It has been demonstrated that in ALS patients, the sHsp proteins, Hsp25 and αB-crystallin are up-regulated and specifically co-precipitate with insoluble SOD1 (*55*). Involvement of copper and crystallins in cataractogenesis (*56*) and modulation of the chaperone activity of crystallin in the presence of transition metal ions (*57*) are other hints for the importance of copper in modulating αB-crystallin activity. However, until now the influence of copper on αB-crystallin and its implication upon substrate binding are not known.
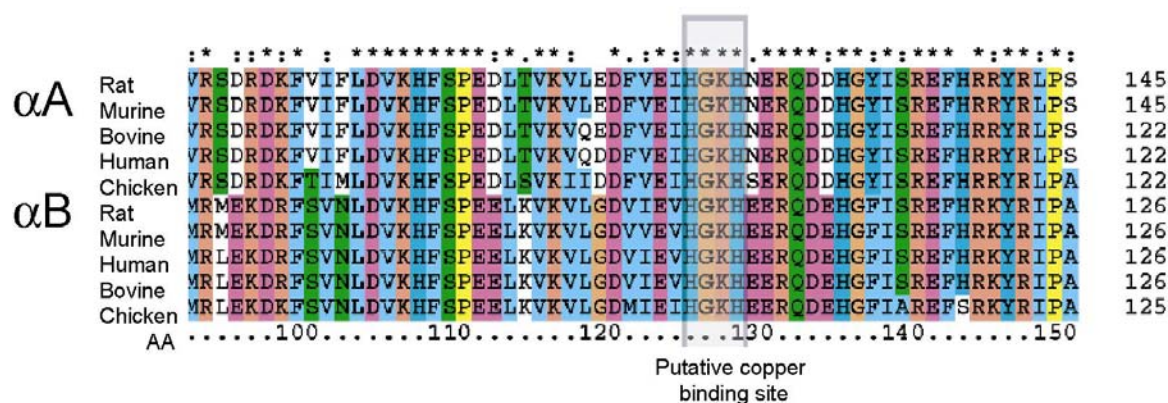
Figure 4. 26. Crystallin family showed conserved HXXH motif

Careful inspection of the primary structure of αB-crystallin shows that αB-crystallin contains an amino acid stretch HXXH (residues 126-129) which is also found in the SOD1 copper chaperone CCS, there it is a CXXC motif (*58*). Notably, this putative copper regulatory motif is highly conserved among all αB-crystallin and αA-crystallin of mammalian origin (Fig. 4.26). A further interesting result came from the NMR experiments, in which copper was added to the GSH and αB-crystallin solution. Upon addition of copper, we observed a strong influence on the redox potential of αB-crystallin. Addition of Cu (II) to GSH should slow down the oxidation of GSH in the presence of αB-crystallin (Fig. 4.27). This can be interpreted in two ways: Either copper directly influences the redox pathway of GSH or changes the redox activity of αB-crystallin.
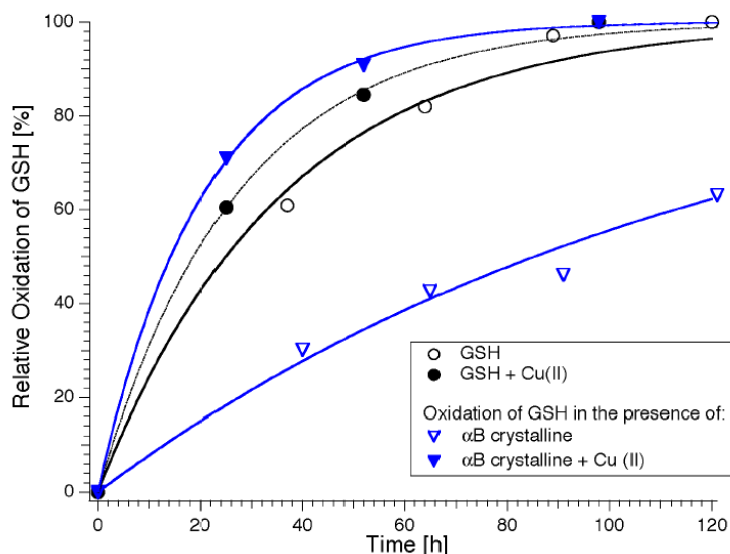
Figure 4. 27. Copper influenced the redox activity of αB–crystallin as observed from NMR based assay

Cu(II) alone does not affect the auto-oxidation of GSH. We conclude therefore that Cu(II) binding to αB-crystallin alters the redox potential of the complex. The above results are confirmed using the colorimetric redox activity assay described above. In the presence of Cu(II), the reduction of *Uptiblue* was dramatically reduced (Fig. 4.28).
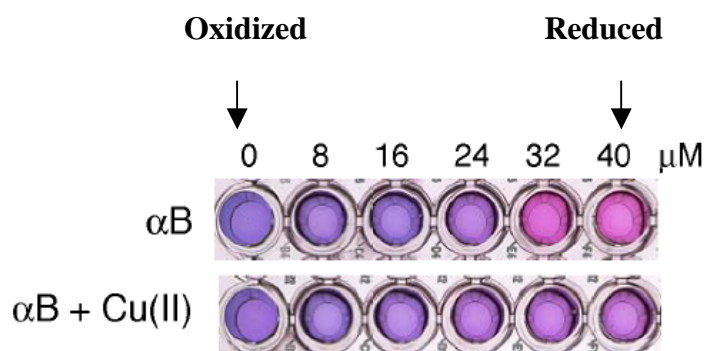
Figure 4.28. Colorimetric assay confirmed the influence of copper on redox activity of αB–crystallin

In order to understand the influence of Cu(II) on substrate binding, we performed ITC experiments. In the experiment shown above, αB-crystallin was titrated into a Aβ(1-40) solution (Fig. 4.29a). We could show that the binding of Aβ(1-40) to αB-crystallin is driven hydrophobically. As we use catalytic amount of αB-crystallin, the endothermic profile of binding was observed upon all the injections of αB-crystallin. In the following experiments, the influence of copper upon binding to αB-crystallin is investigated. In the experiment, αB-crystallin was pre-incubated with copper at a molar ratio of 1:2 and the ITC titration was performed by injecting αB:Cu(II) mixture into the Aβ(1-40) solution. The binding profile was

markedly different (Fig. 4.29b). Superposition of four effects might account for this observation: (1) Dissociation of αB-crystallin and Cu(II), (2) Formation of new Aβ(1−40) and Cu(II) complex, (3) Change of the oligomeric state of αB-crystallin and (4) Different affinities of Aβ(1−40) to the Cu(II): αB-crystallin complex. However, the results show that the number of binding sites on the protein are drastically reduced (Fig. 4.29b).
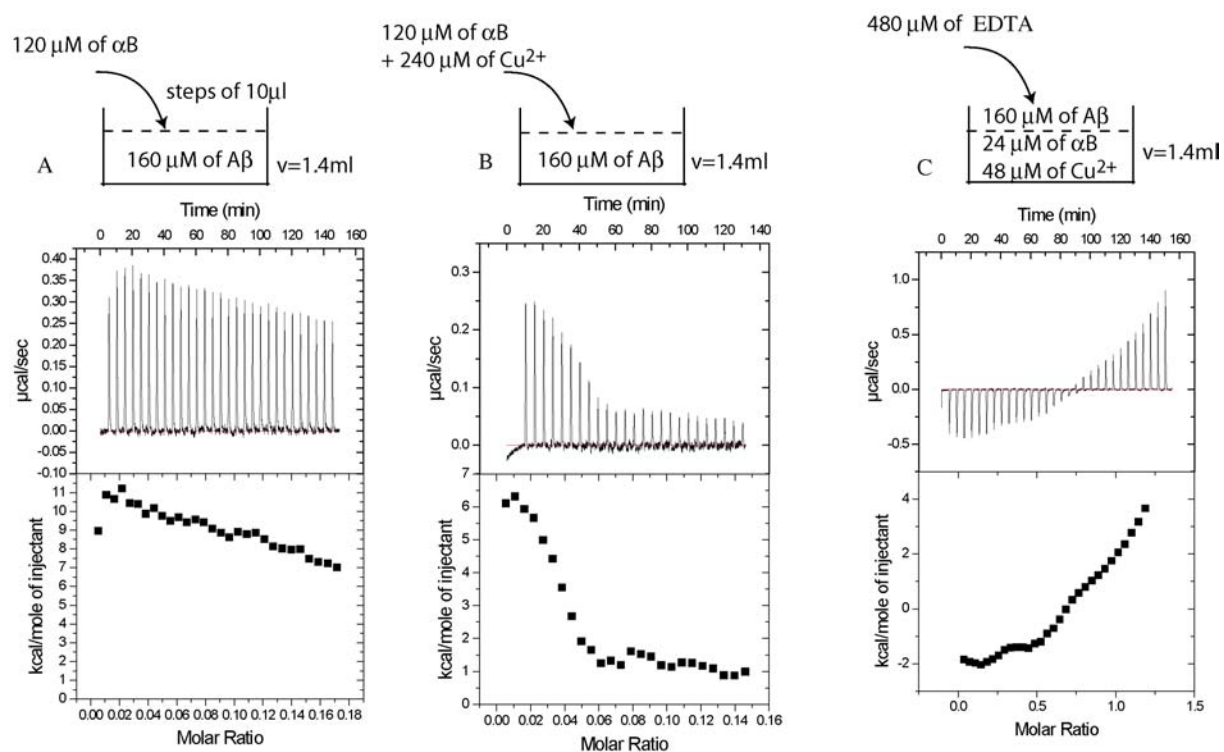
Figure 4.29. a) ITC experiments showed that the binding enthalpy between αB–crystallin and amyloid is endothermic, implied hydrophobic binding; b) Copper addition to αB–crystallin reduced the binding sites in the protein as ITC experiments observed reduced binding affinity; c) Hydrophobic binding restored once the copper was removed from the mixture by EDTA

The dissociation of the complex αB-crystallin ·Cu(II) is initially exothermic, releasing a heat of ca. −0.30 μcal/sec (Fig 4.30b).

Injection of Cu(II) alone into Aβ(1−40) solution yielded a large endothermic effect which is linearly decaying in the concentration ranges employed in the experiments (+0.13 μcal/sec) (Fig. 4.30a). This endothermic effect is attributed to the induction of aggregation of Aβ(1−40) due to Cu(II) ions.

The sum of the two effects should therefore yield an induced heat of approximately −0.17 μcal/sec at the beginning of the titration. However, an endothermic effect of about +0.35 μcal/sec for the complex formation of αB-crystallin and Aβ(1−40) was observed (Fig. 4.29a).

The difference (+0.52 µcal/sec) cannot be explained only by dissociation of αB-crystallin ·αB-crystallin complexes into monomers. The heat of dilution for this dissociation was smaller than −0.02 µcal/sec (Fig 4.18, red). Therefore, the difference must be due to a decreased affinity of the complex of αB-crystallin·Cu(II) to Aβ(1−40), compared to the interaction between αB-crystallin and Aβ(1−40) in the absence of copper.

Alternatively, a model, in which only a certain oligomeric state of Aβ(1−40) is interacting with αB-crystallin is also plausible. The small molar ratio (0.03) of the concentration of αB-crystallin with respect to Aβ(1−40) might be in favour of this interpretation. An unambiguous evaluation of the ITC isotherms is, however, not possible, since the stoichiometry and the dissociation constants are coupled (c= n/K_d) (*48*). However, we can qualitatively state that copper treatment of αB-crystallin yields a decreased binding affinity to Aβ(1-40).
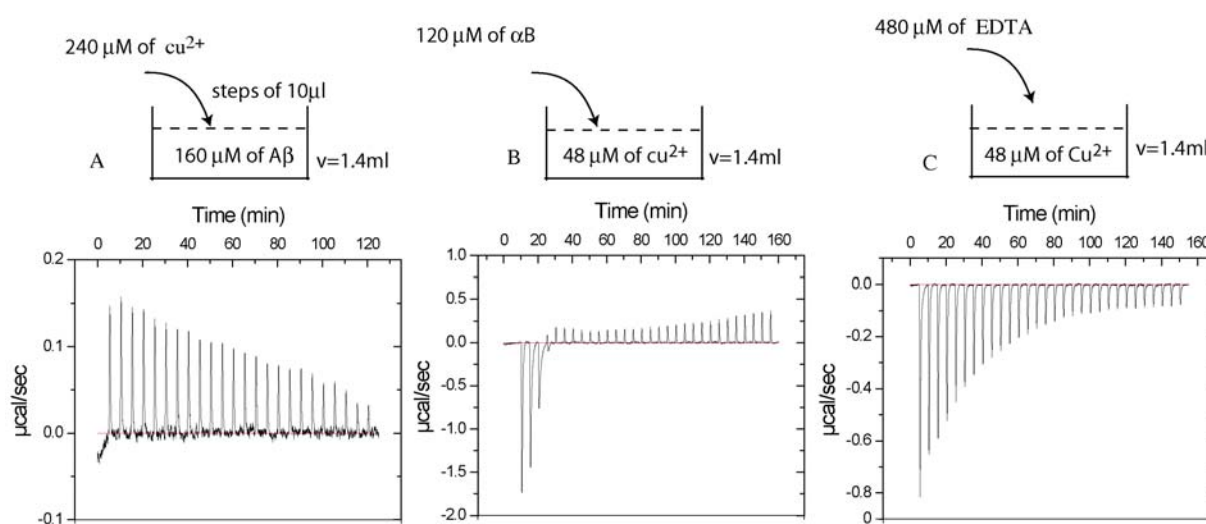
Figure 4.30. a) Titration of copper into amyloid showed slight endothermic profile ; b) titration of αB–crystallin into copper generated heat initially and subsequently aggregated c) Titration of EDTA into copper

It has been shown that copper can be removed from Aβ (1-40) by EDTA even the dissociation constant of copper with respect to Aβ(1-40) is higher (*59*). In order to understand whether the removal of copper from αB-crystallin can restore substrate binding, the αB-crystallin:Aβ(1-40):Cu(II) mixture was titrated with EDTA (final molar equivalent of 2 with respect to copper). Surprisingly, the endothermic substrate binding profile was restored, after removal of copper from the αB(crystallin):Cu(II) complex (Fig.4.29c). Quantitative analysis of the integrated ITC peaks as a function of the molar ratio indicates superposition of several sigmoidal curves that might be due to the binding of EDTA to the Cu(II), dissociation of

Aβ(1-40) oligomeric complexes, and the formation of Aβ(1-40):αB-crystallin complex. ITC titration of EDTA against Cu(II) (Fig. 4.30c) was also performed in order to confirm the substrate specificity.

From the DOSY-NMR experiments, the change in water viscosity upon addition of a catalytic amount of copper into the NMR tube indicates the presence of higher molecular weight species compared to the Apo form of αB-crystallin (Fig. 4.31a). In order to confirm this finding, we performed gel filtration experiments. We found that upon addition of copper into the αB-crystallin, the molecular weight of αB-crystallin is increased. This indicated that copper modulated the quaternary structure of αB-crystallin (Fig. 4.31b).
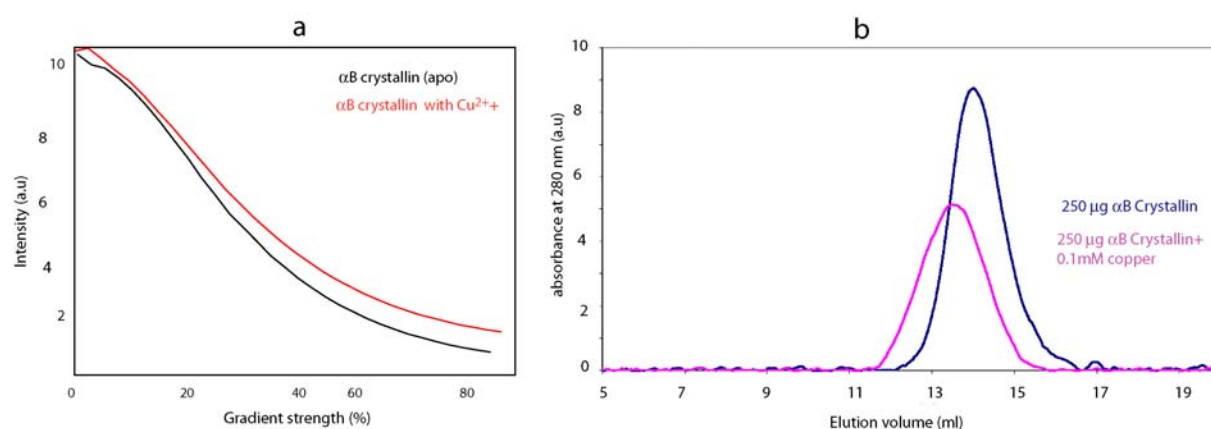
Figure 4.31. a) Copper addition increases the viscosity of water as judged by DOSY; b) Gel filtration studies show that addition of copper shifts the αB-crystallin to the higher molecular weight.

The above, *in vitro* findings were carefully analysed in the *in vivo* conditions (results from Dr. Wilbert Boelens, University of Nijmegen, Netherlands). The expression level of proteins that are induced by over expression of αB-crystallin was monitored. Surprisingly, transcriptional up or down regulation of other redox proteins was observed during the over expression of αB-crystallin. In particular, down-regulation of the cellular quality control protein, disulfide isomerase protein, Erp57, and a down regulation of cellular redox state balancing protein, Glutathione-S transferase was observed. An earlier report also suggests that the protective activity of human αB-crystallin against TNF-α mediated cell death results from its conserved ability to raise the intracellular concentration of glutathione (*60*).

Figure 4.32. Colorimetric redox assay of various oligomeric species of Aβ(1-40). Various oligomers of Ab(1-40) were generated at various anionic buffer conditions. Dutch βA , represents the Aβ(1-40) containing the E22Q substitution. Refer Material and Methods for the redox assay.

What causes the increased neurotoxicity of Aβ(1-40) upon interaction with αB crystallin?

1. All the experimental facts confirm that αB-crystallin possesses a biological redox activity. *in-vivo,* αB-crystallin down-or up regulated various redox proteins. It can be conceivable that the chaperone activity can also be influenced via its redox activity. Also, αB-crystallin may act as a redox regulator in the cellular redox machinery.

2. Is oxidation of methionine responsible for the increased neurotoxicity?
    From the NMR results, we conclude that Aβ(1-40) which is solubilized in chloride buffer generate higher order oligomers compared to the one solubilized in sulphate buffer. It was reflected in the calorimetric assay that, Aβ(1-40) solubilized in sulphate buffer, reduced the *Uptiblue* faster than the one solubilized in chloride buffer (Fig. 4.32). In addition, it was surprising to see that solubilized and fibrillized Aβ(1-40) have different redox potential. From the above facts, we conclude that αB-crystallin is capable to modulate this redox potential of Aβ(1-40) oligomers and/or fibrils (Fig. 4.32).

Figure 4.33. Schematic representation of modulation of redox activity of amyloid induced by αB–crystallin

3. The above scheme summarizes the consequence of the redox potential changes observed in the presence and absence of αB-crystallin (Fig. 4.33). The arrows placed at the far left are indicative for molecules that can give electrons to the molecules that are placed at the right side of the scale. Hence, it is evident that αB-crystallin has the ability to give electrons with respect to Aβ(1-40). However, we observed by NMR experiments that the Met35 of Aβ (1-40) changes its oxidation state upon interaction with αB-crystallin (Fig. 4.25). Notably, from the STD NMR experiments we have no indication of direct interaction between αB-crystallin and Met-35 of Aβ(1-40). The interacting amino acids are, Gln^{15}-Asp^{23} and Ile^{31}- Ile^{32} (Fig.4.17) . We conclude therefore that the oxidation of Met35 in Aβ(1-40) is a consequence of the change in redox potential of Aβ(1-40) .

4. Met-35 in the αB-crystallin·Aβ(1-40) complex has strong tendency to get oxidized compared to the free Aβ(1-40) as the average redox potential of αB-crystallin bound Aβ(1-40) moved away from molecular oxygen compared to the free Aβ(1-40) (Fig. 4.33). Notably, Met-35 was found to be involved in side-chain back bone interactions under non-oxidizing condition (PDB code: 1BA4). Met-35 in oxidized Aβ(1-40) repositions itself and makes contact with the aqueous environment (PDB code: 1BA6) (Fig. 4.34). This causes a change in dipole moment from 1.60 Debye to 5.28 Debye (*61*).
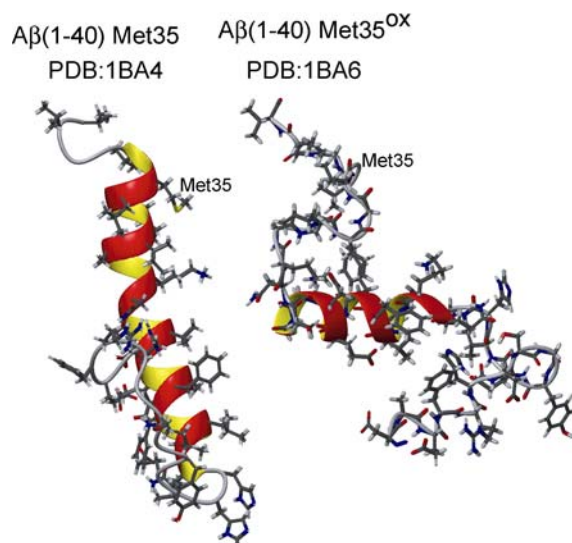
Figure 4.34. Normally buried Met35 in the Aβ(1-40) structure in water-micelle environment(Coles, M., et al., Biochemistry, (1998), 37, 11064) (a) repositions itself in the oxidized form of Aβ(1-40)(Watson, A.., et al.,Biochemistry, (1998), 37, 12700) and contacts aqueous environment due to the change in dipole moment from 1.6 Debye to 5.8 Debye.

5. Earlier reports consider oxidation of methionine residues as an important event in inducing neurotoxicity in Alzheimer's disease. Met35 oxidation was found to be associated with transition metals to serve as an electron source (*62*). The importance of methionine residues continues to grow ever since it had been proposed that they may function as an endogenous antioxidant defence system (*63*) and regulate several cellular functions (*64*). Met35 oxidation increases the polarity of Aβ(1-40) which might act as a switch to populate soluble oligomers through para-nuclei destabilization was suggested earlier (*61*). There are discrepancies regarding whether the reduced or oxidized form of Aβ(1-40) is neurotoxic. Met35(ox) might be more neurotoxic than its reduced form since it can propagate the free radical to adjacent residues (*65*). On the other hand, Met35(ox) is considered to be less neurotoxic compared to Met35 (red) as it destabilizes neurotoxic fibrils and protofibrillar structure (*61*). Confusingly, other investigations report that Met35 (ox) has the same neurotoxic properties as Met35(red), but differs in its lipid membrane penetration ability (*66*).

6. From our experimental data, it can be concluded that oxidation of Met35 of Aβ(1-40) is rather a consequence of the redox properties induced by αB-crystallin than a reason

of neurotoxicity. Accumulation of either soluble or protofibrilar Aβ(1-40) molecules increases the redox burden. This redox burden might provide an explanation for the neurotoxicity. The colorimetric data on soluble and fibrillar Aβ(1-40) indicate that different oligomeric states of Aβ(1-40) have different redox properties. How the redox potential related to neurotoxicity?
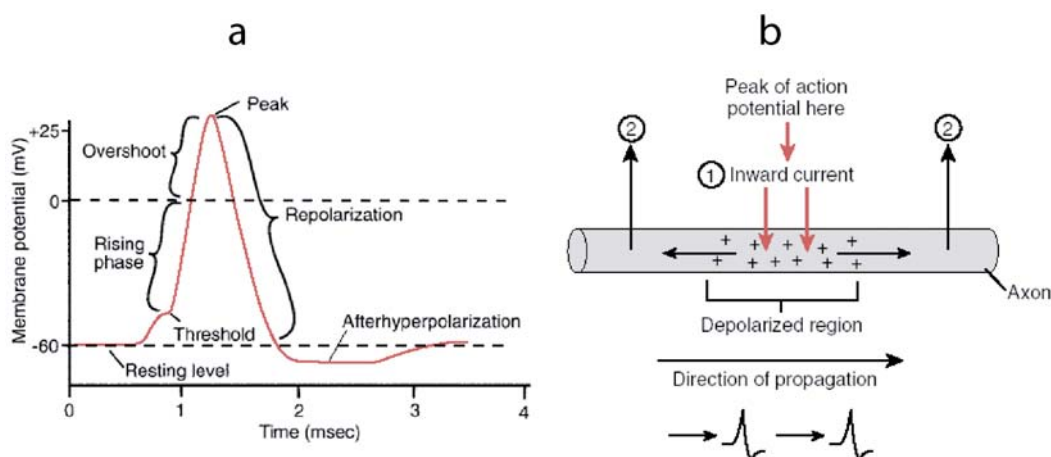
Figure 4.35. a) Neuronal communication is transmitted via action potential (A flow of ions changes the electrical potential difference across the membrane and caused action potential). This mechanism is depicted in three phases: Rising phase, Repolarization and hyperpolarization. Propagation of the action potential either occurs due to continuous channel openings or thorough rapid jumping of action potential from node to node in myelinated neurons (saltatory conduction) as represented in (b)

7. Alzheimer's disease begins with subtle alterations of hippocampus synaptic efficacy prior to neuronal degeneration. It has been shown that the synaptic dysfunction is caused by diffusible oligomeric assemblies of the Aβ peptides (*67*). The efficiency of synaptic connection depends upon the neurotransmitters (chemicals that are responsible for the synaptic transmission) and the propagation of the impulse from one point to another. Within the neurons, the membrane potential is kept at rest at −70 mV (*68*). The propagation of current, the velocity of the nerve impulse or saltatory conduction is varying over a wide range by altering the electric resistance of the external fluid. Thus, above a certain threshold, the action potential shoots up. This influences propagation of synaptic transmission. All the above facts showed that the increased redox burden either secreted or intra-cellular would have an impact in changing the action potential.

8. In addition, we propose that all other events like ROS production, membrane depolarization, mitochondrial dysfunction, metal ion reduction, nitroxyl-ion production and apoptosis may be the consequence of the redox burden caused by

Aβ(1-40). Hence, the increase in redox burden may be considered as an early event in Alzheimer's neurotoxicity.

9. Are Aβ(1-40), αB-crystallin and neurotoxicity connected via copper?

It is known that redox enzymes often use a CXXC (*69*) or CXXS (*70*) motif for metal binding and electron transfer reactions. Many copper chaperones that are actively involved in transport of copper across the mitochondria use a CXXC sequence as copper binding domain (*71*). We postulate that the conserved HXXH motif in crystallins (Fig. 4.26) possesses both copper binding and redox activity. The "Histidine cycle mechanism" exploits the protonation state of the histidine imidazole ring as one step of the mechanisms which governs proton translocation coupled with electron transport(*72*). It has been postulated that many redox proteins use this mechanism (*73*). Many copper and zinc binding proteins use solely histidines as metal acceptor (*74*). Extensive studies on copper binding proteins suggest that copper binds to the HX_mH ($m \leq 4$) motifs involving contacts to $N^{\varepsilon 2}$ (*75*). Importantly, this nitrogen binds copper less tightly, but provides more space to carry out enzymatic activity of the protein as if $N^{\delta 1}$ would be the metal acceptor. The strength of copper binding is exploited by copper chaperones to shuttle copper from one protein to another. Also, in the cell, copper always binds to a protein instead of being in the solution (*71*).

10. The observed termination of redox activity after copper addition observed in NMR-based redox assay (Fig.4.27) can only be explained by the inactivation of proton transfer within the imidazole ring of the histidine side chains, since copper blocks the tautomeric nature of the imidazole ring. Notably, it has been shown that αB-crystallin uses two stretches of amino acid sequences, A^{57}PSWIDTGLSEMR^{69} and V^{93}LGDVIEVHGKHEER^{107}, to bind its substrates (*76*). Interestingly, These sequences also include the putative motif H^{101}GKH^{104}, which is involved in the observed catalytic redox activity (*76*).

11. Taking results from the redox activity assay, ITC experiment and the gel filtration experiments together, we conclude that the αB-crystallin uses copper as a switch to regulate binding to Aβ(1-40). It has been postulated that the Aβ(1-40) producing APP plays a role in maintaining neuronal copper homeostasis (*66*). The role of copper in Alzheimer's pathogenesis is known and recently it has been shown that the elevation of the intracellular copper concentration stabilizes the activity of SOD1 and decreases the production of Aβ(1-40) (*77*). Also, it has been reported that increase in intracellular copper level through over expressing mutated copper chaperone protein,

reduces the production of Aβ(1-40) (*78*). From the above studies, we conclude that neuronal copper homeostasis is essential for the neuronal functions. Importantly, SOD1 activity, not only restricted to copper homeostasis but is also regulated by means of cellular redox condition. Recently published reports suggest that cellular GSH level alter the ROS clearance pathway by activating SOD1 (*79*). All the above facts indicate a clear correlation between the copper transport, redox activity and the neurodegenerative disease. We propose here that αB-crystallin has a cellular redox function as it modulates the activity of other redox active proteins. The redox activity might be modulated by copper homeostasis. Copper homeostasis has an influence on the substrate binding affinity of αB-crystallin.

4.5   Conclusion

αB-crystallin recognizes the same hydrophobic residues in Aβ(1-40) which are also involved in Aβ(1-40)·Aβ(1-40) interactions. The interaction between Aβ(1-40) and αB-crystallin is found to be regulated through copper. In the presence of copper, the interaction between αB-crystallin and Aβ(1-40) is reduced and Aβ(1-40) aggregation occurs. The molecular chaperone αB-crystallin could function as a redox regulatory protein as it is found to regulate other cellular redox proteins. αB-crystallin modulates the redox activity of Aβ(1-40) by directly interacting with the misfolded peptide. The neurotoxicity of Aβ(1-40) induced by αB-crystallin, could be attributed due to the changes in neuronal redox homeostasis. Hence, we postulate here that the early event in the neurodegenerative diseases is a breakdown of redox homeostasis and the disturbance of excitatory and inhibitory synaptic potentials directly or indirectly due to the redox burden. The above findings provide a new concept in the understanding of neuronal degeneration.

4.6   References

1.   Selkoe, D. J. (2001) *Physiol.Rev. 81*, 741-766.
2.   Hardy, J. A., Higgins, G.A. (1992) *Science 256*, 184-185.
3.   Yankner, B. A., Dawes, L.R., Fisher, S., Villa-Komaroff, L., Oster-Granite, M.L., Neve, R.L. (1989) *Science 245*, 417-420.
4.   Esler, W. P., Wolfe, M.S. (2001) *Science 293*, 1449-1454.
5.   Haass, C., Koo, E.H., Capell, A., Teplow, D.B., Selkoe, D.J. (1995) *J.Cell.Biol. 128*, 537-47.
6.   Parvathy, S., Hussain, I., Karran, E.H., Turner, A.J., Hooper, N.M. (1999) *Biochemistry 38*, 9728-9734.
7.   Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., Teplow, D.B., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M.A., Biere, A.L., Curran, E., Burgess, T., Louis, J.C., Collins, F., Treanor, J., Rogers, G., Citron, M. (1999) *Science 286*, 735-741.
8.   Iwata, N., Tsubuki, S., Takaki, Y., Watanabe, K., Sekiguchi, M., Hosoki, E., Kawashima-Morishima, M., Lee, H.J., Hama, E., Sekine-Aizawa, Y., Saido, T.C. (2000) *Nat.Med. 6*, 143-150.
9.   Hardy, J. A., Selkoe, D.J. (2002) *Science 298*, 962-964.
10.   Lansbury, P. T. J. (1997) *Neuron 19*, 1151-1154.
11.   Selkoe, D. J. (2003) *Nature 426*, 900-904.
12.   Haass, C. (2001) *Nat.Neurosci. 4*, 859-860.
13.   Caughey, B., Lansbury, P.T. Jr. (2003) *Ann.Rev.Neurosci. 26*, 267-298.
14.   Kim, H. J., Chae, S.C., Lee, D.K., Chromy, B., Lee, S.C., Park, Y.C., Klein, W.L., Krafft, G.A., Hong, S.T. (2003) *Faseb J. 17*, 118.
15.   Hartley, D. M., Walsh, D.M., Ye, C.P., Diehl, T., Vasquez, S., Vassilev, P.M., Teplow, D.B., Selkoe D,J. (1999) *J.Neurosci. 19*, 8876-8884.
16.   Mattson, M. P. (1997) *Physiol.Rev. 77*, 1081-1132.
17.   Lashuel, H. A., Hartley, D., Petre, B.M., Walz, T., Lansbury, P.T. Jr. (2002) *Nature 418*, 291.
18.   Butterfield, D. A., Yatin, S.M., Varadarajan, S., Koppal, T. (1999) *Methods Enzymol. 309*, 746-768.
19.   Abramov, A. Y., Canevari, L., Duchen, M.R. (2004) *J. Neurosci. 24*, 565-575.
20.   Nakagawa, T., Zhu, H., Morishima, N., Li, E., Xu, J., Yankner, B.A., Yuan, J. (2000) *Nature 403*, 98-103.
21.   Fonte, V., Kapulkin, V., Taft, A., Fluet, A., Friedman, D., Link, C.D. (2002) *Proc.Natl.Acad.Sci.U.S.A. 99*, 9439-9444.
22.   Yoo, B. C., Kim, S.H., Cairns, N., Fountoulakis, M., Lubec, G. (2001) *Biochem Biophys Res Commun. 280*, 249-58.
23.   Shinohara, H., Inaguma, Y., Goto, S., Inagaki, T., Kato, K. (1993) *J. Neurol sci. 119*, 203-208.
24.   Gaestel, M., Vierling, E., Buchner, J. (1997) *Guidebook to molecular chaperones and protein foldign catalysts Gething, M., ed.*, pp.269-272, oxford university press, New york.
25.   van Montfort, R. M., Basha, E., Friedrich, K.L., Slingsby, C., Vierling, E. (2001) *Nat.Struct.Biol. 8*, 1025-1030.
26.   Renkawek, K., van Workum, F.A.P., de Jong, W.W. (1995) *Dev.Brain.Dysfunct. 8*, 35-39.
27.   Siegel, G. J., Agranoff, B.W., Fisher, S.K., Albers, R.W., Uhler, M.D. (Eds.). (1999) *Basic Neurochemistry- Molecular, cellular and medical aspects 6th Ed.*, Lippincott williams & wilkins.

28. Goldstein, L. E., Muffat, J.A., Cherny, R.A., Moir, R.D., Ericsson, M.H., Huang, X., Mavros, C., Coccia, J.A., Faget, K.Y., Fitch, K.A., Masters, C.L., Tanz,i R.E., Chylack, L.T. Jr., Bush, A.I. (2003) *Lancet 361*, 1258-1265.

29. van Rijk, A. F., Bloemendal, H. (2000) *Ophthalmologica 214*, 7-12.

30. Lowe, J., Landon, M., Pike, I., Spendlove, I., Mc Dermott, H., Mayer, R.J. (1990) *Lancet 336*, 515-516.

31. Benndorf, R., Welsh, M.J. (2004) *Nat.Genet. 36*, 547-8.

32. Liang, J. J. (2000) *FEBS Lett. 484*, 98-101.

33. Kudva, Y. C., Hiddinga, H.J., Butler, P.C., Mueske, C.S., Eberhardt, N.L. (1997) *FEBS Lett. 416*, 117-21.

34. Stege, G. J., Renkawek, K., Overkamp, P.S., Verschuure, P., van Rijk, A.F., Reijnen-Aalbers, A., Boelens, W.C., Bosman, G.J., de Jong, W.W. (1999) *Biochem Biophys Res Commun. 262*, 152-156.

35. Gibbs, S. J., Johnson, C.S.Jr. (1991) *J.Magn.Reson. 93*, 395-402.

36. Wilkins, D. K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Johnes, J.A., Smith, L.J.,. (1999) *Biochemistry 38*, 16424-16431.

37. Sreerama, N., Woody, R.W. (2000) *Anal.Biochem. 287*, 252-260.

38. Bianco-Peled, H., Gryc, S. (2004) *Langmuir 20*, 169-174.

39. Goto, Y., Aimoto, S. (1991) *J.Mol.Biol. 218*, 387-396.

40. Barrow, C. J., Zagorski, M.G. (1991) *Science 253*, 179-182.

41. Sticht, H., Bayer, P., Willbold, D., Dames, S., Hilbich, C., Beyreuther, K., Frank, R.W., Rösch, P. (1995) *Eur.J.Biochem. 233*, 293-298.

42. Coles, M., Bicknell, W., Watson, A., Fairlie, D.P., Craik, D.J. (1998) *Biochemistry 37*, 11064-11077.

43. Petkova, A. T., Ishii, Y., Balbach, J.J., Antzutkin, O.N., Leapman, R.D., Delaglio, F., Tycko, R. (2002) *Proc.Natl.Acad.Sci.U.S.A. 99*, 16742-16747.

44. LeBlanc, A., Xue, R., Gambetti, P. (1996) *J.Neurochem. 66*, 2300-2310.

45. van Noort, J. M., van Sechel, A.C., Bajramovic, J.J., el Ouagmiri, M., Polman, C.H., Lassmann, H., Ravid, R. (1995) *Nature 375*, 798-801.

46. Bachoo, R. M., Kim, R.S., Ligon, K.L., Maher, E.A., Brennan, C., Billings, N., Chan, S., Li, C., Rowitch, D.H., Wong, W.H., DePinho, R.A. (2004) *Proc.Natl.Acad.Sci.U.S.A. 101*, 8384-8389.

47. Clement, A. M., Nguyen, M.D., Roberts, E.A., Garcia, M.L., Boillée, S., Rule, M., McMahon, A.P., Doucette, W., Siwek, D., Ferrante, R.J., Brown, R.H. Jr., Julien, J.P., Goldstein, L.S.B., Cleveland, D.W. (2003) *Science 302*, 113-117.

48. Leavitt, S., Freire, E. (2001) *Current opin.Struct.Biol. 11*, 560-566.

49. Multhaup, G., Schlicksupp, A., Hesse, L., Ruppert, T., Masters, C.L., Beyreuther, K. (1996) *Science 271*, 1406-1409.

50. Tabner, B. J., Turnbull, S.T., El-Agnaf, O.M.A., Allsop, D. (2002) *Free radic.Biol.Med 32*, 1076-1083.

51. Narayanan, S., Reif, B. (Communicated) *FEBS Lett.*

52. Hou, L., Shao, H., Zhang, Y., Li, H., Menon, N.K., Neuhaus, E.B., Brewer, J.M., Byeon, I.L., Ray, D.G., Vitek, M.P., Iwashita, T., Makula, R.A., Przybyla, A.B., Zagorski, M.G. (2004) *J.Am.Chem.Soc. 126*, 1992-2005.

53. Bush, A. I., Masters, C.L., Tanzi, R.E. (2003) *Proc.Natl.Acad.Sci.U.S.A. 100*, 11193-11194.

54. Gurney, M. E., Pu, H., Chiu, A.Y., Dal Canto, M.C., Polchow, C.Y., Alexander, D.D., Caliendo, J.,Hentati, A., Kwon, Y.W., Deng, H.X, Chen,W.J., Zhai, P., Sufit, R.L., Siddique, T. (1994) *Science 264*, 1772-1775.

55. Wang, J., Slunt, H., Gonzales, V., Fromholt, D., Coonfield, M., Copeland, N.G., Jenkins, N.A., Borchelt, D.R. (2003) *Hum.Mol.Genet. 12*, 2753-2764.

56. Goldstein, L. E., Leopold, M.C., Huang, X., Atwood, C.S., Saunders, A.J., Hartshorn, M., Lim, J.T., Faget, K.Y., Muffat, J.A., Scarpa, R.C., Chylack, L.T. Jr., Bowden, E.F., Tanzi., R.E., Bush, A.I. (2000) *Biochemistry 39*, 7266-7275.

57. Ganadu, M. L., Aru, M., Mura, G.M., Coi, A., Mlynarz, P., Kozlowski, H. (2004) *J.Inorg.Biochem. 98*, 1103-1109.

58. Culotta, V. C., Klomp, L.W., Strain, J., Casareno, R.L., Krems, B., Gitlin, J.D. (1997) *J.Biol.Chem. 272*, 23469.

59. Atwood, C. S., Scarpa, R.C., Huang, X., Moir, R.D., Jones, W.D., Fairlie, D.P., Tanzi, R.E., Bush, A.I. (2000) *J.Neurochem. 75*, 1219-1233.

60. Mehlen, P., Kretz-Remy, C., Preville, X., Arrigo, A.P. (1996) *Embo.J 15*, 2695-2706.

61. Bitan, G., Tarus, B., Vollers, S.S., Lashuel, H.A., Condron, M.M.,Straub, J.E.,Teplow, D.B. (2003) *J.Am.Chem.Soc. 125*, 15359-15365.

62. Schöneich, C., Pogocki, D., Hug, G.L., Bobrowski, K. (2003) *J.Am.Chem.Soc. 125*, 13700-13713.

63. Levine, R. L., Mosoni, L., Berlett, B.S., Stadtman, E.R. (1996) *Proc.Natl.Acad.Sci.U.S.A. 93*, 15036-15040.

64. Hoshi, T., Heinemann, S.H. (2001) *J.Physiol. 531*, 1-11.

65. Kanski, J., Varadarajan, S., Aksenova, M., Butterfield, D.A. (2001) *Biochim.Biophy.Acta 1586*, 190-198.

66. Barnham, K. J., Ciccotosto, G.D., Tickler, A.K., Ali, F.E., Smith, D.G., Williamson, N.A., Lam, Y.H., Carrington, D., Tew, D., Kocak, G., Volitakis, I., Separovic, F., Barrow, C.J., Wade, J.D., Masters, C.L., Cherny, R.A., Curtain, C.C., Bush, A.I., Cappai, R. (2003) *J.Biol.Chem. 278*, 42959-42965.

67. Selkoe, D. J. (2002) *Science 298*, 789-791.

68. Hodgkin, A. L., Huxley, A.F. (1939) *Nature 144*, 710.

69. Chivers, P. T., Prehoda, K.E., Raines, R.T. (1997) *Biochemistry 36*, 4061-4066.

70. Fomenko, D. E., Gladyshev, V.N. (2002) *Protein Sci. 11*.

71. O'Halloran, T. V., Culotta, V.C. (2000) *J.Biol.Chem. 275*, 25057-25060.

72. Wilkstrom, M., Bogachev, A., Finel, M., Morgan, J.E., Puustinen, A., Raitio, M., Verkhovskaya, M., Verkhovsky, M.I. (1994) *Biochim Biophys Acta 1182*, 106-111.

73. Wilkstöm, M. (2000) *Biochim.Biophy.Acta 1458*, 188-198.

74. Holm, R. H., Kennepohl, P., Solomon, E.I. (1996) *Chem.Rev. 96*, 2239-2314.

75. Karlin, S., Zhu, Z., Karlin, K. (1997) *Proc.Natl.Acad.Sci.U.S.A. 94*, 14225-14236.

76. Sharma, K. K., Kaur, H., Kester, K. (1997) *Biochem Biophys Res Commun. 239*, 217-222.

77. Bayer, T. A., Schäfer, S., Simons, A., Kamer, T., Tepest, R., Eckert, A., Schüssel, K., Eikenberg, O., Sturchler-Pierrat, C., Abramowski, D., Staufenbiel, M., Multhaup,G. (2003) *Proc.Natl.Acad.Sci.U.S.A. 100*, 14187-14192.

78. Phinney, A., Drisaldi, B., Schmidt, S.D., Lugowski, S., Coronado, V., Liang,Y., Horne, P., Yang, J., Sekoulidis, J., Commaraswamy, J., Chishti, M.A.,Cox, D.W., Mathews, P.M., Nixon, R.A., Carlson, G.A., St George-Hyslop, P., Westaway, D. (2003) *Proc.Natl.Acad.Sci.U.S.A. 100*, 14193-14198.

79. Carroll, M. C., Girouard, J.B., Ulloa, J.L., Subramaniam, J.R.,Wong, P.C., Valentine, J.S.,Culotta, V.C. (2004) *Proc.Natl.Acad.Sci.U.S.A. 101*, 5964-5969.

5.0   Solid-state and solution-state NMR investigation of SH3 fibril formation

5.1   Background

Fibril formation has often been viewed as a general property of polypeptide chains. Several studies have shown that under a variety of non-physiological conditions, synthetic peptide sequences and various proteins can form ordered fibrils (*1-4*). It was postulated that under appropriate conditions, any globular protein can form amyloid fibrils *in vitro* (*1*). It is intriguing to find out whether these apply to certain protein structures and folds? One way to test this is to analyze evolutionary related proteins with similar structures and low sequence homology, where the fibrils are obtained under certain conditions for a protein, whereas its homologues do not fibrillize at all. Especially interesting is to pick up a protein family whose members are characterized thermodynamically and kinetically.
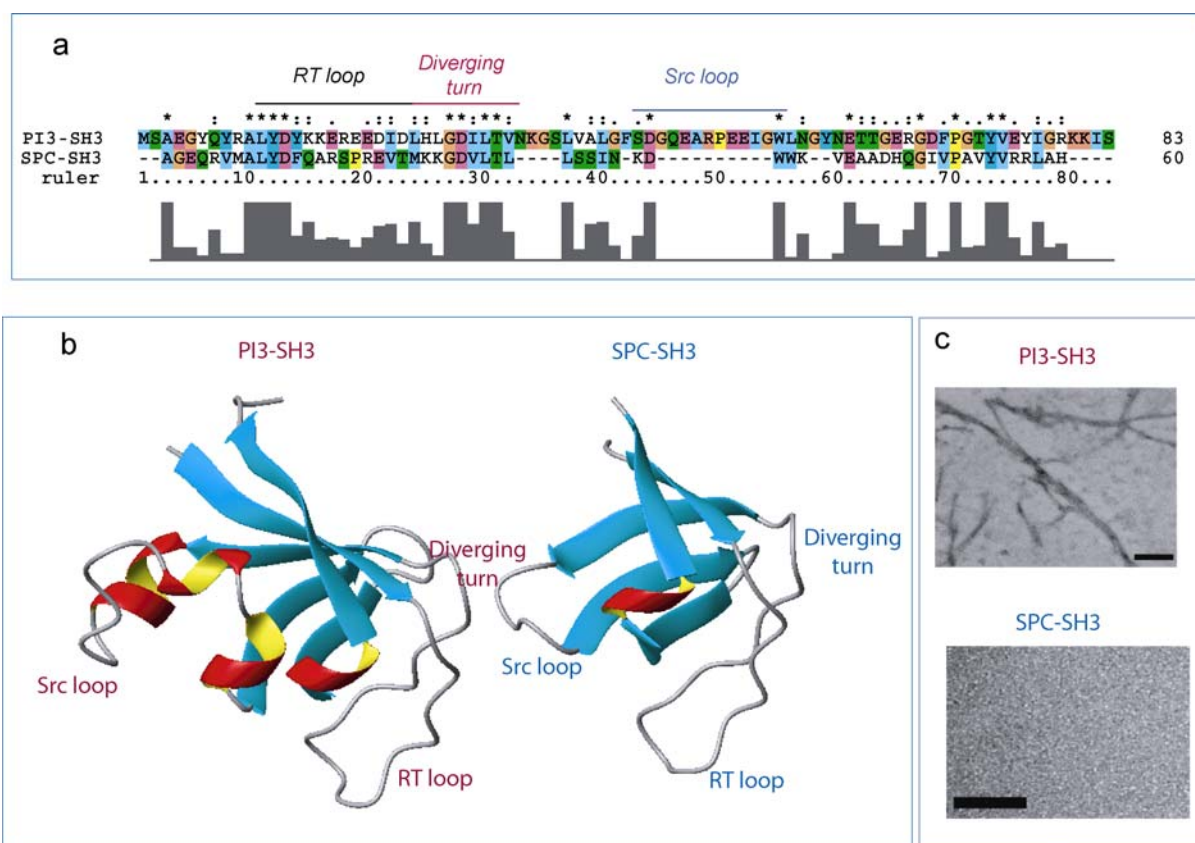
Figure 5.1. (a) Sequence comparison of PI3-SH3 and SPC-SH3;(b) Structural comparison of PI3-SH3 and SPC-SH3 proteins. c) Fibril forming ability of PI3-SH3 (top) and inability of SPC-SH3 (bottom).

One of such group of family of proteins is the family of Src homology 3 (SH3) domain proteins that had been characterized regarding its folding behaviour and thermodynamic characteristics .

SH3 domains are conserved protein modules comprising of 50-70 amino acids which are found in a variety of proteins with important roles in signal transduction. These domains have been shown to mediate protein-protein interactions by binding to short proline-rich sequences (*5*). The SH3 domain of the p85α subunit of phosphatidylinositol-3 kinase (PI3-SH3) is known to form amyloid fibrils *in vitro* under acidic conditions (Fig. 5.1c, top). By contrast, the related α-spectrin-SH3 (SPC-SH3), which has the same fold and some sequence identity with PI3-SH3 (Fig. 5.1a,b), does not form amyloid fibrils under conditions that have been tested so far (*6*) (Fig. 5.1c, bottom). Hence, it is interesting to know what governs the fibril formation of PI3-SH3.

5.2 Aim of the project

Understanding the molecular mechanism of fibril formation and identification of the amino acids that are involved in this process.

5.3 Material and Methods

All ^{15}N, ^{13}C-uniformly or differentially labelled proteins were obtained from our collaborator Dr. Salvador Ventura, Universitat Autonoma de Barcelona, Spain. Protein was expressed and purified as described in the literature (*3*).

5.3.1 Fibril preparation protocol

Signal intensities in a 1D-^{15}N MAS solid-state NMR spectroscopy are indicative for the quality of fibril formation. Disordered fibrils adopt in several possible conformers so that the resonance lines become broad and may eventually disappear. In order to obtain well ordered fibrils, existing methods of fibril formation employ pepsin digestion or protein incubation at pH 2.0 for longer periods in time (*7*). We screened several methods for fibril formation and found the following method to be superior to any other method in the literature. Using this method, we formed ordered fibrils as judged from the linewidth of the histidine side chain signals (Fig. 5.2). We could increase the signal intensity at least by a factor of two. Best NMR results were obtained for 10 mg of protein dissolved in 1 ml of 20 mM of phosphate buffer (pH ~8.5). The resulting solution was transferred to a 3.5-kDa dialysis bag (*float-a-Lyzer,* Spectrapor, Germany). The protein solution was dialyzed against 2 liter of 5 mM phosphate buffer, pH 5.0 which was kept at 50°C for 4 hours. The dialysis bag was then transferred to another flask containing 2 liter of 5 mM glycine buffer at pH 2.0, which was kept at room temperature. A transition from the semi-solid white substance to colourless glassy substance

was obtained. The dialysis bag was kept inside the flask for 15 days with constant stirring. After 15 days, the protein was centrifuged at 45000 g. The sedimented protein was filled into a 4-mm rotor to carry out the NMR experiments.
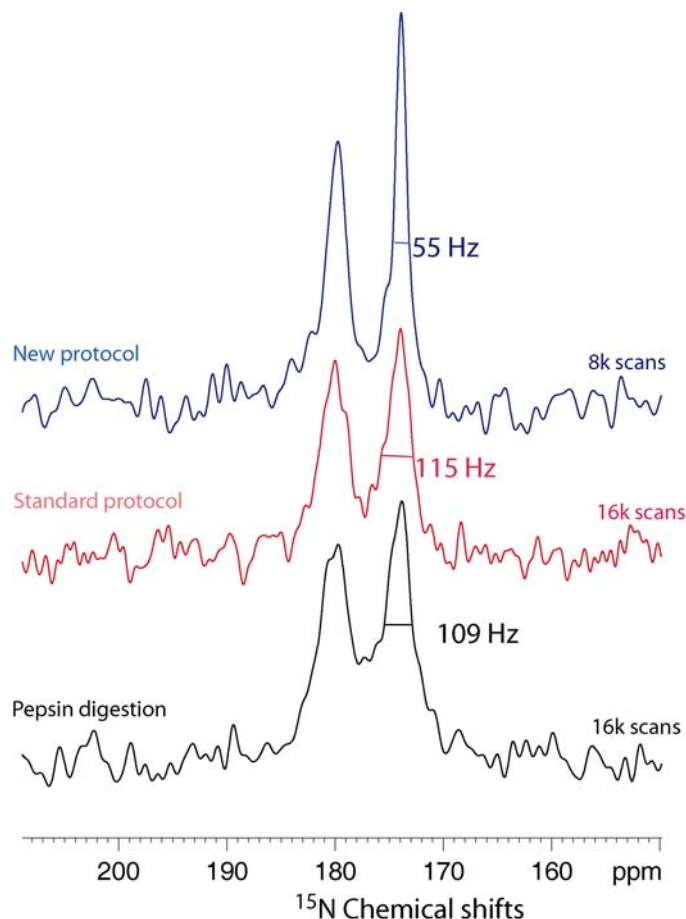
Figure 5.2. ^{15}N-Histidine side chain resonances in PI3-SH3. A novel protocol for fibril formation increased the signal intensity atleast by a factor of two (blue) compared to traditional methods like pepsin digestion (black) and 2 weeks room temperature incubation (red). All protocol uses glycine as final buffer at pH 2.3.

5.3.2 Solid-state NMR spectroscopy

NMR experiments were performed on a 499.974-MHz wide-bore Bruker spectrometer using a commercial triple-resonance probe.

5.3.2.1   1D-^{15}N MAS-SS NMR spectroscopy

Typically 16k scans were accumulated for each 1D-^{15}N spectrum. Pseudo-2D experiments were also performed by incrementing the cross-polarization (CP) time by corresponding units (initial value of 20 µs with 20 µs increments for crystalline ^{15}N-histidine hydrochloride and

initial value of 100 μs with 200 μs increments for the PI3-SH3) in order to probe the proton environment of the histidine side chains in PI3-SH3.

5.3.2.2 Proton-driven ^{15}N-^{15}N correlation experiment

Proton driven ^{15}N-^{15}N correlation experiment was performed in order to get sequential ^{15}N-^{15}N correlations of the protein backbone (*8*). Cross peak build-up using the proton spin diffusion mixing scheme is usually very fast which yields an equilibration of magnetization after ~ 500 ms for strongly coupled protons (*9*). Correlation spectrum was recorded using 1024 points in the direct dimension and 256 points in the indirect dimension with the mixing time of 3 sec. The experiment was carried out at 270 K. A CP contact time of 2 ms was used to transfer the magnetization from the proton to nitrogen under Hartman-Hahn matching condition ($\gamma_H B_{1(H)} = \gamma_N B_{1(N)} \pm n\omega_r$) at a MAS rotation frequency of 10 kHz (Fig. 5.3b).
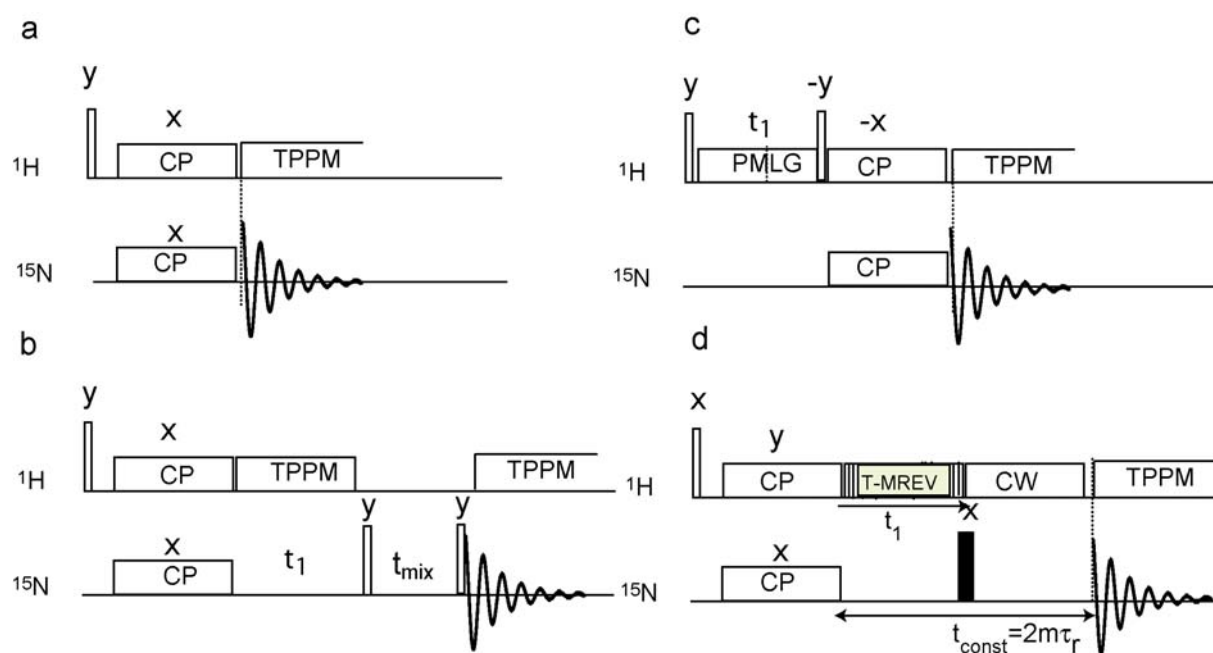
Figure 5.3. a)^{15}N-1D-CP-MAS NMR; b) ^{15}N-^{15}N correlation using proton driven spin diffusion; c) HETCOR-using phase modulated Lee-Goldburg to obtain ^1H,^1H dipolar decoupling; d) T-MREV.

5.3.2.3 Heteronuclear correlation experiments (HETCOR)

Heteronuclear ^1H-^{15}N correlation experiment were performed using decoupling sequences. Frequency switched Lee-Goldburg (FSLG) and its modified version, phase-modulated Lee-Goldburg pulse sequences (PMLG), improved line narrowing by elimination of proton-proton couplings in the indirect dimension (*10, 11*). 150 μs was used as CP contact time to transfer the magnetization from proton to its directly bonded nitrogen and vice-versa at a

MAS rotation frequency of 14 kHz. A Sweep width of 10 kHz and 12.5 kHz was used for the direct (^{1}H) and indirect dimension (^{15}N), respectively. For the FSLG experiments, 256 increments were accumulated in the indirect dimension (^{1}H) for both the ^{15}N-histidine and PI3-SH3 samples. CP time of 150 μs and 150 μs or 2500 μs were used for the ^{15}N-histidine hydrochloride and PI3-SH3 samples, respectively. The FSLG decoupling uses a ^{1}H nutation frequency of 100 kHz and a frequency offset of ±43 kHz in ^{1}H dimension. The scaling factor of $\sqrt{3}$ of the LG sequence was also taken into consideration while calibrating the spectra (Fig. 5.3c).

5.3.2.4 Transverse MREV experiment (T-MREV)

T-MREV experiment offers exact measurement of N-H bond lengths and H-N-H bond angles without relying on the MAS frequency (*12*). Under actual experimental conditions, the spinning frequency was adjusted to 8333 Hz to yield a ^{1}H 90° pulse length of 2 μs in the T-MREV multiple-pulse sequence. The T-MREV ^{15}N-^{1}H dephasing period in the t_1 evolution period was incremented by 24 μs (for the recoupling time over 5 rotor periods) and 48 μs (for the recoupling time over 10 rotor periods) and a total of 25 experiments were recorded in the indirect dimension. The recycle delay was set to 3 sec. TPPM decoupling was used in the direct evolution periods with a ^{1}H RF field of about 60 kHz (Fig 5.3d). The experimental data were fitted to a Bessel-type function using a Fortran code (Courtesy of Dr. M. Hohwy, MIT, Boston).

5.3.3   Biochemical experiments

5.3.3.1 Histidine modification using diethyl pyrocarbonate

Reversible histidine modification in PI3-SH3 was carried out using diethyl pyrocarbonate (DEPC) as a modifier (*13*) (Fig. 5.4). This modification does not alter the net charge in the diverging turn, but blocks the tautomeric nature of protons in the histidine side chains (Fig. 5.4). Considering the favourable protonation of $N^{\epsilon 2}$ as compared to $N^{\delta 1}$, the $H^{\delta 1}$ is more likely to replaced by the carbethoxy group of DEPC (*14, 15*). The modification is reversible as hydroxylamine can remove the carbethoxylation of the histidine side chain (Fig. 5.4). The modification was performed by incubating DEPC at a molar ratio of 1:100 of [protein]: [DEPC], pH 7.5 at 4 °C for 2 hours. A concentration of 1ml of 500 M of PI3-SH3 was used for this purpose. After two hours, the un-ligated DEPC was washed away either by overnight dialysis or by quick buffer exchange using gel filtration columns (PD-10, Amersham

Biosciences, Germany). The level of histidine modification was monitored using a UV-spectrophotometer by measuring the ratio of the absorbance of the histidyl-carbethoxyl group at 240 nm relative to 270 nm with an absorption coefficient of 3200 M^{-1} cm^{-1} (16). An equimolar ratio of hydroxylamine was used to restore the unmodified PI3-SH3 from the modified one. MALDI-MS analyses were carried out using limited-proteolysis approach with trypsin and chymotrypsin as proteases in order to know the exact location of modification.
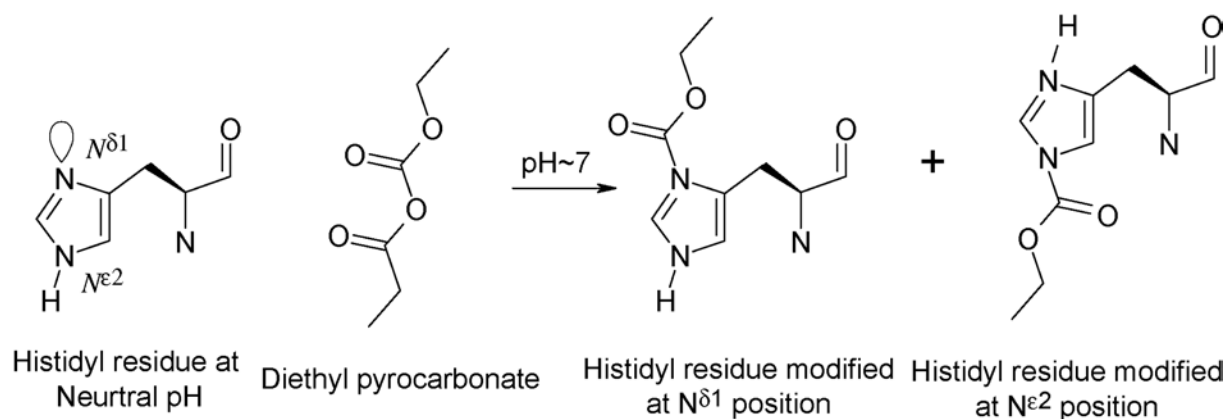
Figure 5.4. DEPC modification of histidine residues.

5.3.3.2 Stability of the modified PI3-SH3 and fibril morphology

The stability of unmodified- and modified-PI3-SH3 upon acid denaturation was determined by measuring the exposure of hydrophobic amino acids. All the experiments were performed at 25° C. Bis-ANS (Sigma), a hydrophobic amino acid-specific dye, was used to probe the exposure of hydrophobic areas when the proteins undergo acid denaturation (17). Pre-sterile filtered 20 mM KH_2PO_4 / Na_2HPO_4 and 20 mM sodium citrate solution was used as a buffer in order to allow for a linear pH relationship during the titration. The initial pH was set to 7.44. Each μl of 4N NaOH / 4M HCl yields an increase or decrease of pH in a regular fashion in 1 ml of buffer (Fig. 5.5). While keeping the final protein to Bis-ANS molar ratio of at 1:10, fluorescence was measured using an excitation wavelength of 395 nm (slit width 5 nm). The emission spectrum was measured by scanning the fluorescence from 445-495 nm (slit width 5 nm). The mixture of Bis-ANS and unmodified or modified PI3-SH3 was incubated for 2 minutes prior to recording. A protein concentration of 1 μM was used for each step in the titration. A concentration of 5 μM of a short peptide (LPFFD) was taken as a reference.

In order to assess the ability of modified PI3-SH3 to form fibrils, 0.2 mM of both unmodified and modified PI3-SH3 were incubated separately in 10 mM glycine buffer, pH 2.3 for 15 days.

The fibril morphology was analyzed for both samples using electron microscopy. A calibrated Superdex-75 column was used to follow the change in molecular weight of the PI3-SH3 upon modification with DEPC. A volume of 100 μl of a 500 μM solution of PI3-SH3 was used in this study.
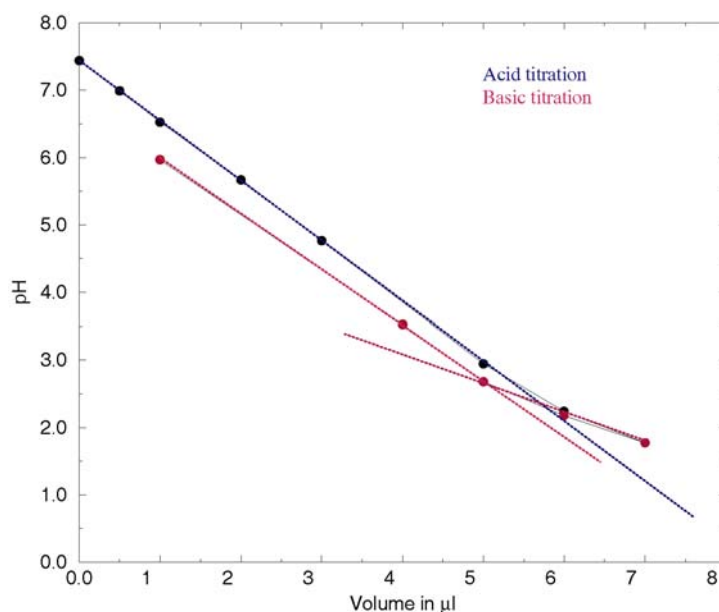
Figure 5.5. Acid (blue) or basic (red) titration of the phosphate/citrate buffer show linear relationship. Hence can be used for the pH titration studies.

5.3.4   Solution- state NMR Spectroscopy

5.3.4.1   HSQC, DOSY, 3D-HNCA, 3D-HNHA NMR experiments

As the fibril formation occurs under acidic conditions, early events in fibrilization can be monitored by acid titration of PI3-SH3. The amino acids that show large chemical shifts perturbations upon acid titration were identified using standard ^{1}H-^{15}N correlation experiments (HSQC). Experiments were performed on both modified and un-modified PI3-SH3. Oligomerization of both modified and un-modified PI3-SH3 were simultaneously monitored using ^{1}H-DOSY experiments taking citrate as an internal standard. Heteronuclear DOSY experiments which exploit the longer spin-lattice relaxation of ^{15}N were also performed (*18*). Magnetization was stored in the ^{15}N z-axis for one second prior to detection. A delay of 3 ms was used as gradient length and 120 ms was used as delay during which molecules were allowed to diffuse. The experimental setup was similar to the one employed for the amyloid peptide (ref. Section 4.3). Though ^{15}N and ^{1}H chemical shifts of PI3-SH3

were obtained from the literature (*19, 20*), 2D-HCCH TOCSY, 3D-HNCA and 3D-HNHA spectra were recorded in order to accomplish H^{α}, C^{α} and side chain assignments (*21-23*).

5.3.4.2 3D-(^{13}C)^{1}H-TOCSY-^{1}H,^{15}N-HSQC and 3D-(^{13}C)-TOCSY –^{1}H,^{15}N- HSQC and intermolecular NOE experiments

Uniformly ^{15}N,^{13}C-labelled protein was used for the sequential assignments of amino acid residues (Fig. 5.6). Selective ^{15}N labelled and selective ^{13}C labelled PI3-SH3 samples were mixed at the ratio of 1:1 to yield differentially labelled samples. If intermolecular interactions take place, then the 25% of the PI3-SH3 exists in ^{15}N:^{13}C form. In order to identify the amino acids that are involved in the intermolecular contacts, differentially labelled samples were used. Delays of 120 ms and 14 ms were used as NOESY and CC-TOCSY mixing time. The MOCCA-XY(16) isotropic-mixing sequence was used for proton-proton mixing (mixing time of 40 ms; delay/pulse length= 2.2) (*24*). Spectra were analyzed in Sparky (*25*). Identification of intermolecular contacts was achieved by the isotopic filtering elements (*26*). A 2D-version of a 3D ^{13}C/^{15}N-editing technique was performed initially. A delay of 200 ms was used as NOE mixing time. 3D-simultaneous ^{13}C-edited and ^{15}N-edited experiments were performed to identify the intermolecular contacts .
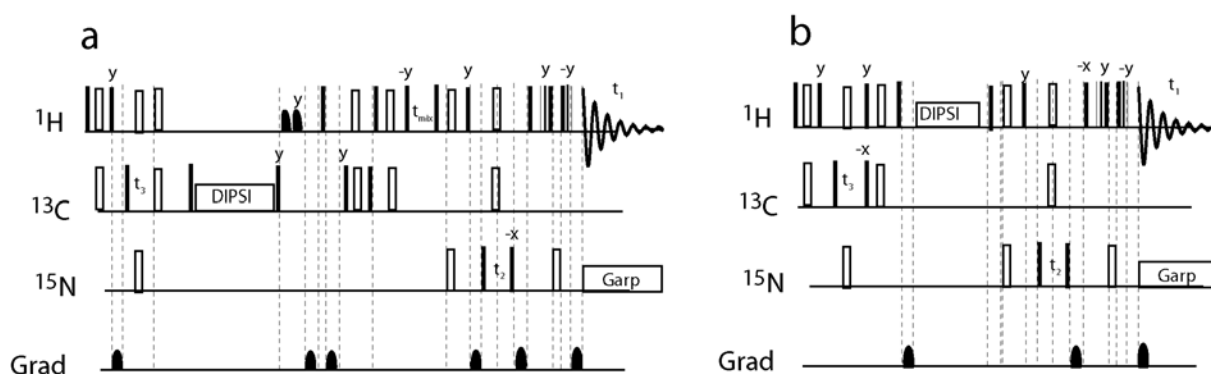
Figure 5.6. In the depicted pulse programs, filled and open rectangular blocks indicate 90 and 180 ° pulses. Gaussian shaped pulses at the proton axis represent soft water flip back pulses (2ms). a) 3D-^{13}Cedited-^{13}C-TOC-NOE-^{15}N-HSQC pulse program. Setting the t-mix as 0 or 120 ms select intra and both inter and intra molecular NOE respectively. TOCSY mixing was removed to increase the intensity of Intermolecular NOE. Differentially labelled samples were used to identify intermolecular NOE. b) 3D-13C-e-1H-TOC-15N-HSQC used to identify carbon resonances of corresponding amino acids.

5.4 Results and Discussion

5.4.1 Solid-state NMR results

Solid-state NMR techniques are capable of providing high-resolution, site-specific structural constraints for amyloid fibrils (27). Uniformly ^{15}N-labelled PI3-SH3 fibrils were analyzed by solid-state NMR. The efficiency of cross polarization of the magnetization transfer from ^{1}H to ^{15}N reflects the local structural and dynamical information. ^{15}N-magnetization build-up was followed in solid-state NMR experiments by incrementing the cross polarization mixing time. Fastest magnetization build-up is expected for the fully protonated non-exchangeable N-H bonds (28). At pH 2.0, the histidine side chain of imidazole ring should be protonated. This was supported by the variable contact time experiment applied to crystalline ^{15}N-histidine hydrochloride. The maximum magnetization transfer was reached using a CP contact time of 120 µs (Fig. 5.7). For PI3-SH3, the backbone magnetization build-up possesses a similar profile as observed for ^{15}N-histidine hydrochloride (Fig. 5.7, blue). However, for the histidine side chain resonances of PI3-SH3, the maximum efficiency of magnetization transfer is only reached at a CP time of 1.5 ms, suggesting the involvement of internal motion in the fibrillar states.
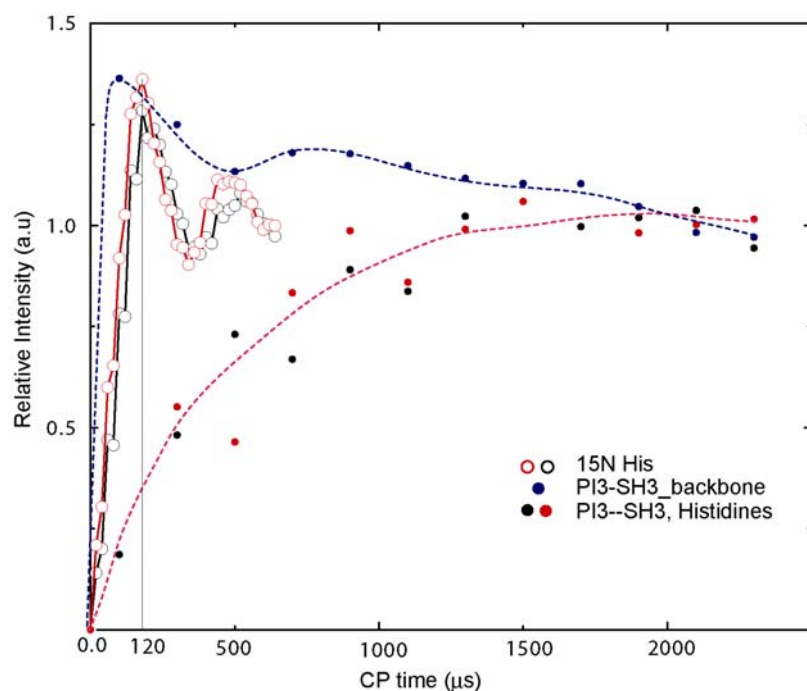
Figure 5.7. Magnetization transfer from ^{1}H to ^{15}N of PI3-SH3 (closed circles) and of crystalline 15N-histidine hydrochloride (open circles). 1D-^{15}N NMR Intensity versus CP contact time.
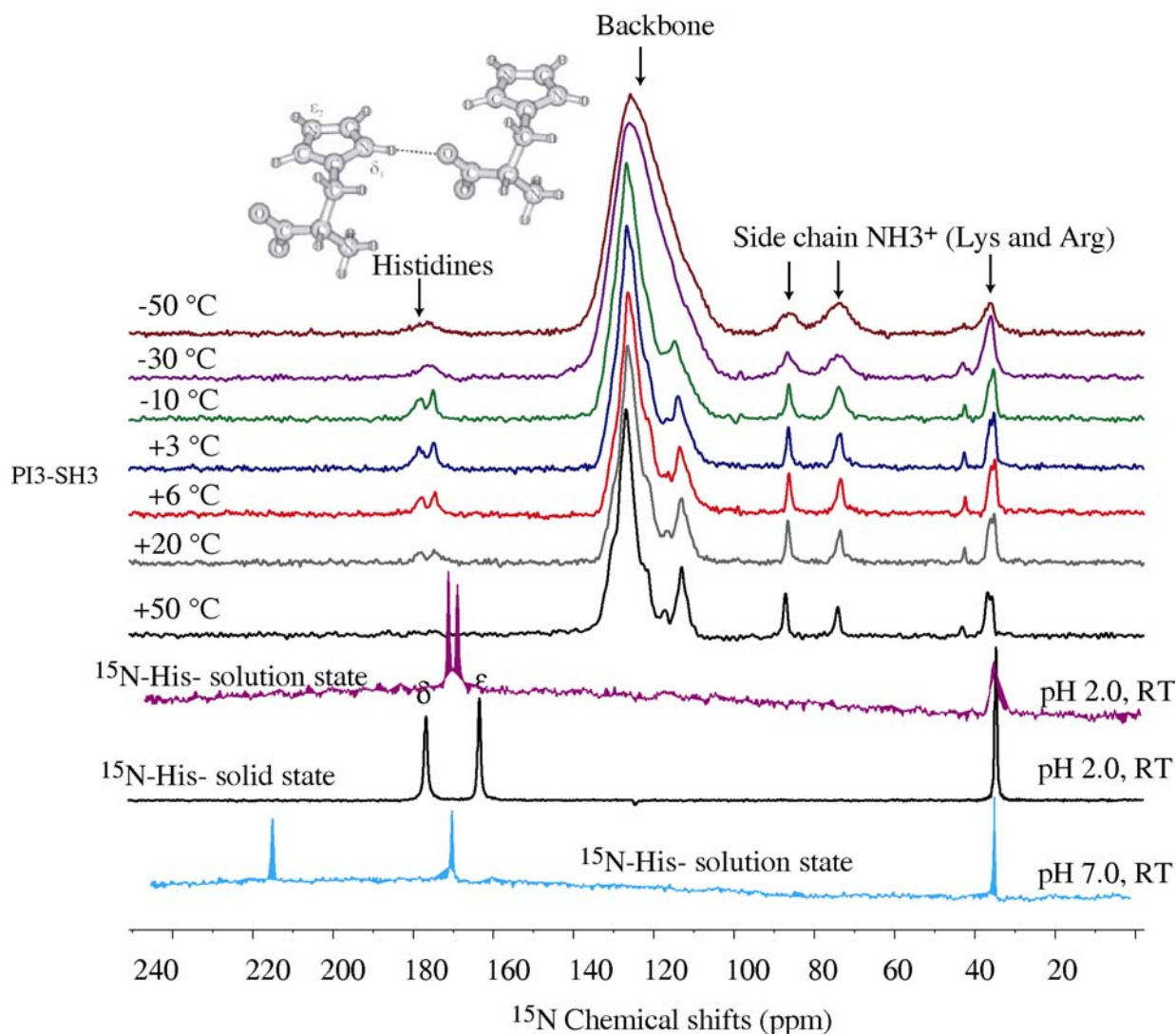
Figure 5.8. Temperature dependent 1D-15N experiment suggested that the side chain histidines were in dynamical motions within fibrils.

The temperature dependence of the 1D-^{15}N spectra also indicates a certain degree of internal motion in the fibrillar state (Fig. 5.8). Reversible line broadening of the resolved histidine side chain resonance was observed at temperatures below freezing or above 50ºC. This suggests that the fibril distortion or the population of alternative conformers. This is also supported by the fact that there were no cross peaks in the proton driven ^{15}N-^{15}N correlation as it provides the information about the structured region (Fig. 5.9). On the contrary, SPC-SH3 has been shown to have several cross peaks (*9*). At low temperature, local dynamics were attenuated and proteins were trapped in different conformational states. Therefore, the side chain resonances for His (185.0 and 181.5 ppm), Arg (ca. 92.1, 79.8 and 48.9 ppm) and

129

Lys (43.0 ppm) were broadened at lower temperatures. For the same reason, signals showed increased resolution at higher temperature, with the exception of His25, the only histidine residue in PI3-SH3. One plausible explanation for this observation is that the environment surrounding His25 is sensitive to temperature heating, as suggested by the line broadening of histidine side chain resonances.
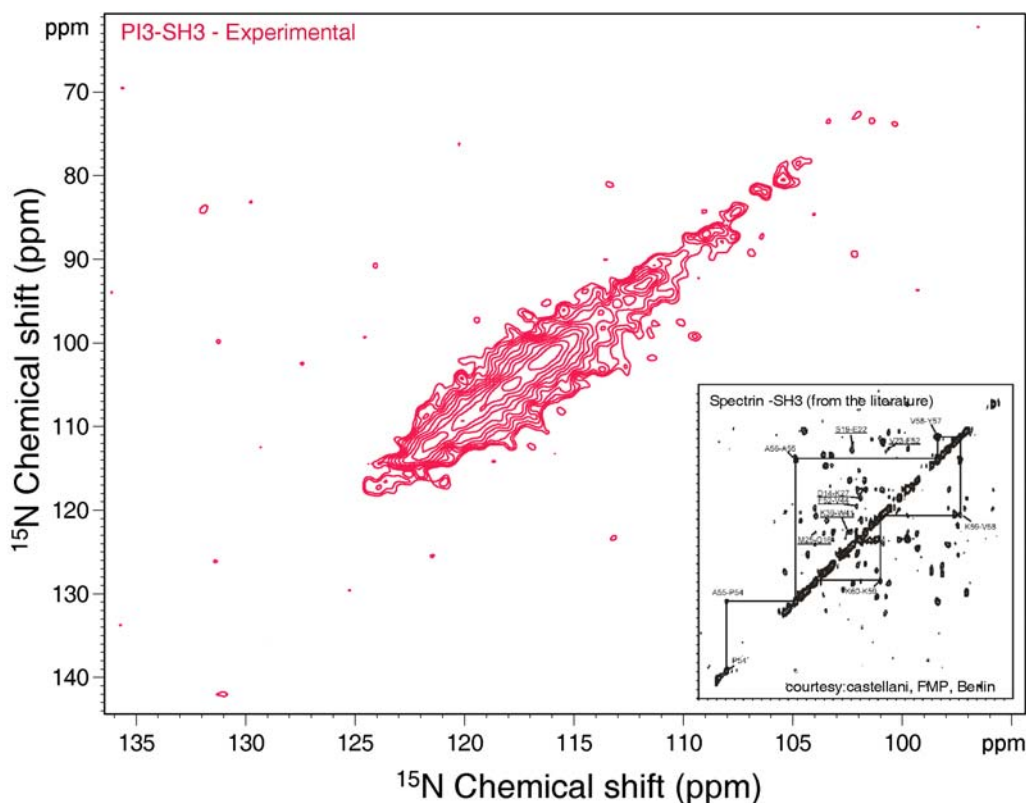
Figure 5.9. No cross peaks could be observed for the proton driven ^{15}N-^{15}N correlation experiment for PI3-SH3. On the contrary spectrin SH3 showed several cross peaks indicated the presence of well-ordered structure.

Comparison of the ^{15}N chemical shifts in the solid state for microcrystalline ^{15}N-histidine-hydrochloride with the shifts of the solubilized amino acid, shows a significant downfield shift for Nδ in the crystalline state (Fig.5.8). In the solid state, Hδ is involved in an intermolecular hydrogen bond, whereas Hε is not (*29*). This is consistent with systematic investigations carried out on model compounds which have been shown that, hydrogen bonding induces a downfield shift of 8 to 10 ppm in the ^{15}N isotropic chemical shift for protonated nitrogens in a cationic imidazole ring (*30*). In its native state structure, His25 of PI3-SH3 is located on the surface of the protein. In the amyloid fibril, however, both His25 Nε and Nδ chemical shifts were significantly downfield shifted, suggesting either a possible

involvement of both nitrogen atoms in hydrogen bonds or perhaps their burial in the interior of the amyloid fibril structure. These non-native features in the fibrillar state of the protein can be indicative of intermolecular interactions that may be present in the cross-β structure of PI3-SH3 amyloid fibrils.

In order to confirm that both histidines are involved in hydrogen bonding, T-MREV experiments were carried out. We expected that involvement of the histidine imidazole in hydrogen bonds should result in a change of bond order which eventually changes the bond length (*31*). Hence, measuring the bond lengths of histidine side chain imidazole H-N should provide information about hydrogen bonding. If the N-H bond length increases, the histidine side chain is involved in a hydrogen bond. The T-MREV experiment recouples the N-H dipolar coupling which is directly related to the N-H distance. The magnetization on ^{15}N (during application of TMREV on ^{1}H) can be described by the Bessel function,

$$\omega(t) = \frac{3\sqrt{3}}{2} b_{NH} \int d\beta \sin\beta \cos[\sin 2\beta t],$$

where b_{NH} is the dipolar coupling constant which can be expressed as

$$b_{NH} = \frac{-\gamma_H \gamma_N \hbar}{(r_{NH})^3} (\mu_0/4\pi),$$

The term r_{NH} represents the inter-nuclear distance.
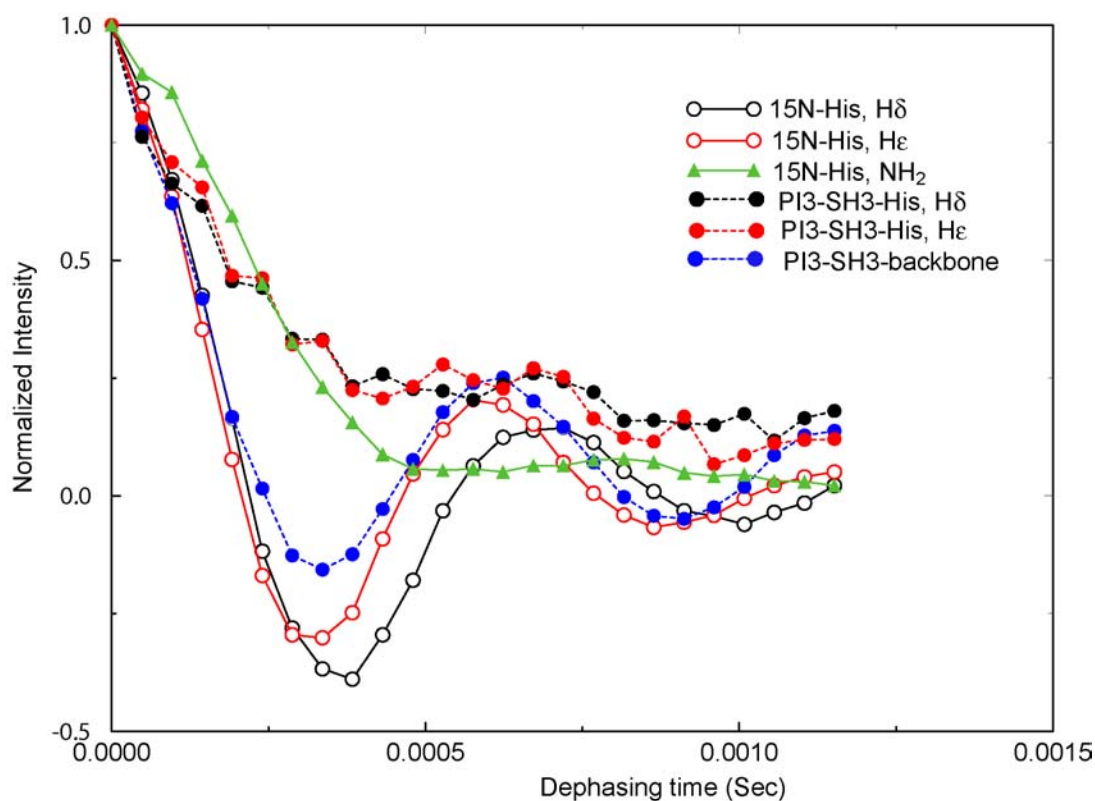
Figure 5.10. T-MREV curves for crystalline ^{15}N-Histidine-monochloridemonohydrate and PI3-SH3-His.

Using this T-MREV element (Fig. 5.3d), two distinct side chain NH bond lengths were obtained for crystalline ^{15}N-L-histidine-hydrochloride. In the case of Hδ, Hε, we obtained a value of 10320Hz and 9000 Hz for the fitted dipolar coupling, respectively (Fig. 5.10). We can therefore conclude that Hδ is involved in hydrogen bond.

In case of PI3-SH3 backbone resonances, the oscillation of magnetization is comparable and the dipolar coupling values could be fit to a value of 11300 Hz. On the contrary, the side chains Hδ and Hε did not show any oscillation of magnetization (Fig. 5.10). The only possible explanation is that side chain dynamics strongly damps the theoretical predicted oscillation.

A heteronuclear [1]H-^{15}N correlation experiment was carried out in order to confirm that both Hδ and Hε protons are involved in hydrogen bonding. The isotropic proton chemical shifts of the histidine Hδ and Hε side chain protons should be indicative of hydrogen bonding. The hydrogen bonded histidine side chain Hδ resonates at 16.4 ppm, where as resonance line of the non-hydrogen bonded Hε proton is located at 12.3 ppm in crystalline ^{15}N-histidine hydrochloride (*32*). Hence, using the side chain proton chemical shifts of histidines, the involvement of histidines in hydrogen bonds can be analyzed. The HETCOR experiment with FSLG / PMLG-type decoupling was used to identify the isotropic chemical shifts of histidine side chain protons. Interestingly, we found that both histidine side chain nitrogens of PI3-SH3 were show correlations with the protons resonating at water frequency (Fig. 5.11). However, careful analysis of the spectra shows a broad peak around 16.5 ppm for both of the protons in the histidine side chains of PI3-SH3 (Fig. 5.11b). A complete description of the various equilibria of intramolecular and intermolecular interactions of histidine in aqueous environment has been presented earlier (*33*). In the study, it was reported that the existence of tautomeric equilibria have an impact on the the population of intermolecular contacts and hence observing hydrogen bonds in such scenario depends upon the population of the species that involve in hydrogen bonds. In the case of PI3-SH3, we consider that the probability of histidine side chain groups (H-N or N, depending upon the tautomeric nature) involved in hydrogen bonding or exchange with water is higher considering the fact that water also do exists in protonated (H_3O^+), neutral (H_2O) in acidic environment. Combining the information from T-MREV experiments and the HETCOR experiment, we can conclude that the histidine side chain protons are in exchange with water or hydrogen bonded to water.

The above explanation was supported by the fact that upon removal of water from the PI3-SH3 using speed-vac, the fibril structure was collapsed as observed by the line width of the side chain and backbone resonances of PI3-SH3 (Fig. 5.12)
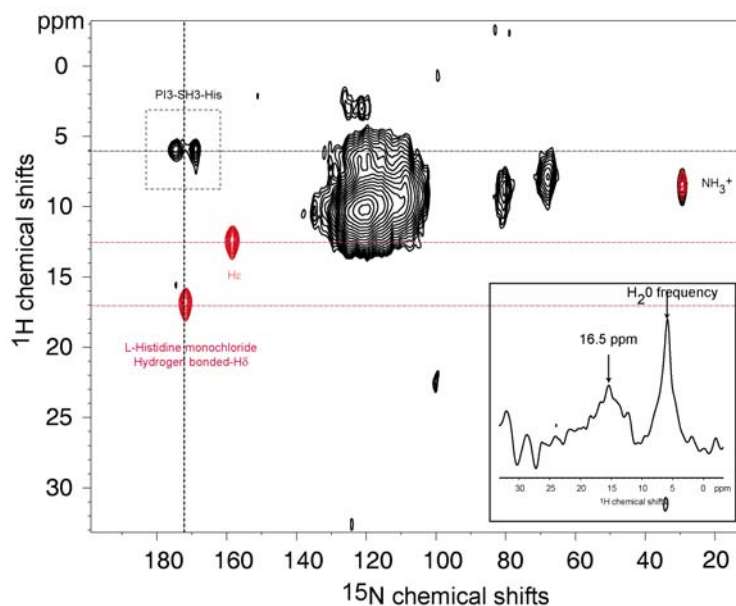
Figure 5.11. ^{1}H-^{15}N heteronuclear correlation experiment of PI3-SH3 (black 2D spectrum). 1D column of the 2D HETCOR spectrum of PI3-SH3 at a ^{15}N chemical shift of 178 ppm indicates a broad resonance at 16.5 ppm in addition to the water resonance frequency (broad peak, inset). The hydrogen bonded proton in L-histidine as approximately the same chemical shift, whereas the non-hydrogen bonded histidine side chain proton adopts a chemical shift of 12.5 ppm (red 2D spectrum).

The above findings are corroborated the postulations that the amyloid fibrils are nothing but, water-filled nano-tubes (*34, 35*) and the molecular packing of the fibrils occurred through water-mediated intermolecular contacts (*36*).
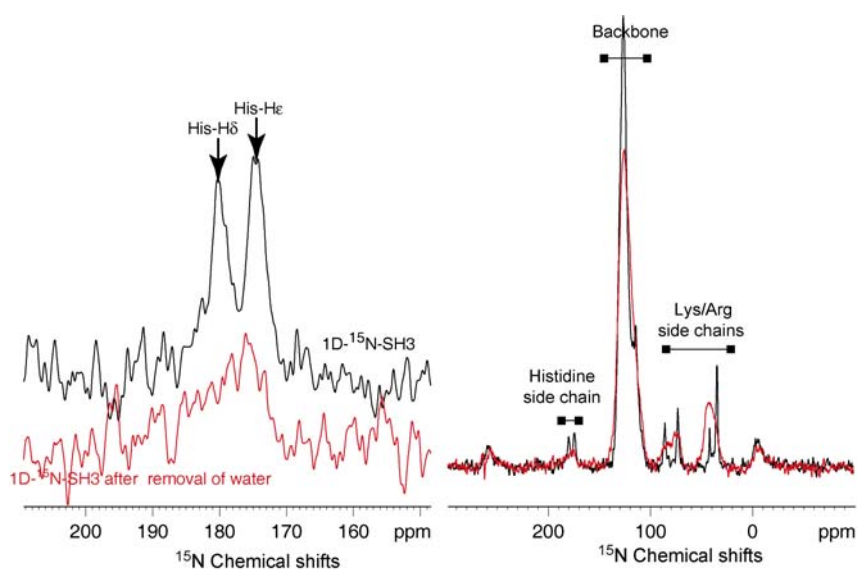
Figure 5.12. 1D-15N-MAS-NMR spectrum of PI3-SH3 before (black) and after water removal (red). In the absence of water the Histidine resonances disappeared suggesting the involvement of water in fibrillar structures.

5.4.2   Sequence conservation in the diverging-turn region of SH3 domains
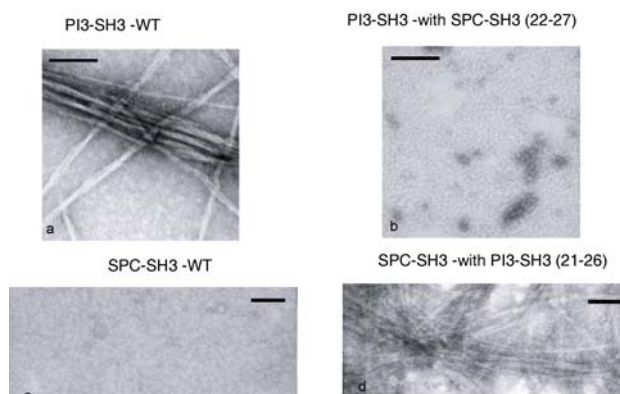
Figure 5.13. Mutation studies have shown the importance of charged amino acids in the diverging turns of PI3-SH3. a) Amyloid fibrils of PI3-SH3 as observed by electron microscopy; b) Replacement of His-Lys to Lys-Lys (as in SPC-SH3) abolished the fibrils in Pi3-SH3. c) Non-amyloidogenic SPC-PI3; d) Formation of amyloid fibrils by SPC-SH3 after incorporation, His-Lys into the diverging turn. The scale bar represents 100 nm (Ventura., S., et al., Proc.Natl.Acad.Sci., (2004), 101, 7258)

The diverging turn, (residues 21-28 of PI3-SH3) in which H25 located, represents a unique subset of type II β–turns stabilized primarily by local interactions. The β-turn is a part of the folding nucleus. When looking at the alignment of the SH3 domain (290 different sequences, for clarity only two are shown) the sequence around His25 in PI3-SH3 appears to be peculiar in the SH3 family (Fig. 5.1a). First, position 26 is occupied in more than 58% of the SH3 sequences by a charged residue, 40% of them present a basic residue at this place, mainly Lys (30%). Leu26 of PI3-SH3 is a rare residue at this position, since it is exposed to the solvent in the native structure of the domain. A hydrophobic residue is found at this position in less than 10% of all SH3 domains, and if so, mainly a smaller residue is incorporated at this position (Val 64%). At position 25 usually a basic residue (40% of total), mostly Lys (30%) is found. The canonical sequence for these two positions is Lys25-Lys26, as it is found in α–spectrin - SH3 (SPC-SH3) (Fig. 5.1a), a domain that does not form amyloid deposits at low pH.   This region of the PI3-SH3 domain(residues 21-26) has a net charge of +1 at low pH, while the corresponding domain (residues 22-27) of α–spectrin has a net charge of +2. Recent studies have shown that the balance of charges could play an important role in driving amyloid fibril formation, and that indeed an increase of the net charge in an amyloidogenic polypeptide can severely interfere with its ability to form fibrils (*36, 37*) .

Taking the information from the NMR results, the mutation was made by swapping the amino acids at the diverging turn of both PI3-SH3 and SPC-SH3 to find out whether or not this region of the PI3-SH3 domain is important in the amyloid formation properties of the intact native protein. Six residues in PI3-SH3 which included the crucial histidine (21-26 of PI3-SH3) were swapped by the six residues in SPC-SH3 (22-27). Surprisingly, it was observed that mutated SPC-SH3 which carried the 6 amino acid sequence of PI3-SH3 formed amyloid fibrils, whereas the PI3-SH3 which carried the amino acids in the diverging turn of SPC-SH3 did not form fibrils under any conditions. These results showed clearly that a short amino acid sequence was sufficient to dramatically increase the amyloidogenicity of a globular protein (*38*).

5.4.3 DEPC modified PI3-SH3
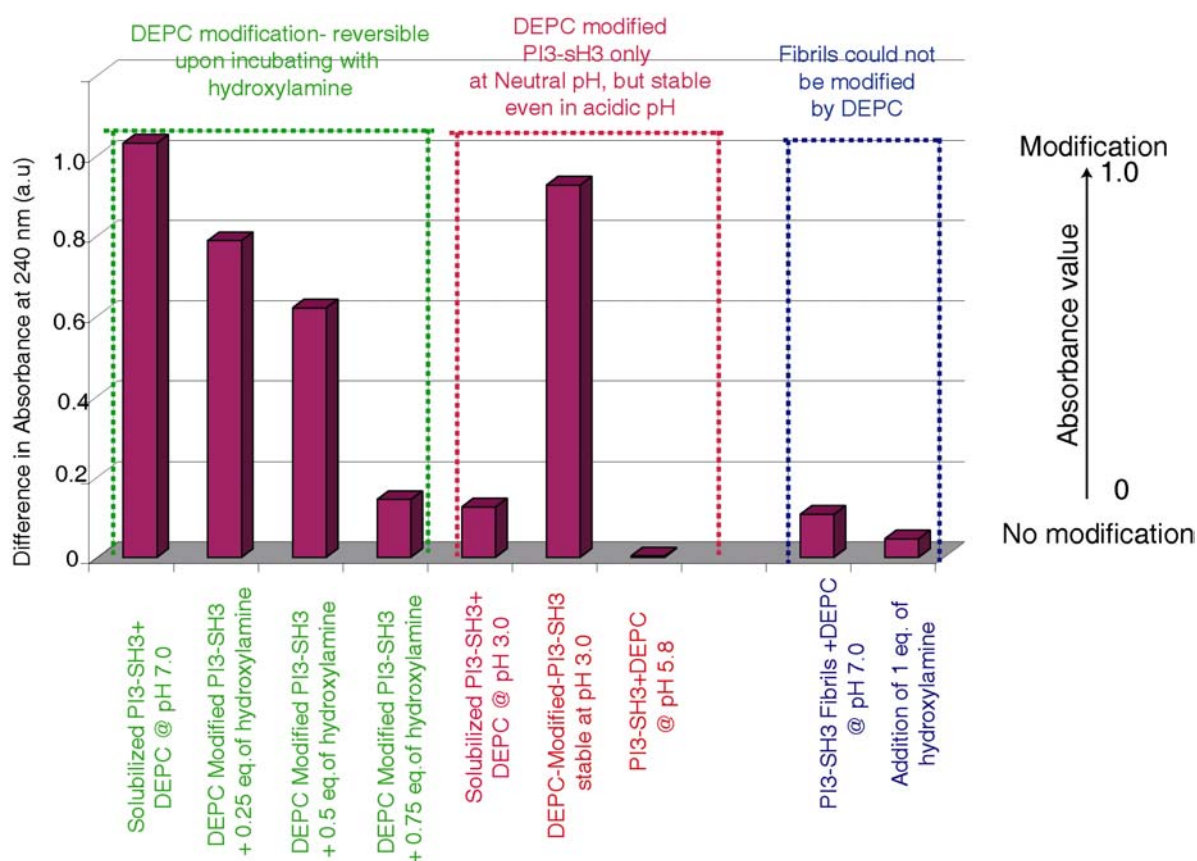
Figure 5.14. UV-spectrophotometer analysis of PI3-SH3 modification.

In order to understand the molecular mechanism of the involvement of histidine in fibril formation, histidine was modified as described in section 5.3. The modification can only be carried out above a pH value of 7. But, once modified, the carbethoxy-histidyl group was

observed to be stable even under acidic conditions, implying selectivity of the modification towards tautomeric histidine side chains (Fig. 5.14). Notably, fibrils of PI3-SH3 could not be modified at pH 7.0, indicating either burial of the histidine residue within the fibril core (Fig. 5.14).

5.4.4 Stability of histidine modified-PI3-SH3

Histidine modification of PI3-SH3 should have an impact on the stability against acid denaturation. From the solid-state NMR studies, it was shown that both the histidine side chains were involved in hydrogen bonding. Hence, replacing one of the hydrogens of the histidine side chains by a carbethoxyl group should either increase or decrease the stability of the protein. Bis-ANS fluorescence experiments were carried out in order to study the exposure of hydrophobic amino acid upon acid titration. The hydrophobic exposure is an indication for protein stability. The results suggested that histidine-modified-PI3-SH3 is slightly more stable than the unmodified PI3-SH3. A short amino acid sequence (ref. Material and Methods) was taken as a reference and as expected, hydrophobic exposure was stable throughout the pH range (Fig. 5.15).
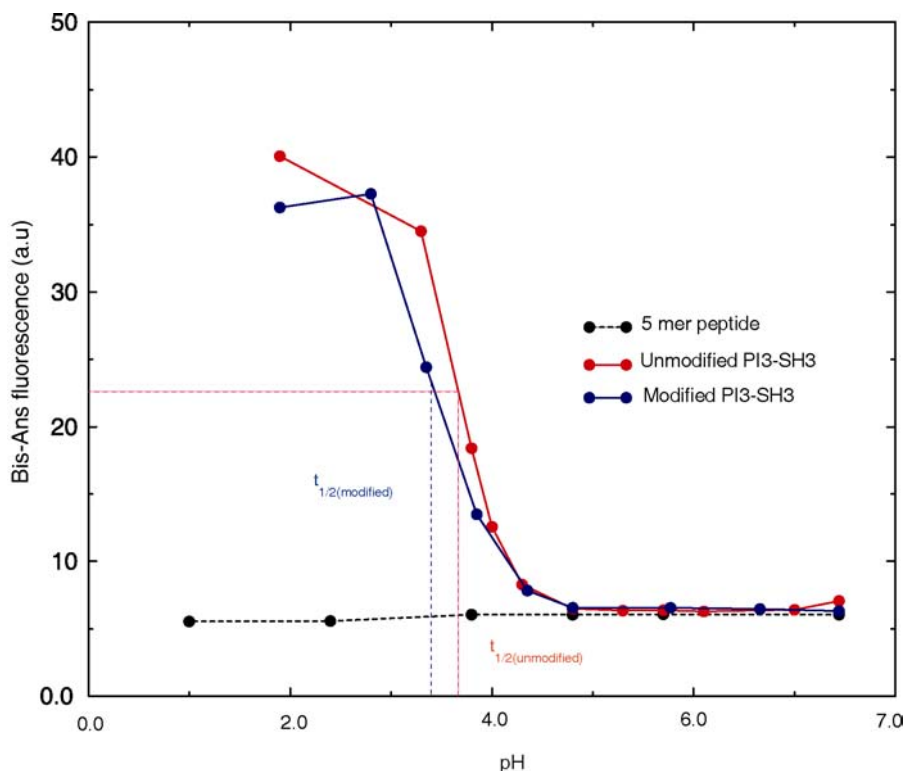
Figure 5.15. Bis-ANS fluorescence studies revealed that the modifed PI3-SH3 is slightly more stable than the unmodified one.

136

5.4.5 Limited proteolysis and gel filtration studies

Mass spectroscopic analysis shows that there are three modification in PI3-SH3 upon incubation with DEPC. The limited-proteolysis approach was used to locate the sites of modification. Trypsin and chymotrypsin digestion followed by MALDI-MS studies reveal that apart from H25, two of the seven tyrosines are also modified (Y8,Y76), even though tyrosines are considered as a poor nucleophile compared to histidines at the pH~7.

Gel filtration studies show that upon modification the molecular weight is increased from dimer (62.7 ml) to trimer (57.7 ml) (Fig.5.16).
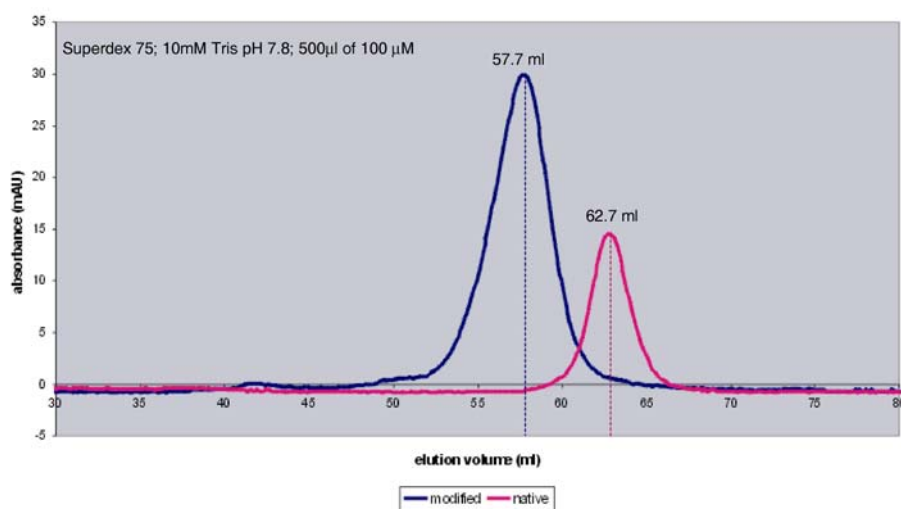
Figure 5.16. Gel filtration studies showed that modified PI3-SH3 existed in higher molecular weight oligomer (approximately a trimer) compared to the unmodified PI3-SH3 (dimeric).

5.4.6 Modified-PI3-SH3 formed fibrils
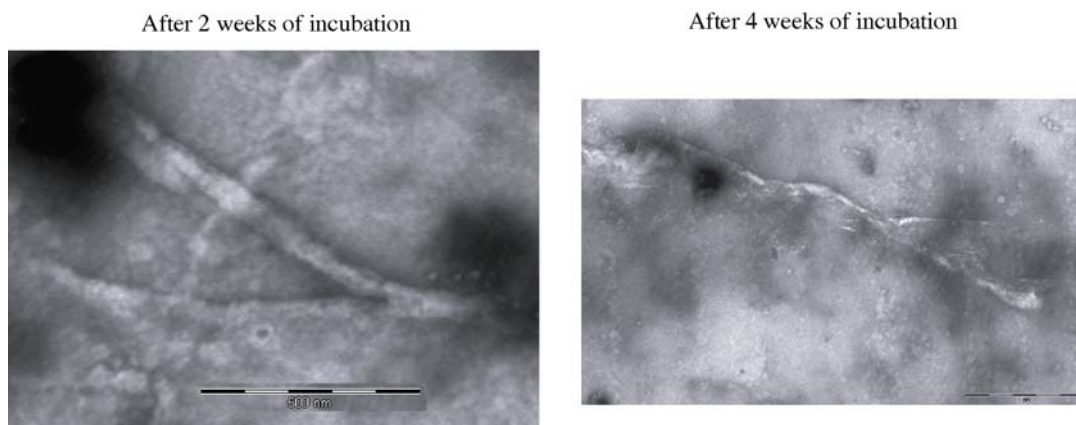
Figure 5.17. Modified PI3-SH3 formed fibrils after 2 weeks of incubation (left) , unmodified PI3-SH3 formed fibrils after 4 weeks (right). The scale represents 500 nm 100 nm for the left and right figures.

We were interested in finding out if the histidine modification of PI3-SH3 affected fibril formation *in vitro*. Therefore fibril formation was observed by electron microscopy for modified and unmodified PI3-SH3. The standard fibril formation protocol (incubation for 15 days at pH 2.0) was used for both the samples. Surprisingly, the modified PI3-SH3 formed fibrils in 2 weeks time, where as the unmodified protein formed ordered fibrils only after 4 weeks (Fig. 5.17)

5.4.7   Solution state NMR and light scattering analysis

a)   DOSY NMR analysis

Analysis of solid-state NMR spectra indicated that H25 is important in the fibrillar form to stabilize the structure (Section 5.4.1). Acid titration of both modified and unmodified PI3-SH3 was carried out to probe the initial events in fibril formation. The titration was carried out using the buffer as described in section 5.3.3.2.
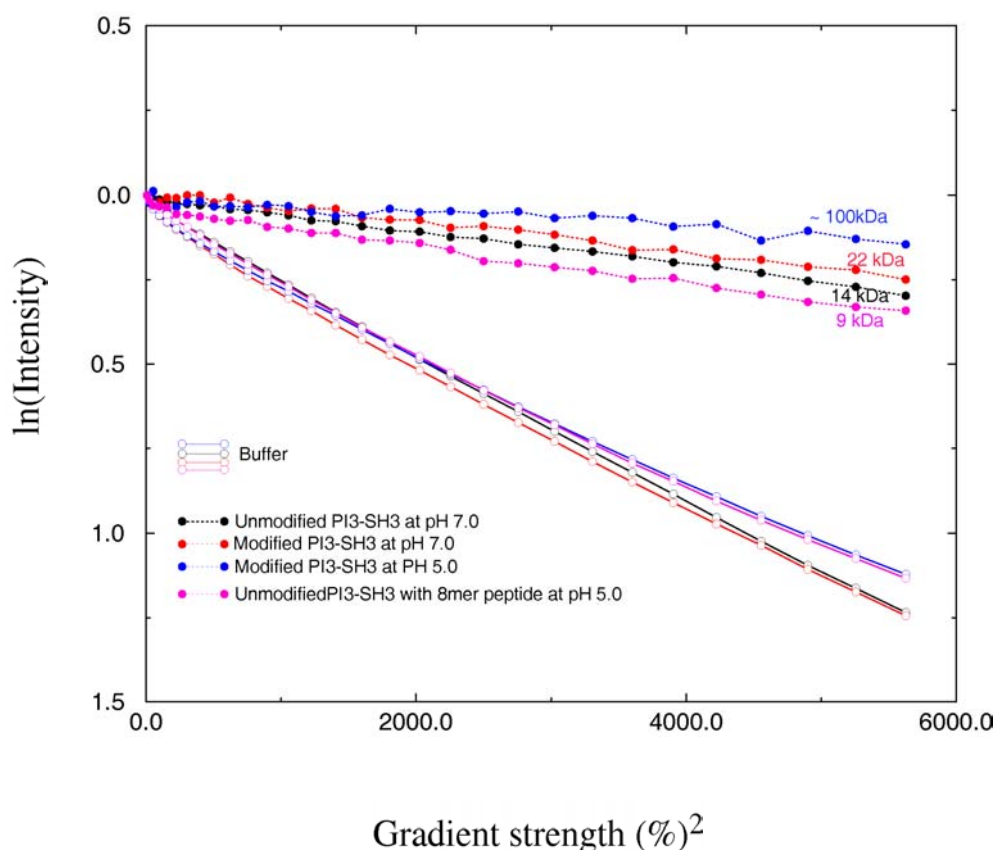
Figure 5.18a. DOSY NMR analysis of PI3-SH3 and the corresponding fitted average molecular weight.

DOSY-NMR experiments were used to analyze the change in molecular weight upon change in pH for the modified and unmodified-PI3-SH3 protein. Interestingly, we observed that the

unmodified PI3-SH3 at neutral pH diffuses with a molecular weight of 14±1 kDa (the expected value is 9±1 kDa).

Upon histidine modification, PI3-SH3 diffuses even more slowly as compared to the wild type. In this case, the diffusion constant fits to a value of 22±1 kDa. At pH 5.0, the effective diffusion constant of the modified PI3-SH3 was decreases significantly and yields a molecular weight of ~100 kDa. The exact molecular weight of the oligomers could not be obtained as the diffusion was too slow (Fig. 5.18a).

A synthetic short peptide which was derived from the RT loop of PI3-SH3 (residues 21 to 28 of PI3-SH3 which included the key His25) was titrated with unmodified PI3-SH3 and the observed diffusion profile corresponds to the monomeric form of PI3-SH3. This result implies that a small stretch of amino acids derived from the RT-loop can modulate the oligomerization behaviour (Fig. 5.18a). Molecular weight analysis has been performed earlier using HP-SEC in which a value of 14 kDa was fitted under native conditions and a value of 9 kDa was fitted under denaturing conditions, corroborating the DOSY results (39). From these results, we conclude that PI3-SH3 exists in equilibrium between monomer and higher-order oligomers. Therefore, the average molecular weight of the PI3-SH3 is shifted to higher values. At the same time, oligomerization can be inhibited by titrating PI3-SH3 with the 8-residue-peptide derived from the RT loop. Histidine modification tremendously affected the equilibrium as it pushed equilibrium towards higher-order oligomers.

b) Light scattering experiments

In order to confirm the NMR results and to obtain a more accurate information about the distribution of the oligomers that lead to fibril formation, light scattering experiments were performed in collaboration with Dr. Klaus Gast, MDC-Berlin, Germany. The experiments were performed initially to understand the differences in oligomeric distributions of modified and unmodified PI3-SH3 at pH 7.4. We obtained a hydrodynamic radius of 1.7 nm for the both modified and unmodified PI3-SH. This value corresponds to the monomeric form of PI3-SH3 (Fig. 5.18b). In addition, higher-order oligomers with hydrodynamic radii of 11 nm in the case of unmodified PI3-SH3 and 4.5 nm and 12.5 nm in the case of modified PI3-SH3 were observed (Fig 5.18b). The population of higher order oligomers was higher in the case of modified PI3-SH3. These results are supported by the DOSY-NMR experiments where the molecular weight of the modified PI3-SH3 was fitted to higher values compared to the diffusion data obtained for unmodified protein (Fig. 5.18a).
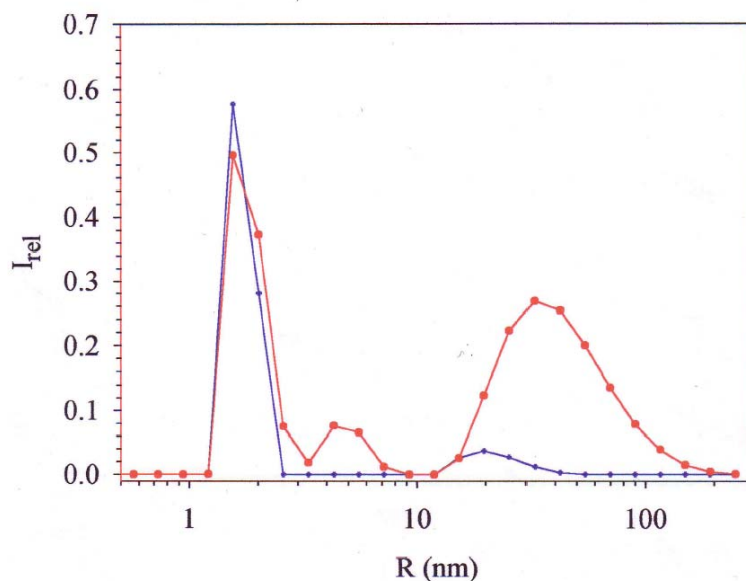
Fig 5.18b Distribution of oligomers of PI3-SH3 as observed by light scattering experiments. The unmodified PI3-SH3 is represented in blue, where as the modified PI3-SH3 is represented in red. The experiments were performed using 5 mg/ml of PI3-SH3 dissolved in 20 mM of phosphate, 20 mM of citrate, pH 7.4 at 27 ° C.

The existence of oligomers with a molecular weight of ~100 kDa for the modified PI3-SH3 at pH 5.5 as observed by DOSY experiments was again confirmed by light scattering experiments. After shifting the pH values from pH 7.4 to pH 5.4, a migration of lower-order oligomers with a hydrodynamic radius of 10.5 nm to higher order oligomers with a hydrodynamic radius of 12.5 was observed as a function of time (Fig. 5.18c).
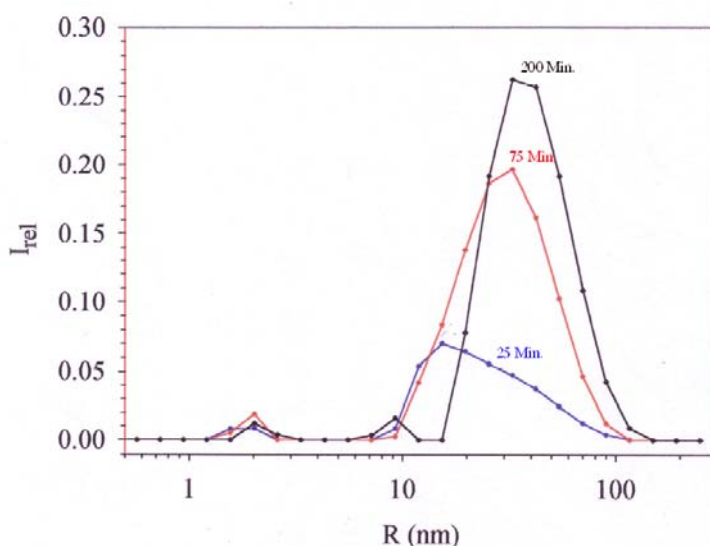
Fig 5.18c Distribution of oligomers of modified-PI3-SH3 at pH 5.4 as a function of hydrodynamic radius. The colour of the light scattering curves corresponds to the scattering intensity obtained at different time points as represented in Fig. 5.18d.

After shifting the pH from 7.4 to 5.4, the hydrodynamic radii of modified PI3-SH3 was monitored as a function of time. We observe an approximately linear increase in the scattering intensity as a function of time (Fig. 5.18d). This implies that fibrils grow by monomer addition.
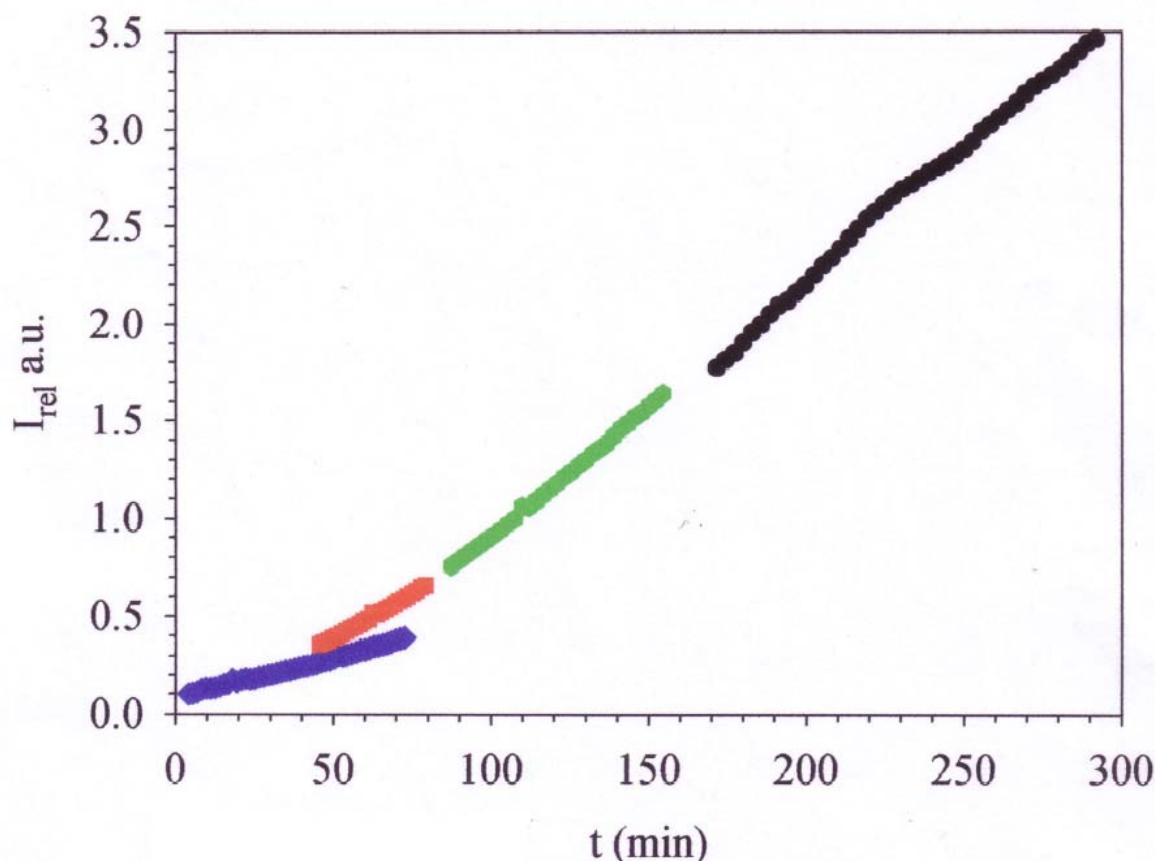
Fig 5.18d Scattering relative intensity as a function of time for the modified PI3-SH3 sample in 20 mM of phosphate, 20 mM of citrate, pH 5.4 at 27 °C . The different colours correspond to the populations of different oligomeric species. (The two different experiments were super imposed, red, green and black correspond to a single experiment and the violet correspond to another experiment for the same sample under the same conditions).

Combining DOSY and light scattering experiments, we conclude that a change in single amino acid can modulate the quaternary arrangement of PI3-SH3.

c) HSQC Titration studies of PI3-SH3

To find out the site specific contributions during the acid titration on PI3-SH3, HSQC spectra were monitored. No major chemical shift changes for PI3-SH3 upon transition from the pH

7.5 to 5 were observed (Fig. 5.19). In the primary structure the amino-acids that are affected during the pH shift are highlighted in yellow.
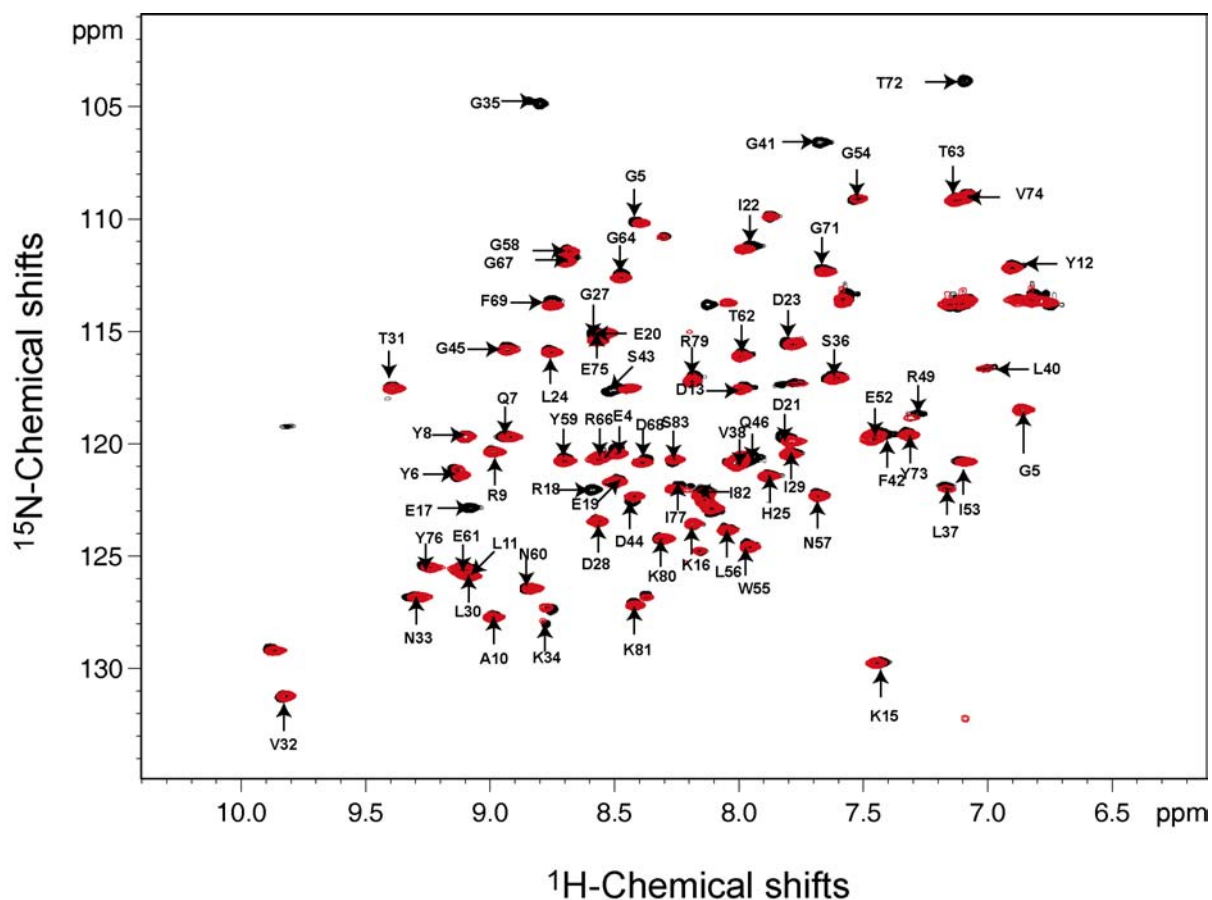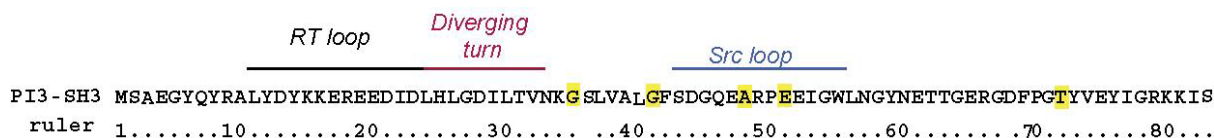
Figure 5.19. HSQC of unmodified PI3-SH3 at pH 7.4 (black) and pH 5.0(black) displays no major chemical shifts.

In addition, we do not observe denaturation of the modified PI3-SH3 while shifting the pH to 3.0. However, the unmodified protein showed a typical acid denaturation spectrum (Fig. 5.20). This result demonstrates that the modified PI3-SH3 is more stable and corroborates the results from the Bis-ANS studies (Fig. 5.15)
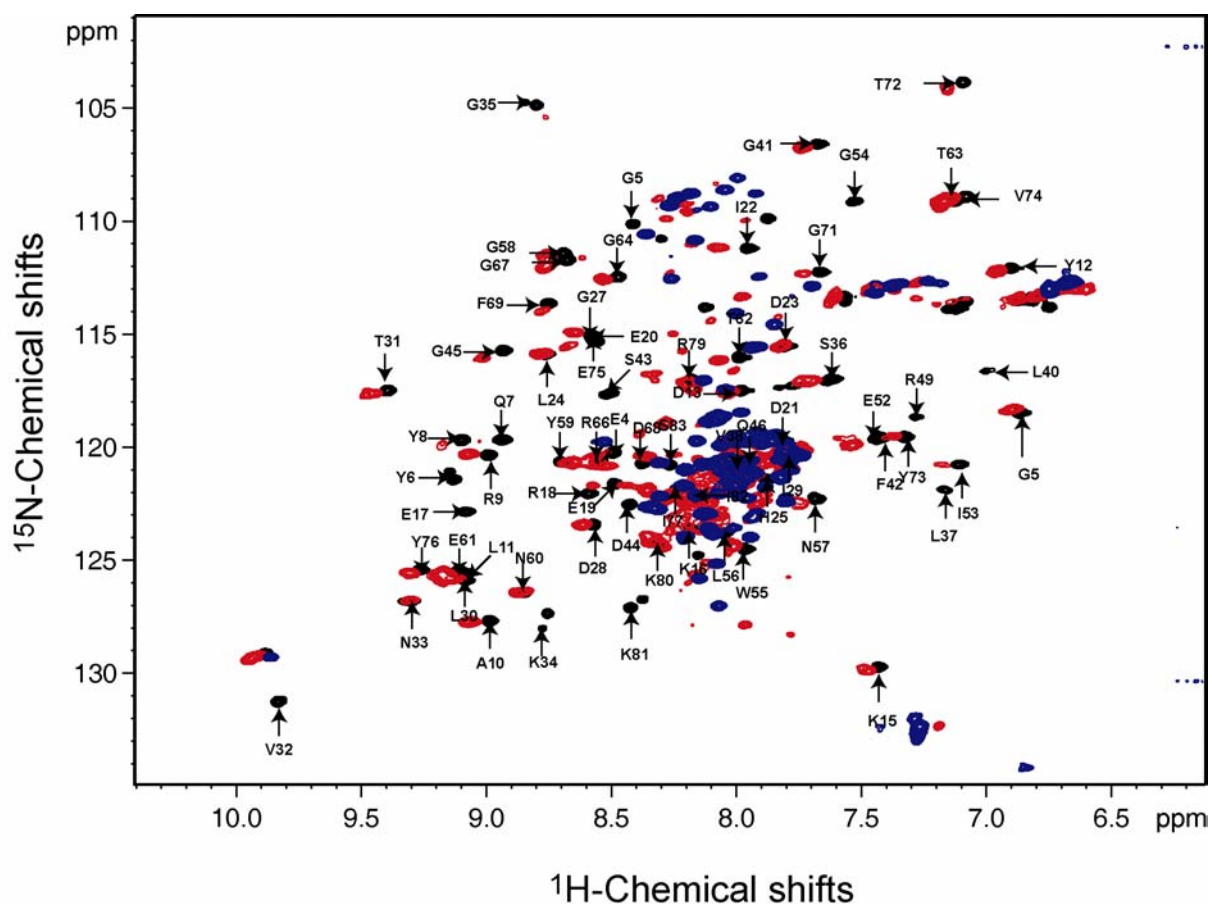
Figure 5.20 HSQC spectra recorded for modified PI3-SH3 (red) and unmodified PI3-SH3 (blue) at pH 3.0. For reference the spectrum at pH 7.0 is represented (black).

From the DOSY analysis, a sample of PI3-SH3 incubated together with, a 8mer peptide (sequence: DIDLHLGD) at pH 5.0 showed a predominantly monomeric diffusion profile (Fig. 5.18a). The peptide was derived from residues 21-28 in PI3-SH3 corresponding to the amino acids from RT-loop (synthesized by Ms. Zhongjing Chen, FMP-Berlin). This indicates that the peptide influenced the protein-protein contacts. Hence, this peptide was used to study sequence specific interactions of the PI3-SH3 during association. Titration of the 8mer peptide with PI3-SH3 at a molar ratio of 2:1 for [8mer]:[PI3-SH3], at pH 7.4 did not induce any chemical shift changes in the HSQC spectrum (Fig. 5.21).
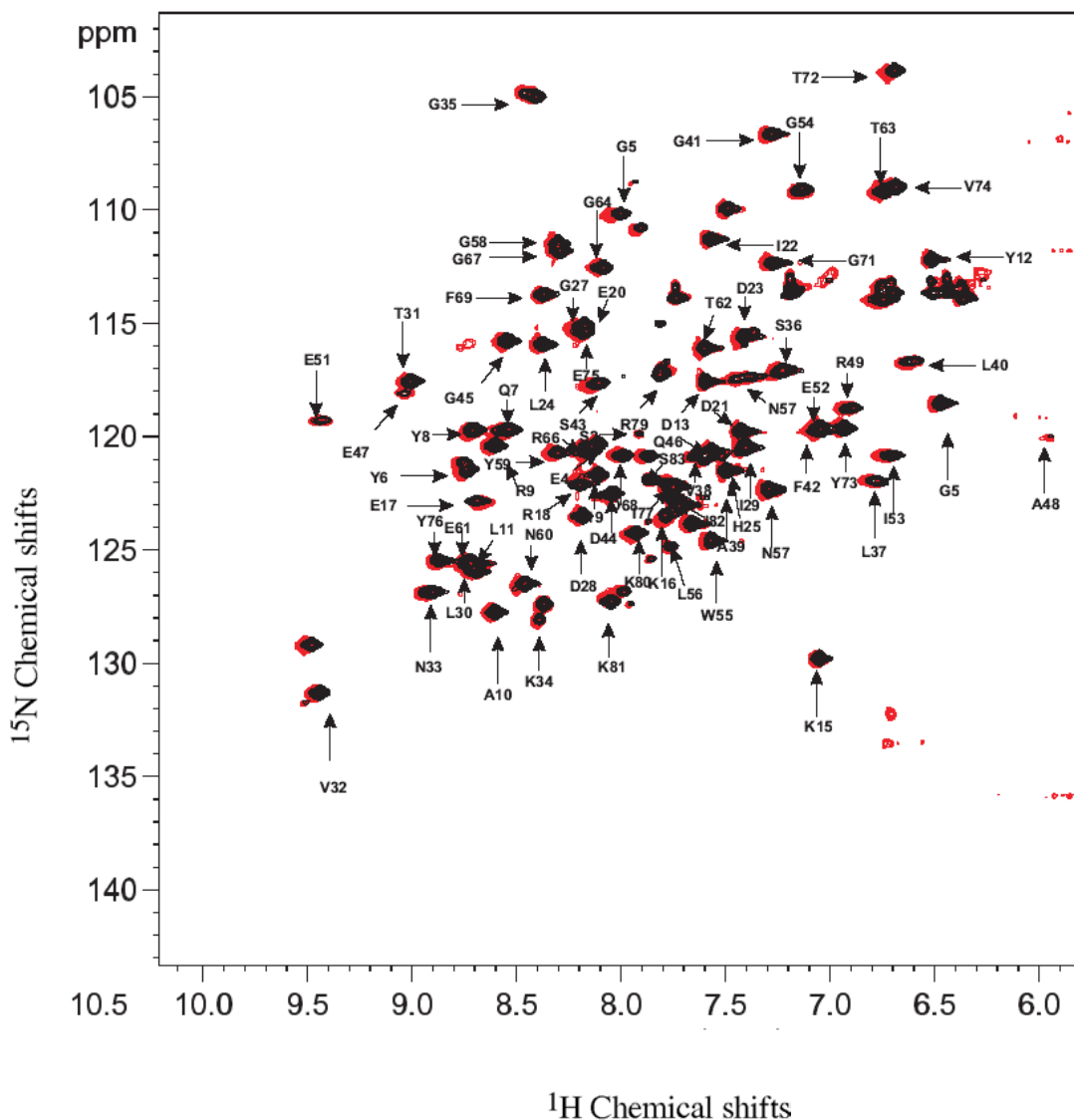
Figure 5.21. No changes in chemical shifts are observed upon addition of the 8mer peptide to PI3-SH3 (red) compared to PI3-SH3 alone (black) at pH 7.4.

But shifting the pH from 7.5 to value of 5.0, changes in chemical shifts of certain amino-acids were observed for the PI3-SH3:8mer mixture (Fig. 5.22). A sequence specific analysis was complicated by the fact that many resonance frequencies changed at pH 5.0, since the pH value is near to the pI of the PI3-SH3 (pI= 4.3). Careful analysis of the chemical shift changes gave an early indication that the region around the RT and Src loops were affected mostly during acid titration (Fig. 5.22). Notably, in absence of 8mer peptide, the PI3-SH3 showed chemical shift differed in only few amino acid residues during this pH shift (Fig. 5.19).

The amino-acids highlighted in yellow in the primary structure of PI3-SH3 given below show the largest chemical shift perturbations:
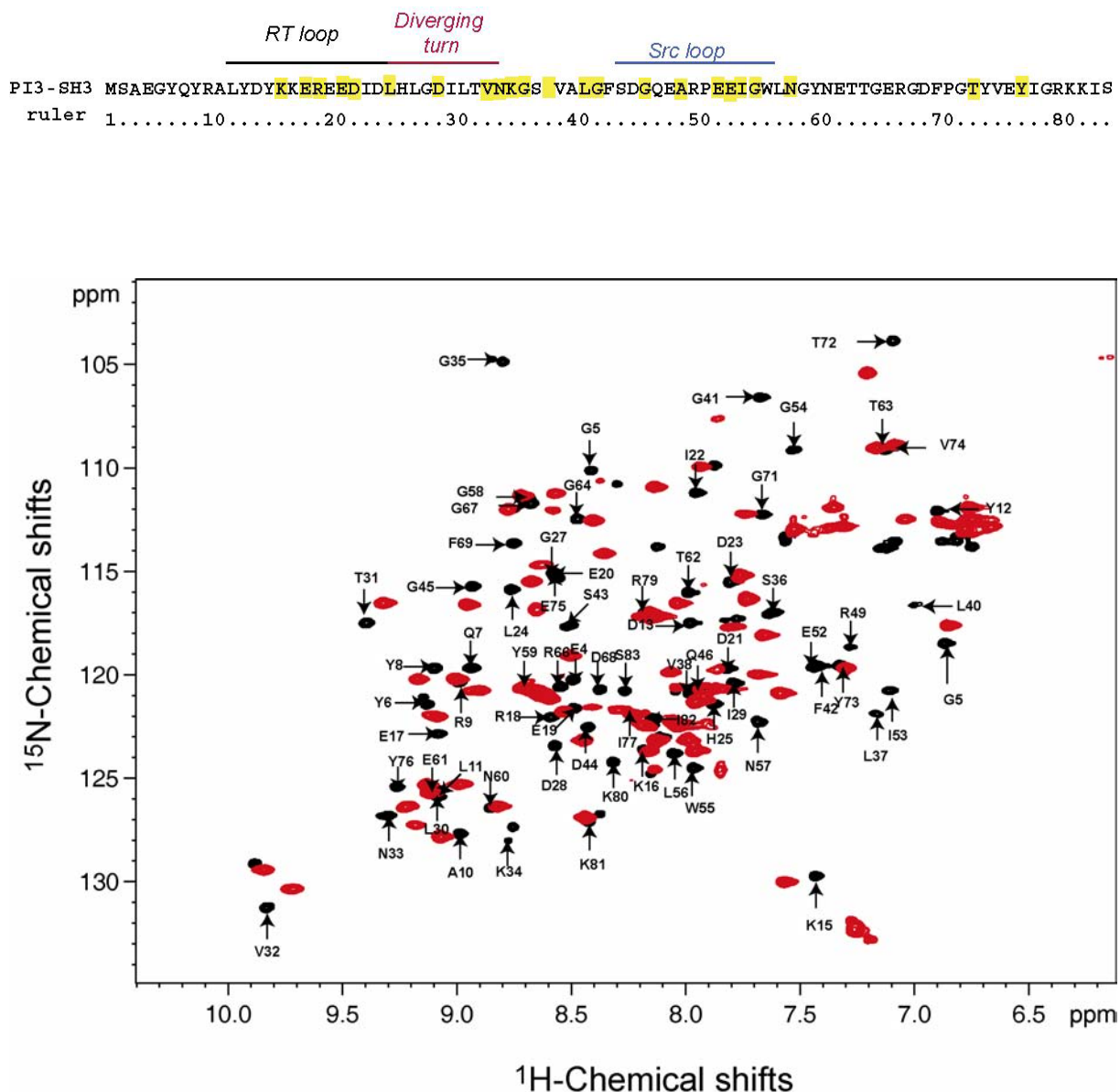
Figure 5.22 The majority of the amino-acids in the RT and Src-loop region of PI3-SH3 are shifted upon addition of 8-mer peptide at pH 5.0 (Red). The reference spectrum (black) was taken at pH 7.4.

Hence we conclude that the titration of PI3-SH3 with the synthetic peptide derived from RT-loop affected the Src loop (Fig. 5.22). But, previous studies observed no influence on fibril formation of PI3-SH3 upon deletion of its Src-loop (*6*). This observations imply the question, if the interaction between RT and Src loop occurs during the early steps of aggregation?

d) **Isotope filtering experiments identified intermolecular contacts**
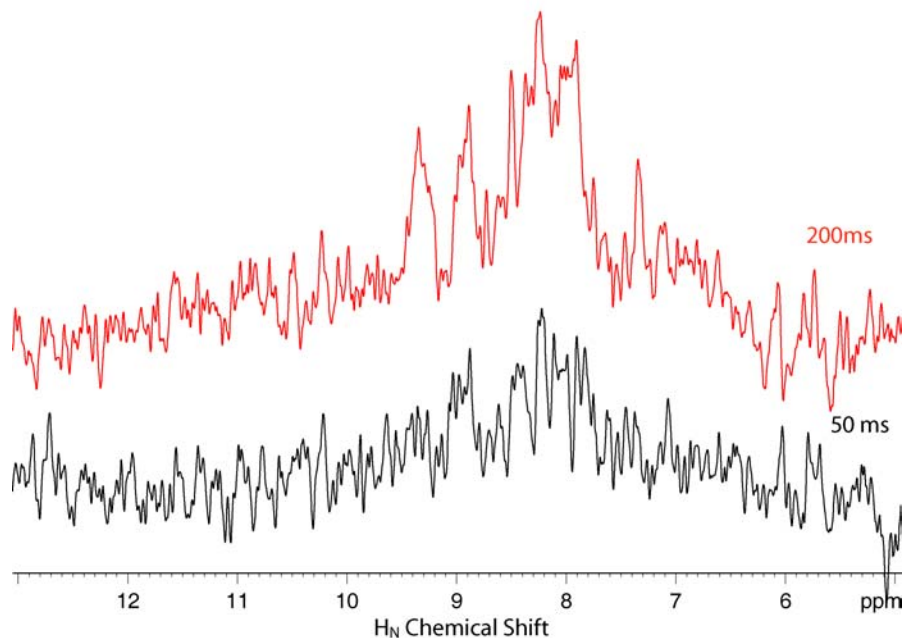
F*igure 5.23. 1D version of a* ^{13}C *edited NOE-*^{15}N *edited 2D* ^1H-^1H *correlation experiment recorded for the modified PI3-SH23 at two different mixing times indicating that the observed signals were purely due to NOE mixing.*

Differentially labelled samples (^{15}N or ^{13}C) using a molar ratio of 1:1 were employed in isotopic editing (editing here means selection). The magnetization of ^{13}C was selected from the ^{13}C labelled molecules and stored along the Z-axis of H_C $\left(H_z^c \right)$. The magnetization was then distributed both inter and intramolecularly through NOE interactions. After the NOE mixing time, the protons that are bound to ^{15}N were selected and detected (Fig. 5.23). Using the modified PI3-SH3 (pH 7.4), the initial experiment was carried out by recording ^1H-^1H correlation experiments editing for ^{13}C prior to t1 and editing for ^{15}N prior to acquisition.. The experiment detects the intermolecular contacts. The majority of the contacts were between H_N and Hα region, suggesting the involvement of backbone-backbone interactions (Fig. 5.24).

146
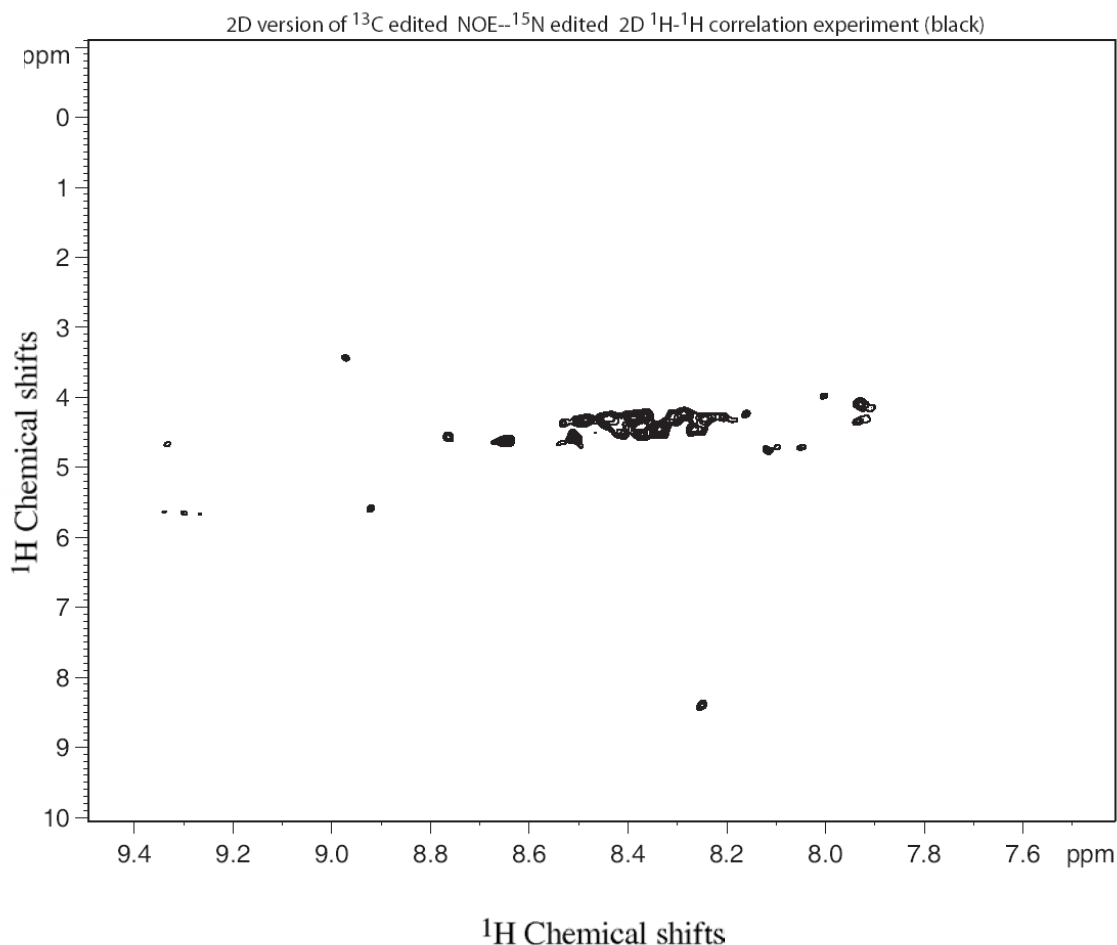
Figure 5.24. 2D-(^{13}C)^{1}H -NOE-(^{15}N)^{1}H correlation indicating backbone-backbone contacts for the modified PI3-SH3 (pH 7.4). In this experiment, no site specific resolution could be obtained due to the poor resolution in the ^{1}H dimension.

No site-specific information could be obtained in this experiment due to the poor resolution in the ^{1}H dimensions of the 2D isotope filtered experiment. Therefore, 3D-experiments were carried out on the unmodified PI3-SH3, including an additional carbon / proton evolution before the NOE period. Unmodified PI3-SH3 was used for this study as it was also shown to exist in equilibrium with higher ordered oligomers in gel filtration studies (Fig. 5.16), DOSY analysis (Fig.5.18a) and scattering experiments (Fig. 5.18 b, c). Though the population of higher order oligomers in unmodified PI3-SH3 at native pH are very low, it could be increased at a higher concentration of PI3-SH3.

2D projection of uniformly labelled (red) and differentially labelled (green) PI3k-SH3 samples
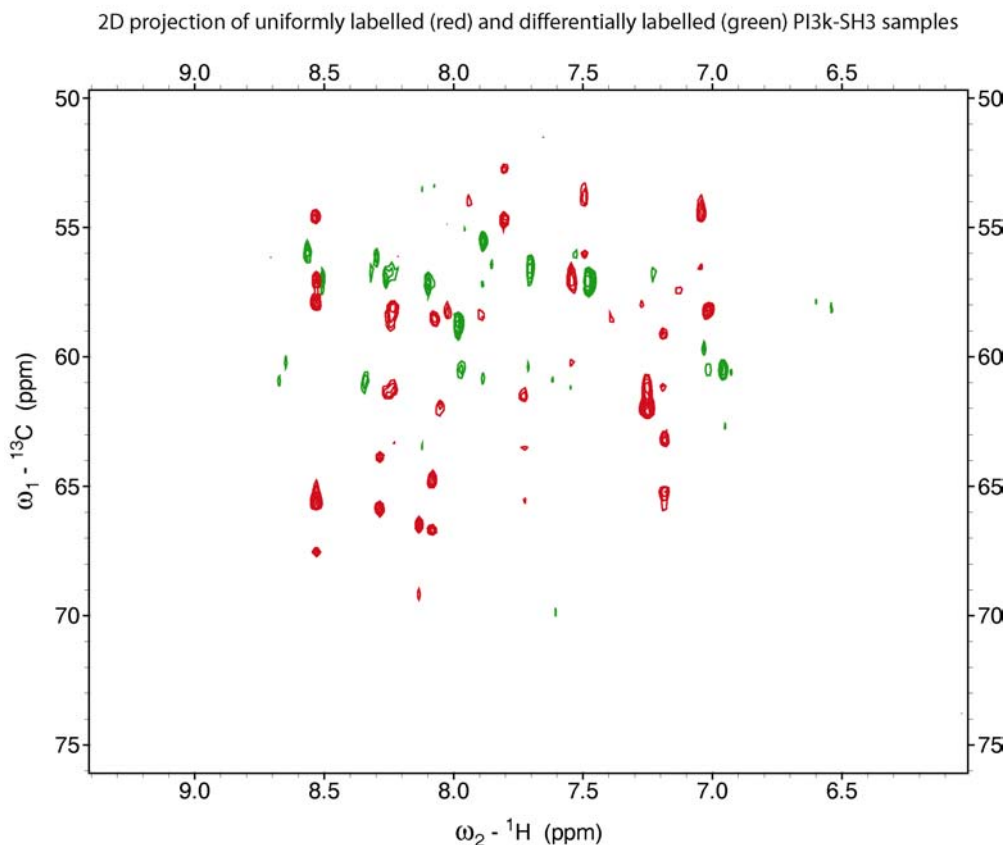
Figure 5.25. 2D projection of the 3D $^{13}C(1H)$- NOE-$^{15}N,^{1}H$ isotope filtered experiment for the uniformly labeled PI3-SH3 (red) and differentially labeled protein (green). The observed cross peaks in the differentially labeled samples must therefore be due to intermolecular contacts.

In the figure 5.25, 2D projections of the 3D $^{13}C(1H)$- NOE-$^{15}N,^{1}H$ isotope filtered experiment for the uniformly labelled PI3-SH3 using the pulse programme as depicted in the Fig 5.6a and the differentially labelled PI3-SH3 using the pulse programme as depicted in the Fig 5.6b are compared. From comparison, it is clear that the observed signals in the mode must be due to intermolecular NOE contacts. Therefore selection of natural abundance ^{13}C can be completely ruled out (Fig. 5.25). Careful analysis of 3D spectra yields site specific information protein contact points : A10, L11, D13, K15, L24, H25, S43, D44,G45, Q46,E52, L46,N57,G58, E61. Interestingly, only the specific amino acids could be identified. Others are not visible at all. The stretch of amino acids are indicated in Fig. 5.26
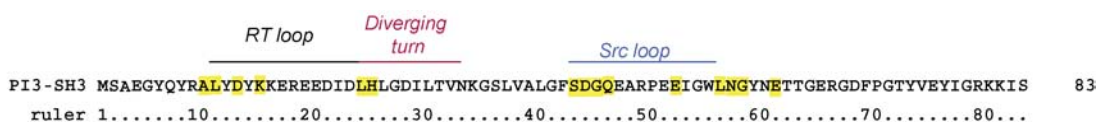
Figure 5.26. Observed intermolecular NOE signals mapped to the primary structure

Surprisingly, they are located in the RT and Src loops. This indicates that the RT and Src loops are involved in inter-molecular interactions suggesting the possible existence of domain swapping involving the RT and Src loop. Detailed analysis of the path is currently going on to identify the interaction partner for each amino acids that showed inter-molecular contacts.

Three-dimensional (3D) domain swapping is a mechanism of exchanging one structural domain of a protein monomer with an identical domain from a second monomer, resulting in an intertwined oligomer. The swapped domain has nearly identical non-covalent interactions in the oligomer as in the monomer. More than a dozen crystal structures have been determined of dimers and trimers that are 3D domain swapped (*40, 41*). 3D domain swapping may also lead to polymerization, and has been suggested as a mechanism of forming protein amyloid fibrils. The amyloid fibril forming proteins Cystatin-C and human-prion proteins exist as domain swapped dimers (*42, 43*). One of the SH3 domain proteins (Eps8-SH3) has also been shown to exist in monomeric and dimeric forms (*44, 45*) (Fig. 5.27).
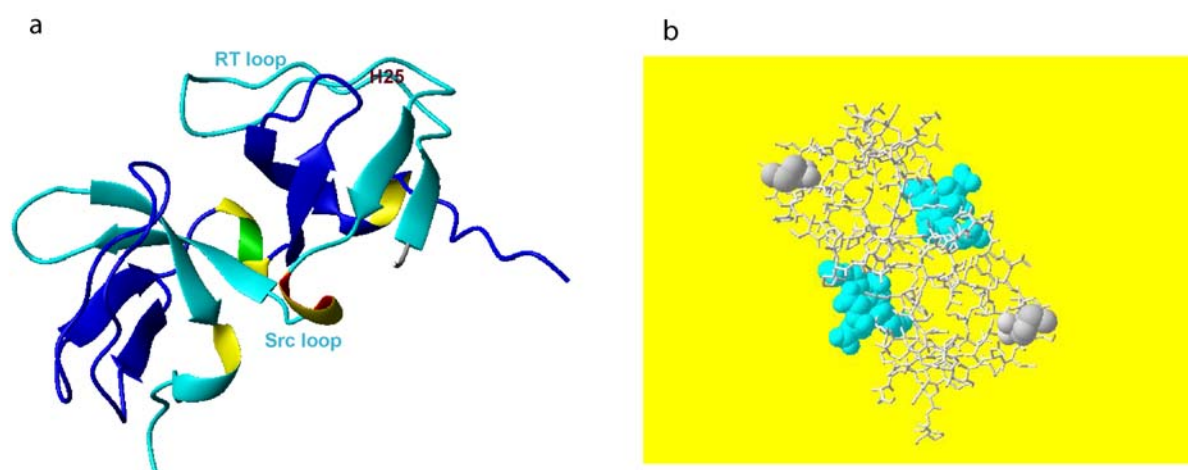
Figure 5.27. Dimeric structure of Eps8-SH3 (PDB:1AOJ). (a). The space filling in blue indicated the Src loop which is located at the interface between two monomers (b) The amino acid 25 of the RT loop in Eps-8.SH3 is indicated in grey (b)

Looking into the structure of Eps8-SH3, the Src-loop lies at the dimeric interface. Likewise, comparing at the monomer and dimer structure, the fold is almost identical looking at the N-terminal to the Src-loop. From this point, the structures diverge (Fig. 5.28). In the case of PI3-SH3, the Src-loop is even longer than Eps8-SH3, making an involvement of the Src-loop in oligomerization more plausible.
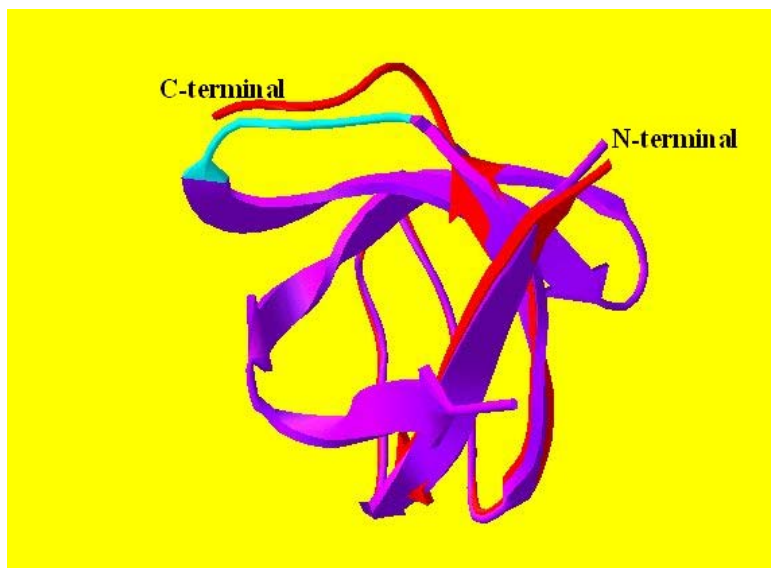
Figure 5.28 Monomeric and dimeric structures of Eps8-SH3 showed identical structure till the beginning of Src loop.

From the light scattering experiments (Fig. 5.18 b-d), monomer addition to the nucleus could be a possible explanation for the fibril formation mechanism of PI3-SH3. From the DOSY NMR experiments (Fig. 5.26), we suggest that the monomers are in equilibrium with low-molecular weight oligomers. We observe contacts between the RT and Src loop regions in isotope filtered NMR experiments.

Can the domain swapping explain the protein polymerization?

Taking all the results together, occurrence of amyloid fibril formation through domain swapping between monomeric species may be plausible. According to the proposal of Bennett and co-workers, the interface in a closed monomer can be disrupted under certain conditions to form an open monomer (*46*). There is a high energy difference, i.e., the activation energy between the closed and the open form of monomers. Two or more open monomers can aggregate to form a domain-swapped dimer or oligomer. The free energy between the closed form of the monomer and the domain-swapped oligomer must be small because they share the same structures except for the hinge loop. Therefore, both structures are thermodynamically feasible. The initial activation energy barrier can be reduced under certain circumstances, such as change of pH, change of temperature, mutation in the protein, presence of denaturants, and binding of a ligand. A high energy barrier can be reduced by a change of the solution conditions. Recent studies on 3D domain swapping showed that there are three factors affecting the free energy difference between the monomer and the domain-swapped oligomer (*47-49*). First, the greater entropy of the monomer makes it more favored thermodynamically. Second, hinge loops may form new interactions in the domain-swapped dimer, which favour

dimerization. Also, there may be strains introduced or relieved when a protein forms a domain-swapped dimer. Therefore, the conformational changes at the hinge loop also contribute to this free energy difference. Third, new interactions at the open interface make the domain-swapped oligomer more favorable thermodynamically. Therefore, by changing the hinge loop or engineering the open interface, one can change the equilibrium between the monomer and the domain-swapped oligomer. In the case of PI3-SH3 it is likely that other mechanisms in addition to domain-swapping is effective as previous results have shown that deletion of the loop did not prevent PI3-SH3 to form amyloid like fibrils (*6*). However, changing two amino acids in the exposed RT-loops abolished fibril formation (*38*). Hence, it could be that interactions are mediated through charged amino acids. Taking all above factors into consideration, we can conclude that fibril formation of PI3-SH3 may occur either through domain swapping in higher-order oligomeric states or through interactions of exposed charged amino acids stretch that accelerate the population of higher-order oligomers.

5.5   Conclusion

Combining solid-state, solution-state NMR and biochemical results, we conclude that a small stretch of residues in a protein is enough to form fibrils from a soluble protein. The fibril formation of PI3-SH3 which occurs at acidic pH may be occurring even at the physiological pH, as intermolecular interactions were observed at higher concentration of PI3-SH3. Modification of a single residue indicates changes in the quaternary arrangement of the protein, indicating the importance of hot spot residues in the protein structure. We postulate that PI3-SH3 may use open domain swapping to form fibrils, in which the net charge at the open end of the RT loop drives the aggregation.

5.6   References

1.   Bucciantini, M., Elisa,G., Chiti, F., Baroni, F., Formigli, L., Zurdo,J., Taddei, N., Ramponi, G., Dobson,C.M., Stefani, M. (2002) *Nature 416*, 507-511.
2.   Glenner, G. G., Eanes, E.D., Bladen, H.A., Linke, R.P., Termine, J.D. (1974) *J.Histochem.Cytochem. 22*, 1141-1158.
3.   Guijarro, I., Sunde, M., Jones, JA., Campbell, ID., Dobson, CM. (1998) *Proc. Natl. Acad. Sci. USA 95*, 4224-4228.
4.   Frandrich, M., Fletcher, M.A., Dobson, C.M. (2001) *Nature 410*, 165-166.
5.   Musacchio, A., Gibson, T., Lehto, V.P., Saraste, M. (1992) *FEBS Lett. 307*, 55-61.
6.   Ventura, S., Lacroix, E., Serrano, L. (2002) *J. Mol. Biol. 322*, 1147-1158.
7.   Zurdo, J., Guijarro, J.I., Dobson, C.M. (2001) *J.Am.Chem.Soc. 123*, 8141-8142.
8.   Cross, T. A., Frey, M.H., Opella, S.J. (1983) *J.Am.Chem.Soc. 105*, 7471-7473.
9.   Reif, B., van Rossum, B.J., Castellani, F., Rehbein, K., Diehl, A., Oschkinat, H. (2003) *J.Am.Chem.Soc. 125*, 1488-1489.
10.   Goldburg, W., Lee, M. (1963) *Phys.Rev.Lett. 11*, 255-260.
11.   Vinogradow, E., Madhu, P.K., Vega, S. (1999) *Chem.Phys.Lett. 1999*, 443-450.
12.   Hohwy, M., Jaroniec, C.P., Reif, B., Riesentra, C.M., Griffin, R.G. (2000) *J.Am.Chem.Soc. 122*, 3218-3219.
13.   Miles, E. W. (1977) *Methods Enzymol. 47*, 431-442.
14.   Shimba, N., Serber, Z., Ledwidge, R., Miller, S.M., Craik, C.S., Dötsch, V. (2003) *Biochemistry 42*, 9227-9234.
15.   Farr-Jones, S., Wong, W.Y.L., Gutheil, W.G., Bachovchin, W.W. (1993) *J.Am.Chem.Soc. 115*, 6813-6819.
16.   Hlavica, P., Lehnerner, M., Eulitz, M. (1996) *Biochem.J 318*, 857.
17.   Smoot, A. L., Panda, M., Brazil, B.T., Buckle, A.M., Fersht, A.R., Horowitz, P.M. (2001) *Biochemistry 40*, 4484-4492.
18.   Ferrage, F., Zoonens, M., Warschawski, D.E., Popot,J.L:, Bodenhausen,G. (2003) *J.Am.Chem.Soc. 125*, 2541-2545.
19.   Yu, H., Rosen, M.K., Schreiber, S.L. (1993) *FEBS Lett. 324*, 87-92.
20.   Booker, G. W., Gout, I., Downing, A.K., Driscoll, P.C., Boyd, J., Waterfield, M.D., Campbell, I.D. (1993) *Cell 73*, 813-822.
21.   Kay, L. E., Ikura, M., Tschudin, R., Bax.A. (1990) *J.Magn.Reson. 89*, 496-514.
22.   Vuister, G. W., Bax, A. (1993) *J.Am.Chem.Soc. 115*, 7772-7777.
23.   Kay, L. E., Ikura, M., Bax.A. (1990) *J.Am.Chem.Soc. 112*, 888-889.
24.   Furrer, J., Kramer, F., Marino, J.P., Glaser, S.J., Luy, B. (2004) *J.Magn.Reson. 166*, 39-46.
25.   Goddard, T. D., Kneller, D.G. *SPARKY 3*, University of California, San Francisco.
26.   Breeze, A. L. (2000) *Prog.Nucl.Magn.Reson.Spect 36*, 323-372.
27.   Tycko, R. (2001) *Methods Enzymol. 339*, 390-413.
28.   Smernik, R. J., Baldock, J.A., Oades, J.M. (2002) *Solid state Nucl.Mag.Reson. 22*, 71-82.
29.   Fuess, H., Hohlwein, D., Mason, S.A. (1977) *Acta Cryst B33*, 654-659.
30.   Wei, Y., Dios, A.C.D., Mc Dermott, A.E. (1999) *J.Am.Chem.Soc. 121*, 10389-10394.
31.   Orchin, M., Jaffé, H. (1967) *The importance of antibonding orbitals*, Houghton mifflin, Boston.
32.   Song, X., Rienstra, C.M., McDermott, A.E. (2001) *Magn.Reson.Chem. 39*, S30-S36.
33.   Bloomberg, F., Maurer, W., Rüterjans, H. (1977) *J.Am.Chem.Soc. 99*, 8149-8159.
34.   Perutz, M. F., Finch, J.T., Berriman, J., Lesk,A. (2002) *Proc. Natl. Acad. Sci. USA 99*, 5591.

35.	Elam, J. S., Taylor, A.B., Strange, R., Antonyuk,S., Doucette, P.A., Rodriguez, J.A., Hasnain, S.S., Hayward, L.J., Valentine, J.S., Yeates, T.O., Hart, P.J. (2003) *Nat.Struct.Biol. 10,* 461-467.

36.	Lopez delapaz, M., Serrano, L. (2004) *Proc. Natl. Acad. Sci. USA 101*, 87-92.

37.	Chiti, F., Calamai, M., Taddei, N., Stefani, M., Ramponi,G., Dobson, C.M. (2002) *Proc. Natl. Acad. Sci. USA 99*, 16419-16426.

38.	Ventura, S., Zurdo, J., Narayanan, S., Parreno, M., Mangues, R., Reif, B., Chiti, F., Giannoni, E., Dobson, C.M., Aviles, F.X., Serrano, L. (2004) *Proc. Natl. Acad. Sci. USA 101*, 7258-7263.

39.	Harpur, A., Layton, MJ., Das, P, Bottomley, MJ., Panayotou, G., Driscoll, PC, Waterfield, MD. (1999) *J.Biol.Chem. 274*, 12323–12332.

40.	Bennett, M. J., Choe, S., Eisenberg, D. (1994) *Proc. Natl. Acad. Sci. USA 91*, 3127-3131.

41.	Liu, Y., Eisenberg, D. (2002) *Prot.Sci. 11*, 1285–1299.

42.	Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., Jaskolski, M. (2001) *Nat.Struct.Biol. 8*, 316-320.

43.	Knaus, K. J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W.K., Yee, V.C. (2001) *Nat.Struct.Biol. 8*, 770-774.

44.	Kishan, K. V., Scita, G., Wong, W. T., Di Fiore, P. P., Newcomer, M. E. (1997) *Nat.Struct.Biol. 4*, 739.

45.	Kishan, K. V. R., Newcomer, M.E., Rhodes, T.H., Guilliot, S.D. (2001) *Prot.Sci. 10*, 1046-1055.

46.	Bennett, M. J., Schlunegger, M.P., Eisenberg, D. (1995) *Protein Sci. 4*, 2455-2468.

47.	Kuhlman, B., O'Neill, J.W., Kim, D.E., Zhang, K.Y., and Baker, D. (2001) *Proc. Natl. Acad. Sci. USA 98*, 10687–10691.

48.	Rousseau, F., Schymkowitz, J.W., Wilkinson, H.R., and Itzhaki, L.S. (2001) *Proc. Natl. Acad. Sci. USA 98*.

49.	Schymkowitz, J. W., Rousseau, F., Wilkinson, H.R., Friedler, A., and Itzhaki,L.S. (2001) *Nat. Struct. Biol. 8*, 888-892.

6.0   Cloning and expression of Aβ(1-40) fragment as a fusion system

6.1   Background

Most biomolecular NMR techniques require isotopic labeling (^{2}H, ^{13}C, and/or ^{15}N) of recombinant proteins to alleviate the problems like spectral crowding and fast relaxation (*1*). Isotopic labeling can be achieved for short peptides using appropriate labeled precursors (*2*). But it is troublesome to synthesize macromolecules in this manner, due to the limitations in solid-phase peptide synthesis (*3*). Also, many of the NMR investigations require large quantities of isotopically labeled proteins, the production of which is often a costly and time-consuming aspect in NMR studies. Currently, most isotopically-labeled recombinant proteins are expressed in the bacterial host *E. coli*. Exploiting the glycolytic and citric acid cycle, selective amino acids can be labeled and have been proven useful for the structure determination of proteins (*4*). Recombinant DNA technology was used to produce the desired proteins in self-replicating living system. It can be done by incorporating plasmid DNA, encoding the genes of interest into the host system (*5*). Methods of generating heteronuclear-labeled samples in *E. coli* commonly use standard or modified versions of M9 minimal media employing ^{13}C glucose or ^{13}C glycerol as carbon source, ^{15}N ammonium sulfate or ^{15}N ammonium chloride as nitrogen source and deuterium oxide for deuteration (*5*). It is possible to enhance the solubility and survival of the protein of interest by combining the gene of interest, with the genes that encode various proteins (*6*). These fusion systems were initially used for the purification and handling of proteins of interest. Use of molecular chaperones as fusion partners, not only enhances the solubility of the protein, but also keeps the proteins at the proper fold (*6*). These fusion tags can be removed using proteases that can specifically recognize the engineered sequences that connect the two fusion partners (*7*).

6.2   Aim of the project

To clone and express the amyloid fragment as a fusion system

6.3   Material and Methods

DH5α (Novagen), BL21[DE3] (Novagen), were used as bacterial strains for the plasmid preparation and protein expression, respectively. Oligonucleotides were synthesized by MWG biotech GmbH (Ebersberg, Germany). Taq polymerase was purchased from Applied Biosystems. Restriction enzymes were purchased from New England Biolabs. pET-9 vector (Stratagene) was used for the preparation of the plasmid. pETM-41 (EMBL-Heidelberg) vector, which contains His tag (6 histidines) with maltose binding protein encoding gene,

kanamycin antibiotic resistance gene and tobacco etch virus (Tev) protease cleavage site, was used as a vector for the insertion of the amyloid gene. Tev cleavage site was incorporated in order to cleave the protein of interest from the fusion partner. Tev protease recognizes a seven amino acid consensus sequence, Glu-X-X-Tyr-X-Gln/Ser, where X can be varied amino acyl residues (*8*). Cleavage occurs between the conserved Gln and Ser residues. QIAEX II (Qiagen) gel extraction kits were used for the DNA fragment extraction from agarose gel. Plasmid midiprep kit (Qiagen) was used for the plasmid DNA isolation. Econo-columns(Biorad) filled with Ni-chelate agarose (Qiagen) were used for the fast purification of His-tag proteins. Biospin columns(Biorad) were also used instead of econo-columns to check the solubility of expressed proteins in a faster way. PD-10 columns (Amersham Biosciences) were used for the desalting of buffers-containing proteins.

6.4 Results and discussion

The nucleotide sequence of Aβ(1-40) comprising, 120 base pairs (bp), was designed (1-75 bp as sense strand and 69-120 bp as antisense strand) and the restriction sites NcoI or BbsI restriction site on 5' end of sense fragment and Acc65 restriction site of 3' end of antisense fragment were incorporated (NcoIf/Accrev or Bbsf/Accrev). 1-75 bp fragment without NcoI/BbsI was also synthesized to have control reaction (Amyf). Rare codons for *E.coli* were considered and replaced. The fragments were amplified by PCR with minimum cycles at 94-60-72 °C. The strategy is shown in the figure 6.1.
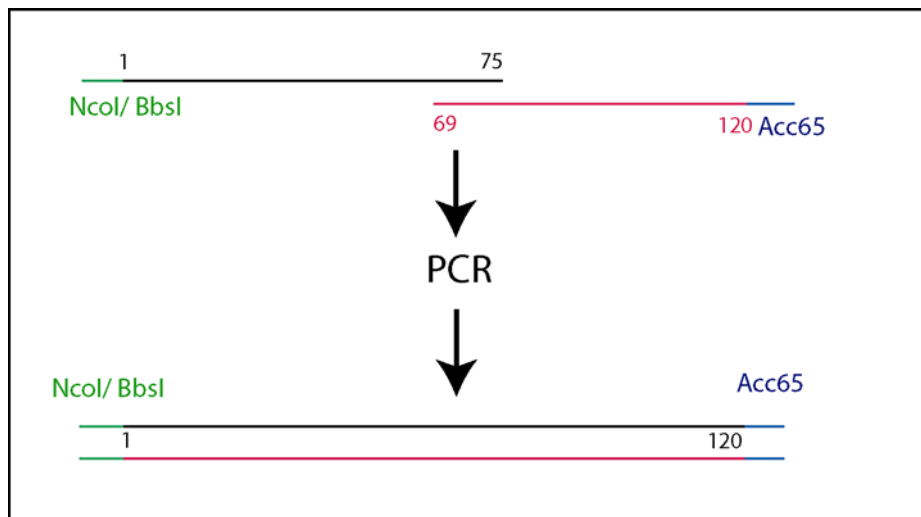
Figure 6.1 Synthesis and amplification of Aβ(1-40) nucleotide with restriction sites (NcoI/BbsI at 5' Acc65 at 3' of Aβ(1-40).

The PCR mix were prepared by the following formula: n x 30 μl for n PCR reactions

10 times buffer	:	n x 3 μl
dNTP mix (10mM)	:	n x 0.5 μl
strand –one (50 μM)	:	n x 1.5 μl
strand –two (50 μM)	:	n x 1.5 μl
Taq polymerase	:	n x 0.25 μl
Water (rest)	:	n x 23 μl

The amplification produced the Aβ(1-40) with overhangs at both the ends with corresponding restriction sites. Control experiment was performed by amplifying Aβ(1-40)-Acc65 fragments. After amplification, the DNA was analyzed by agarose gel electrophoresis. It was seen that all the reactions produced corresponding base pairs (NcoI/BbsI-Aβ(1-40)-Acc65) (Fig. 6.2). In the figure 6.2, NcoI/BbsI-Aβ(1-40)-Acc65 fragments show higher molecular weight corresponding to the Aβ(1-40)-Acc65 DNA fragments (lane 6 and 12).

Figure 6.2. PCR fragments showed correct fragments. lane 2-5 and 8-11 are overhanged PCR fragments. Lanes 6 and 12 are control DNA full length fragment without including NcoI or BbsI at 5' end of Aβ(1-40) nucleotide.

The corresponding PCR fragments were excised and the DNA was extracted from the agarose gel using standard extraction protocol (from Qiagen). The purified fragment was ligated into precut pET-9 (Acc65/NcoI) vector using standard protocol (*5*). After the ligation, the transformation was carried out in DH5α strain and plated out into the kanamycin plates and kept overnight at 37 °C. Large scale plasmid production and purification were carried out using standard procedure (from Qiagen). The purified plasmid were analyzed by PCR using plasmid specific primers (specific for T7 promotor site) and DNA specific primers (Fig. 6.3)
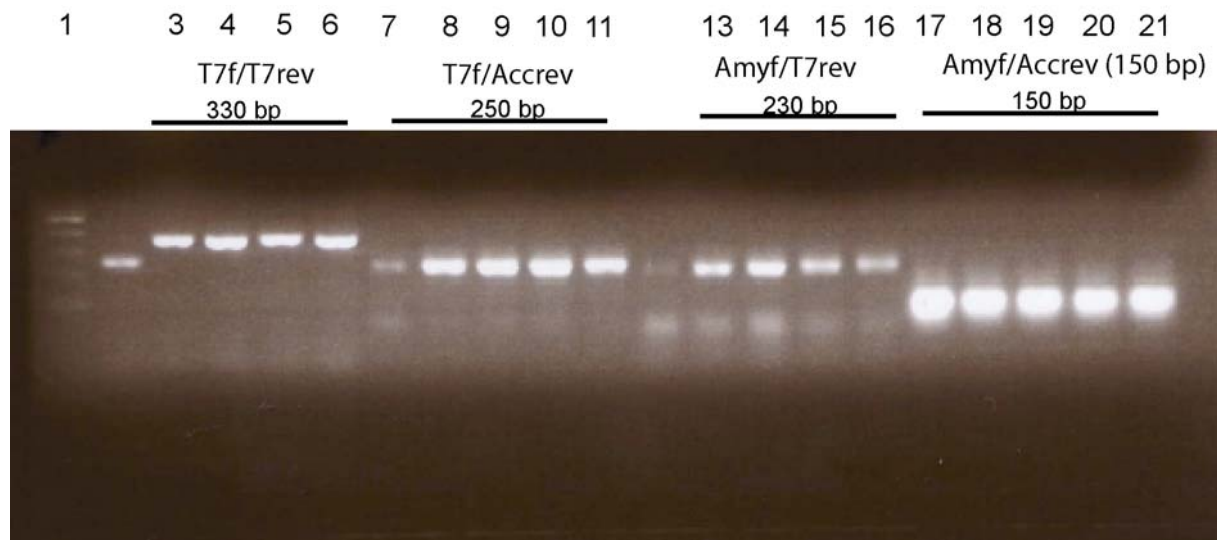
Figure 6.3. Agarose gel electrophoresis showed that the DNA fragment was inserted into the vector. PCR reactions were carried out using the combination of DNA specific and vector specific primers. Plasmid maxiprep of pET-9 vector was taken as a source DNA.

Subsequently, the plasmid was sequenced (SeQlab, Göttingen) and it proved that the vector was carrying the gene of interest. The gene was sub-cloned into the expression vector pETM-41 using appropriate restriction enzymes (Fig.6.4) and transformed into the expression strain BL21(DE3).
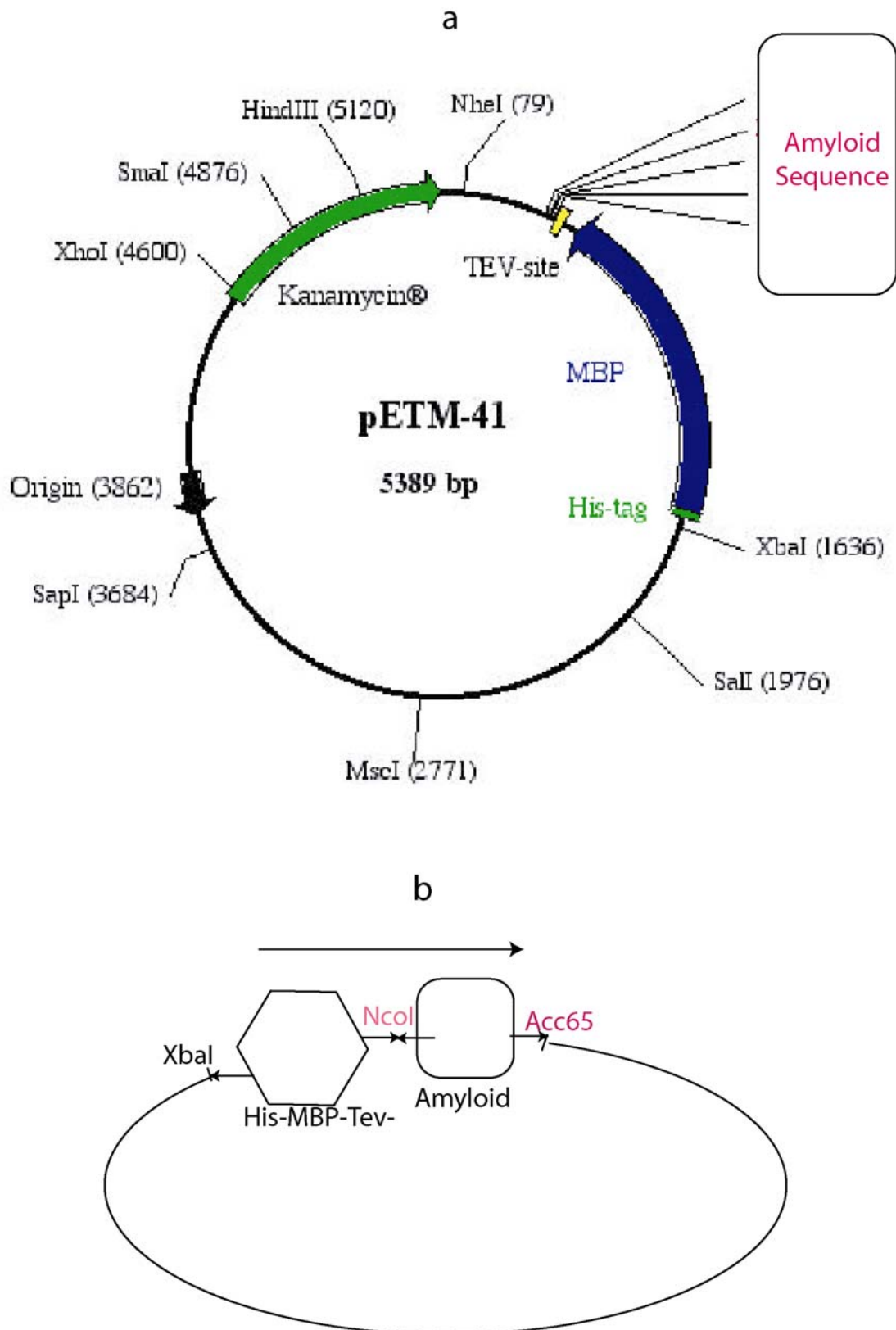
Figure 6.4. a) pETM-41 with the gene of interest; b) Restriction enzymes used for the ligation reaction for pETM-41 series vector.

After transformation the cells were plated out onto the kanamycin plates and incubated overnight at 37°C. The protein expression was carried out using the standard protocol (*5*). The cultured cells were incubated at 37 °C for 3 hours prior induction and overnight at 25 °C after induction in order to increase the solubility. The expression profile of the protein was monitored and the solubility screen was performed using standardized protocol (from Qiagen) (*9*). Rapid purification of His-MBP-Amyloid fusion protein was carried out using Ni-NTA agarose on econo-columns support. After SDS-PAGE analysis, expression of Aβ(1-40) fused with MBP was seen (~50 kDa) (Fig. 6.5). Work is in progress in our laboratory to purify the Aβ(1-40) fragment after digestion with Tev Protease.
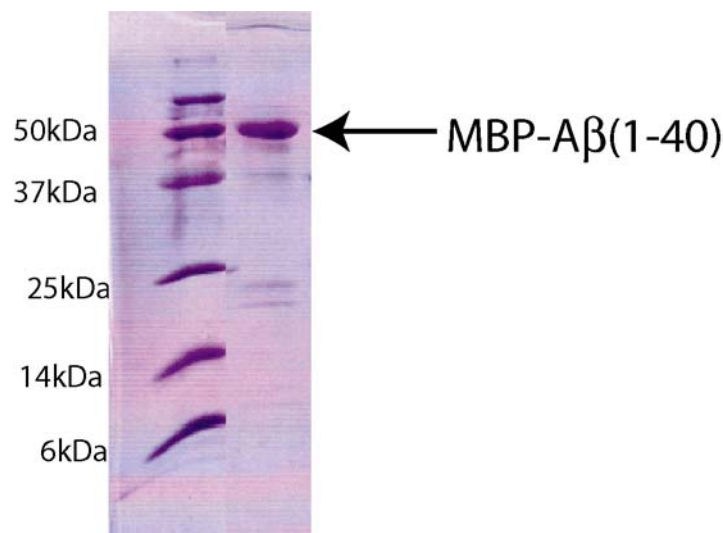
Figure 6.5. Expression profile of Aβ(1-40) fused with Maltose binding protein

6.5   References

1.   Gardner, K. H., Kay, L.E. (1998) *Annu.Rev.Biophys.Biomol.Struct. 27*, 357-406.
2.   Merrifield, B. (1997) *Methods Enzymol. 289*, 3-13.
3.   Wellings, D. A., Atherton, E. (1997) *Methods Enzymol. 289*, 44-67.
4.   Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., & Oschkinat, H. (2002) *Nature 420*, 98-102.
5.   Sambrook, J., Russell, D.W. (2001) *Molecular cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press.
6.   Baneyx, F. (1999) *Curr.Opin.Biotech. 10*, 411-421.
7.   Jenny, R. J., Mann,K.G., Lundblad, R.L. (2003) *Protein Expr.Purif. 31*, 1-11.
8.   Parks, T. D., Leuther, K.K., Howard, E.D., Johnston,S.A., Dougherty, W.G. (1994) *Anal.Biochem. 216*, 413-417.
9.   Stier, G. (2002) *EMBO practical course on protein expression, purification and crystallization*, EMBL, Hamburg outstation, 2002.

7.0   Stereo-selectivity of Hsp104 as observed in ^{19}F NMR

7.1   Background

Biological systems continue to fascinate the spectators as they show tremendous discipline in exerting their functions. To address the complexity in biological systems, the knowledge of structural, sequence properties, evolutionary history, cell type specificity, abundance, protein localization and post-translational modification are important. All these factors contribute to functions. These biological functions often occur via highly-specific multi-component assembly. These occur due to the interactions of proteins with other proteins or nucleic acids. To study the importance of protein-protein interactions, various techniques from molecular biology (*1*), NMR techniques (*2, 3*) to computational methods(*4*) have been employed (reviewed in chapter 1). *In vivo* methods for identifying and characterizing biomolecular interactions have greatly improved over the past few years. To detect and identify the physiologically relevant protein-protein interactions, traditional methods like co-immunoprecipitation and affinity chromatography are used *in vivo* (*5*). Yeast and bacterial hybrid systems (*6*) and protein micro array techniques (*7*) provide excellent methods for the modern post-genomic era where large number of libraries of proteins need to be probed. But, none of the above methods would specify site-specific information concerning the interactions of protein-protein interactions. Site-specific information are necessary to reveal the segment of proteins involved in signal transduction and disease.

^{19}F-Nuclear magnetic resonance spectroscopy has proved to be a powerful technique in the study of protein-protein interactions and dynamics in site-specific manner. Where high-resolution NMR experiments cannot be useful, ^{19}F NMR generates valuable information. In past, it has been used extensively in structural analysis of membrane proteins (*8*) and protein-protein interactions (*9*). There are several factors that contribute to the study of protein-protein interactions by ^{19}F NMR (*10*).

1. ^{19}F nucleus occurs at 100% natural abundance and has 83% the sensitivity of ^{1}H.

2.   As ^{19}F does not occur naturally in proteins, no background signals are observed.

3. ^{19}F chemical shift range is 100-fold larger than that of ^{1}H. In addition, it is influenced by lone pairs, which provide large paramagnetic term in shielding constant

4.   The chemical shift is therefore sensitive to changes in the local van der Waals environment, as well as local electrostatic fields and hence making the technique highly sensitive for the detection of protein conformational changes and interactions with other partners.

5. Incorporating fluorine into the protein can be easily performed by growing the cells in enriched fluoro-amino acids (*11*)

When biosynthetic incorporation of fluorinated amino acids is impractical due to the cell toxicity, fluorination of specific protein side chains using thiol modifying agents can be used. There are several thiol modifying agents available commercially. The thiol-reactive functional groups are primarily alkylating reagents, including iodoacetamides, maleimides, benzylic halides and bromomethylketones. Selectivity, specificity and stability of modified thiol group should be taken into account while choosing the modifying agents. Maleimides are excellent reagents for thiol-selective modification, quantification and analysis (*12*). The reaction involves addition of the thiol across the double bond of the maleimide to yield a thioether (Fig.7.1). The optimum pH for the maleimide to react with thiol is 7.0.
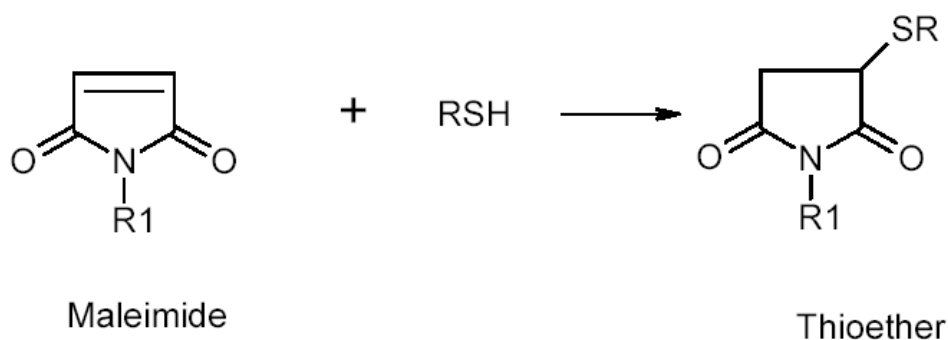
Figure 7.1. Reaction scheme of maleimide with thiols

Quantification of thiol modification can be done by simple Ellman's test. It uses 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) as a reagent and the level of modification can be obtained by UV spectroscopic measurements (*13*).

7.2 Aims of the project

1. To synthesize N-2,2,2 Trifluoro N-Methyl maleimide (FNEM) as a cysteine modifier for ^{19}F NMR studies on protein-protein interactions and to standardize the reaction conditions
2. To modify the Hsp104 using the FNEM reagent

7.3   Synthetic Strategy
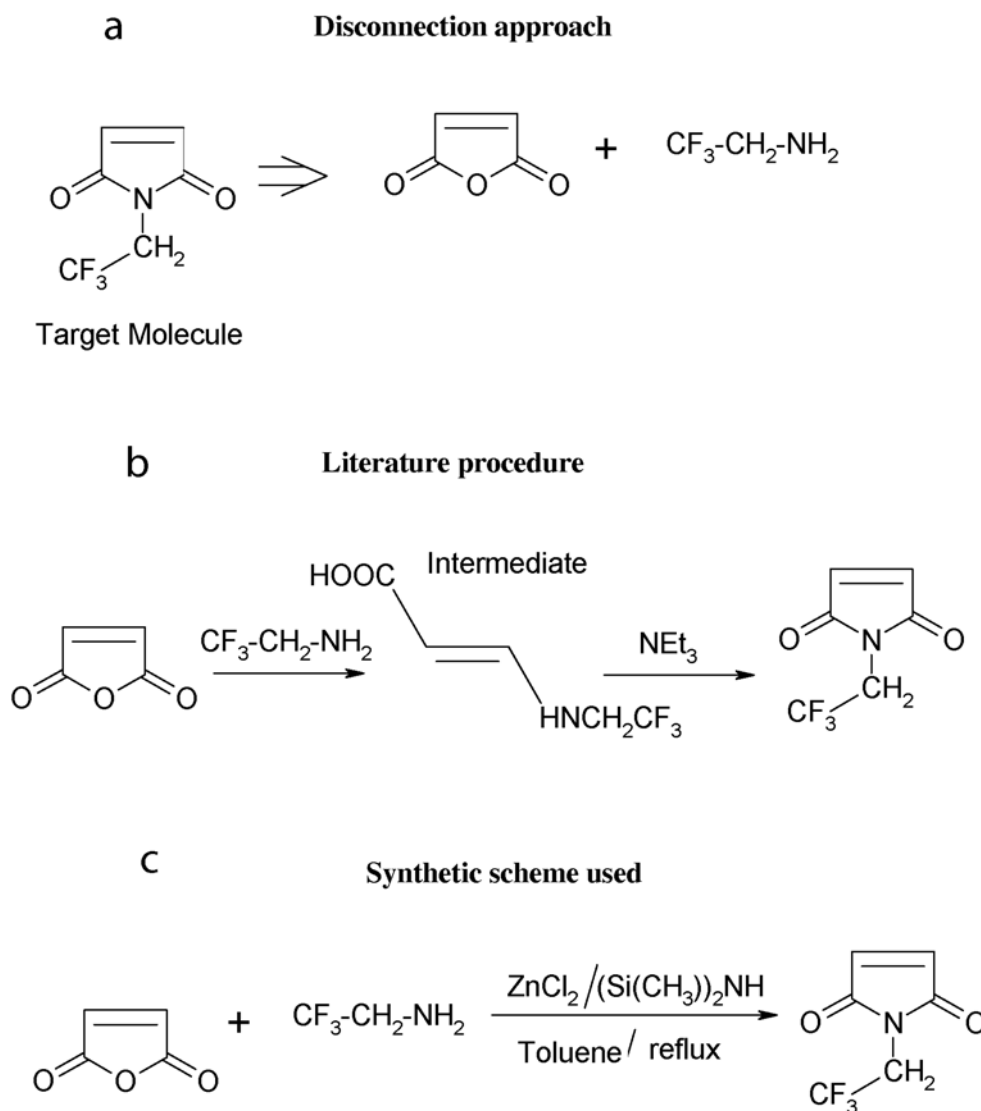
Figure 7.2. a) Maleic anhydride and corresponding amines are used up as starting materials; b) Reaction mechanism involves nucleophilic substitution and followed by intramolecular cyclization; c) For fluoro-amines as nucleophiles, efficient organo-silicon chemistry activated by Lewis acid was used

The synthetic strategy is depicted in the Figure 7.2. As fluoroamines are weak nucleophiles, an efficient synthetic protocol was used. Using 1,1,1,3,3,3-hexamethyldisilazane (HMDS) in presence of a Lewis acid, ZnCl$_2$, the target molecule was synthesized from its corresponding precursors (Fig. 7.2c). This method has been used previously to synthesize N-aryl amides that also contain similar electron withdrawing groups next to the nucleophiles (*14*).

7.4   Material and Methods

All organic solvents were distilled before use. RP-HPLC analysis and semi-scale preparations were carried out on a Waters instrument. NMR experiments were performed on a 600.4 MHz Bruker spectrometer. N-2,2,2-trifluoroethylamine was purchased from ABCR chemicals, Karlshruhe, Germany. HMDS, Maleic anhydride and Zinc(II) chloride were purchased from Sigma-Aldrich. Both benzene and toluene were used as solvents. Hsp104 was purified as described in Chapter 3.

In a round-bottomed reaction flask, Maleic anhydride (0.3 gm, 3.2 mmol) was taken together with N-2,2,2, trifluoro-ethylamine (0.204 ml, 2.58 mmol). A volume of 5 ml of dry toluene was added into the mixture and stirred at room temperature. In order to control the moisture during the mixing, a stream of dry-nitrogen was passed into the reaction vessel. Zinc chloride (0.4gm, 3.2mmol) was added after an hour of mixing the corresponding anhydride and the amine. The dry-nitrogen air inlet was replaced by a calcium chloride guard tube and the whole set-up was mounted into an electrical mantle to heat up for an hour at 85 °C. Afterwards, the reaction mixture was cooled down to 60 °C and solution of HMDS (0.82 ml, 3.8 mmol) in dry toluene was added drop-wise over 30 min. The reaction mixture was refluxed for additional 2 hours. The reaction vessel was cooled down to room temperature and 0.5N HCl (30 ml) was poured in. The aqueous phase was extracted with ethyl acetate (4x25 ml). The combined extracts were washed successively with saturated sodium bicarbonate (2x50ml), sodium chloride (2x50 ml) and finally with water (1x50 ml). The resulting solution was dried over anhydrous sodium sulphate for 12 hours. Afterwards, the solution was filtered and ethyl acetate was pumped off and yielded white powder. Several times, the final compound, FNEM, crystallized out as colourless glassy needles from the reaction vessel after cooling.

HPLC purification was carried out (acetonitrile / water) using analytical reverse phase column and the purified product came out at 16 ml elution volume (5-80% of acetonitrile). ESI-MS correspond to m/z =179 [M+H$^+$], matched exactly the formula weight of the target molecule (Mass = 178). NMR confirmed high level of purification (>98%). [NMR (7.3 ppm (one proton, Singlet, H-double bond), 6.8 ppm (one proton, Singlet, H-double bond), 4.4 ppm, (2 protons, quartet, $^3J_{HF}$=8.59 Hz, CH$_2$ in –N-CH$_2$-CF$_3$)].

In order to modify the thiol containing proteins or peptides, the purified FNEM, was incubated with the thiols for 6 hours at 4°C (ratio 1:2 of thiols: FNEM, pH 7.0 buffer). After the modification, the free ligand was either dialysed out or purified over column depending upon the size of the protein of interest to be modified.

7.5   Results and discussion

7.5.1   FNEM produced stereo-isomeric compounds upon modification

In order to test the efficiency of labelling, a thiol containing short peptide, reduced-glutathione (GSH) was used as a test system. After modification, the sample was subjected to HPLC purification and two fractions were collected with the peak intensity of 50% each. Mass spectroscopy showed that both samples were genuinely labelled with the FNEM (m/z=487 [M+H$^+$]). This can be explained as stereo-isomeric products (Fig. 7.3).
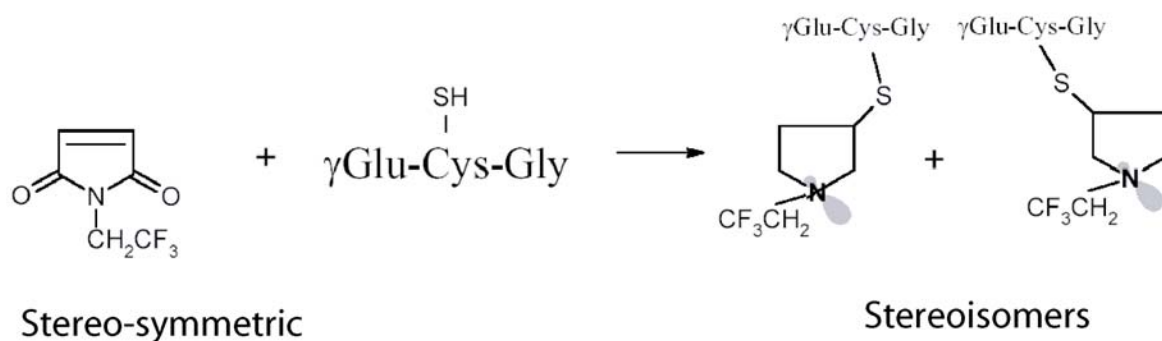
Figure 7.3. Stereo-isomers of FNEM modified Glutathione(reduced)

In order to confirm the stereo-isomers, ^{19}F NMR experiments were carried out. Both the fractions were showed different ^{19}F chemical shifts, as expected (Fig. 7.4). This confirmed that the modification of thiols using FNEM render two different steroisomeric products. This finding is very interesting as it can be used to probe the specificity of proteins containing thiol groups. Existing cysteine-modifying agents are mainly chromophores or fluorescence probes and hence the isomers cannot be observed. Only NMR can be used for the purpose.
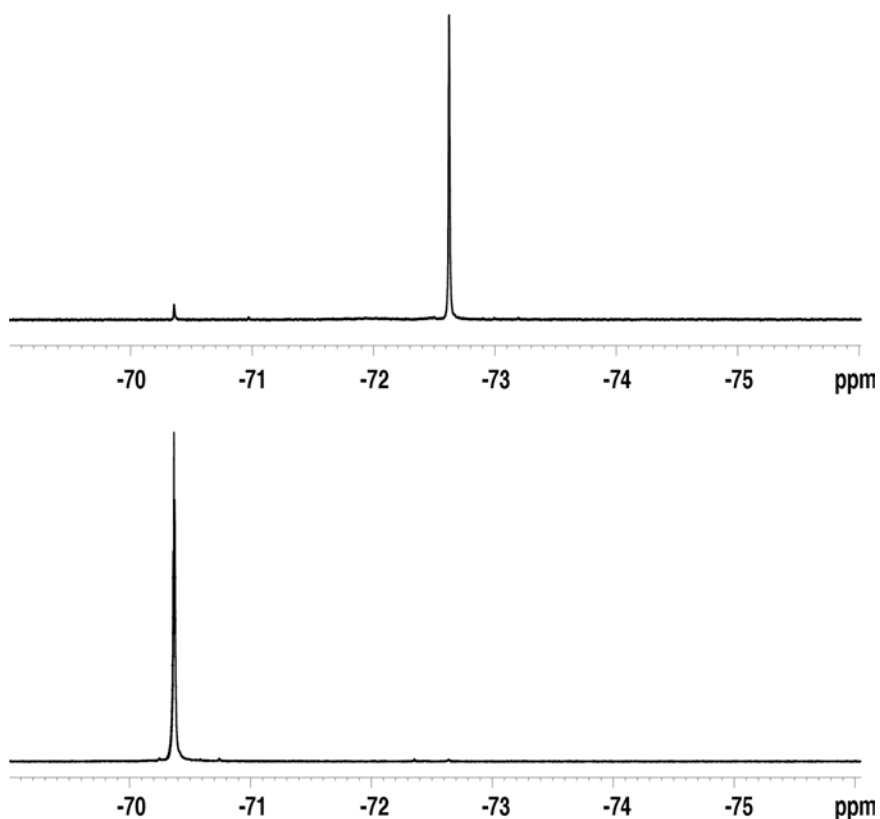
Figure 7.4. Decoupled ^{19}F-NMR spectrum of conformer 1 (top) and conformer 2 (bottom) showed different chemical shifts

7.5.2   Hsp104 modification showed stereo-specificity towards FNEM reagent

Hsp104 contains six cysteines and hence modification of Hsp104 was performed as described in Materials and Methods. After the modification, the Ellman's test suggested the modification of four out of six cysteines in Hsp104. It implies that two other cysteines were non-accessible for the FNEM.

After the modification, the Hsp104 activity was measured according to the colorimetric procedure as described in the literature (*15*) . The results suggest that the modification of cysteines does not impair the activity of Hsp104, as the activity of modified Hsp104 was showed similar profile to the unmodified one (Fig. 7.5).
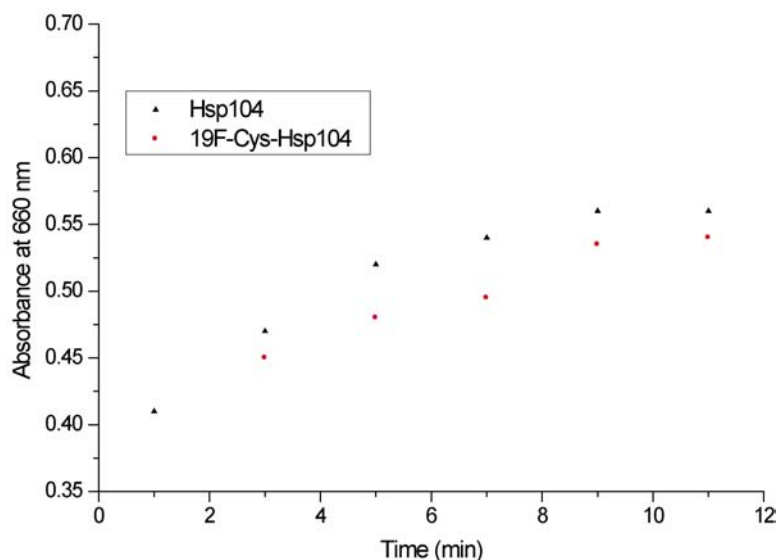
Figure 7.5. ATPase activity of Hsp104 was found to be unaffected after modification

^{19}F-NMR measurements were carried out using the modified Hsp104. The results show that Hsp104 populated only one of the two available stereo-conformations as judged from the ^{19}F-NMR chemical shifts (Fig. 7.6)
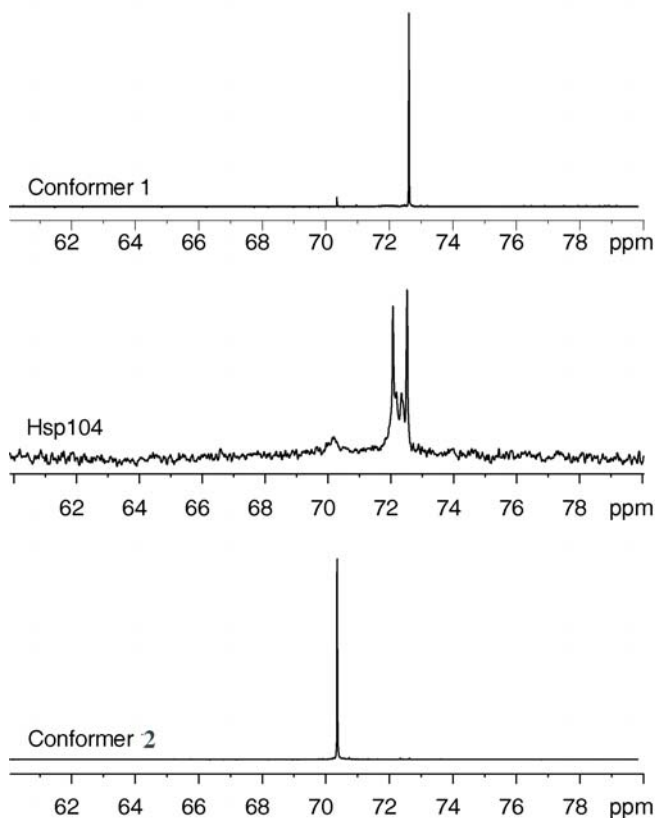
Figure 7.6. 19F NMR showed that Hsp104 preferentially adopted one of the two possible conformation while reacting with the Cysteine modifying agent

It is an interesting result as one would expect an equal population of both the stero-isomeric products. But, it is clear that the Hsp104 shows preference over one conformation to the another. It can be viewed as stereo-selectivity of Hsp104 towards the modifier. It is important to know that, compared to other substrates that Hsp104 encounter, (Ref. Chapter 3) the FNEM is much smaller, but still it showed high degree of specificity. This specificity may be used by the Hsp104 upon substrate binding and activity. Hence, substrate specificity of Hsp104 towards selective species can be possible as described in Chapter 3. Hence, we conclude that, in addition to sequence specific binding, chaperones could also use their stereochemistry to locate certain symmetry-related species.

7.6   Conclusion

In future, FNEM reagent can be very useful in probing protein-protein interactions and enzymatic activity. For the molecular chaperone, Hsp104, this strategy can be useful to study the substrate binding sites in proteins and domain communications.

7.7   References

1.   Fields, S., Song, O. (1989) *Nature 340*, 243-6.
2.   Ludwiczek, M. L., Baminger, B., Konrat, R. (2004) *J.Am.Chem.Soc. 126*, 1636-7.
3.   Zuiderweg, E. R. (2002) *Biochemistry 41*, 1-7.
4.   Marcotte, E. M., Pellegrini, M., Ng, H., Rice,D.W., Yeates, T.O., Eisenberg, D. (1999) *Science 285*, 751-753.
5.   Sambrook, J. a. R., D.W., Sambrook, J. (2002) *Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory; 3rd edition*.
6.   Joung, J. K., Ramm, E.L, Pabo, C.O. (2000) *Proc Natl Acad Sci U S A. 97*, 7382-7.
7.   MacBeath, G. a. S., S.L. (2000) *Science 289*, 1760-3.
8.   Loewen, M. C., Klein-Seetharaman, J., Getmanova, E.V., Reeves, P.J., Schwalbe, H., Khorana, H.G. (2001) *Proc Natl Acad Sci U S A. 98*, 4888-92.
9.   Bann, J. G., Pinkner, J., Hultgren, S.J., Frieden, C. (2002) *Proc Natl Acad Sci U S A. 99*, 709-14.
10.   Gerig, J. T. (1994) *Progress in NMR spectroscopy 26*, 293-370.
11.   Kim, H.-W., Perez, J.A., Ferguson, S.J., Campbell, I.D. (1990) *FEBS Lett. 272*.
12.   Palmer, M., Buchkremer, M., Valeva, A., Bhakdi, S. (1997) *Anal.Biochem. 253*, 175-179.
13.   Ackerman, R. J., Robyt, J.F. (1972) *Anal.Biochem. 50*, 656-659.
14.   Reddy, P. Y., Kondo, S., Toru, T., Ueno, Y. (1997) *J.Org.Chem. 62*, 2652-2654.
15.   Banecki, B., Wawrzynow, A., Puzewicz, J., Georgopoulos, C., Zylicz, M. (2001) *J.Biol.chem. 276*, 18843-18848.

8.0   Factors behind amyloidogensis and the rationale behind chaperones' action

8.1   What drives the proteins to form ordered aggregates ?

From the results, it can be deduced that ordered aggregates can occur due to the interactions of hydrophilic, as in case of Sup35[5-26] (Chapter 3) or hydrophobic, as in the case of Aβ (1-40) (Chapter 4) stretchs of amino acids. In the former case, thermal energy causes prevention of fibril formation and the later case, the enhancement of fibril formation was seen upon heating. This is due to the electrostatic mediated association in the case of Sup35[5-26]. But, through Van der waals interactions in the case of Aβ(1-40). Taking the globular protein PI3-SH3 as an example, which formed fibril due to the charged residue at a stretch of amino acids (Chapter 5). All the above examples can be converged in to a single term called net charge. It was also shown that fibril formation of Aβ(1-40) could be modulated by the anions. These anions are bound to the positively charged loop region in Aβ(1-40) and showed different oligomeric distribution (Chapter 4). Hence, these charges should play a crucial role in the fibril formation, though the hydrophobic interactions may govern the acceleration of fibril formation in the case of Aβ(1-40). It has been widely accepted that protein-protein association occurs after collisions over a long time whenever protein molecules are encountered in the solution (*1*). Protein association is facilitated, if the contact points consist of some charged residues ("hot spots"), which would be an equivalent of partially formed bonds and allow time for the protein to reorganize during fibril assembly. There are some clearance pathways that avoid protein aggregation. Protein aggregation occur via nucleation and the formation of nucleus is concentration-dependent. Thus, the effective clearance pathway would keep the protein out of the aggregation pathway by keeping the concentration of the nucleus to be negligible.

8.2   How do chaperones act?

The results concluded that chaperones use specificity towards certain oligomers and pull them out of the aggregation pathway (Chapter 3). This catalytic cycle would exist until it lowers the fibril forming nucleus concentration, a critical concentration above which the fibril formation occurs. The specificity of molecular chaperones towards these proteins is one of the prime factors in the modulation of aggregation pathway. It is intriguing, what drives the chaperones to recognize these species. Previously, it has been shown that molecular chaperones use their symmetry (*2*), sequences (*3*) and hydrophobicity (*4*) to specifically interact with their targets. Can catalytic redox cycle of chaperones influence the aggregation profile? There was an indication that chaperones could also use their redox activity to recognize the substrates (*5*).

From the obtained results found in this thesis (Chapter 4), it was clear that the heat shock protein αB-crystallin showed redox activity and was modulated by substrate binding. Moreover, the similar behavior was seen in the case of Hsp104 (Fig. 8.1), though more experimental evidences are needed to confirm these finding
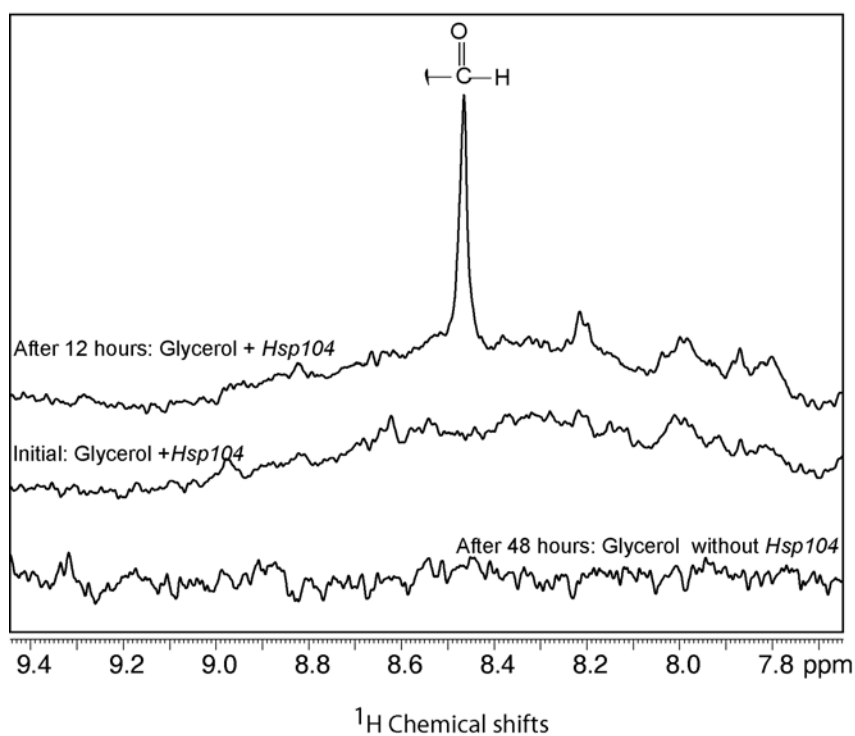
Figure 8.1. Redox activity shown by Hsp104. Hsp104 oxidized glycerol into glyceraldehyde.

Biologically, these redox activity possessed by molecular chaperones could be used for the modulation of net charge in the sequences that are crucial for fibril formation and hence could prevent aggregation. It can also be deduced that the chaperones' action *in vivo* are usually coupled to the various redox proteins (PDI in case of calnexin cycle). Hence, this redox property could be used to recognize selectively misfolded proteins using redox activity as a tag of recognition. It still puzzles how the molecular chaperones and proteosomes work together. Molecular chaperones involve in refolding pathway where as proteosomes work on degradation pathway, together they make an effective quality control machinery. The proteins that are unable to refold are directed to the proteosomes with prefixed ubiqutin tag. The attachment of ubiquitination to the proteins may consider to increase the redox load of the misfolded protein that can only be recognized by the proteosomes. By this way the chaperones and proteosomes may work together. This postulation may be proved or disproved in future. Surely, more data are in need in order to do so.

8.3   References

1. Northrup, S., H., Erickson, H.P. (1992) *Proc. Natl. Acad. Sci. USA 89*, 3338-3342.
2. Akey, C. W., Luger, K. (2003) *Curr.Opin.Struct.Biol 13*, 6-14.
3. Tasab, M., Jenkinson, L., Bulleid, N.J. (2002) *J.Biol.Chem. 277.*
4. Rudiger, S., Mayer, M.P., Schneider-Mergener, J., Bukau, B. (2000) *J.Mol. Biol. 304*, 245-251.
5. Hoffmann, J. H., Linke, K., Graf, P.C.F., Lilie, H., Jakob, U. (2004) *EMBO J 23*, 160-168.

9.0   Appendix: C source codes

-- ----------

Programme #1   Constructs 2D NM R (2rr) files from 1D (1r) files

----

/* Written by Saravanakumar Narayanan, AG Bernd Reif and AK Horst Kessler */

/*Usage after compilation : a.out (time domain) <outputfile name>

/* It creates Ser 2D file by appending several 1D files. Keep all 1r files in a single folder*/

```c
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <math.h>

#define MAXVEC 32768

FILE *fout, *fin ;
int i, a, count, k, td;
int  re1[MAXVEC];
char ser[80],serin[80],sout[80], s2[80], istring[10];
char filename[50],ans;
main(argc,argv)
int argc;
char *argv[];

{

 if (argc != 3)
   { printf("\n pattex: wrong number of arguments\n\n");
    printf("USAGE: a.out <TD> <output file name:ser>\n");
    printf("dummy2 assumes files fid1, fid2, ..., fidn in the current directory\n\n");
    printf("Works for fids up to 8k points\n\n");
    exit(1);
   }

   sscanf(argv[1],"%d",&td);
   if ((fout=fopen(argv[2],"wb"))==NULL)
   { printf("ERROR opening file %s !\n",argv[2]); exit(1); }
  printf("\n do you want to convert a fid file in to a ser file? Y/N\n\n");
  fflush(stdin);
  scanf("%c", &ans);
  while((ans=='Y' || ans=='y'))
{    fseek(fout,0,SEEK_END);
   printf("\n Please append a fid file to include:    ");
   fflush(stdin);
   scanf("%s", filename);
   if ((fin=fopen(filename,"rb"))!=NULL)
   printf("| current fid: %10s |\n",filename);
```

171

```
    fread(&re1[0],4,td,fin);
    fwrite(&re1[0],4,td,fout);
    fclose(fin);
    printf("\nAdd another fid? (Y/N)\n\n");
    fflush(stdin);
    scanf("%c", &ans);

}
{    fclose(fout);
   printf("\n ==> FINISHED !\n");
}
}
```

-- -----------

Programme #2   Plots Time Vs concentration plot
                                Reads out from 2rr NMR file

/* Written by Saravanakumar Narayanan, AG Bernd Reif and AK Horst Kessler */

/*Usage after compilation : a.out <file name>

Input file contains

Number of peaks, Number of points in the indirect dimension, <peak pick file> ,<time file> ,

size in direct dimension, <2rr_file after usage from convert2rr.c source code> , <output file> ,

noise column at the spectra.

<peak pick file> contains

name of the peak, column of the peak in 2rr file, {(column)$_{peakend}$-(column)$_{begin}$} */

<time file> contains
time in linear order

/* concentrationdiff.c */
/* plotting time vs concentration*/

#include <stdio.h>
#include <string.h>
#define MAXVEC 32768

FILE *fout1, *fout2, *fout3, *fin, *fin2, *fin1pp, *fin2temp, *fin3datatemp ;
int l, i, a, cnttemp, cntgrad, j, cntpeak, k, si2, peakpnt, peakpnt2, noise;
int re1[MAXVEC], ns1, ns2, s2, difftime1, max;
float slope1,conc1, conc2, slope, refnoise,temp, refnoise0, refnoise3,refnoise2,
refnoise1,refint, refint6, refint0, refint2, refint1, concentration ;
float timeref0,timeref,timeref1,timeref2;
float intref,intref0,intref1,intref2;
char peaknam[140],data_pp[140],data _temp[140],data _2rrtemp[140];
char data_out[140],s1[140], s3[140];

172

```
main(argc,argv)
int argc;
char *argv[];


{

if (argc != 2)
   { printf("\n concentratindiff: wrong number of arguments\n\n");
    printf("USAGE: concentrationdiff <concmap.in>\n");
    printf("Synthax <concmap.in>: \n");
    printf("<number_of_peaks>   <number_of_temppoints> \n");
    printf(" <peak_pick_file>  <temp_file>  <size> <2rr_tem p> <dataout> <noise> \n");
    exit(1);
   }


 if ((fin=fopen(argv[1],"r"))==NULL)
    { printf("ERROR opening file %s !\n",argv[1]); exit(1); }


fscanf(fin,"%d",&cntpeak);
fscanf(fin,"%d",&cnttemp);
fscanf(fin,"%s",data_pp);
fscanf(fin,"%s",data_temp);
fscanf(fin,"%d",&si2);
fscanf(fin,"%s",data_2rrtemp);
fscanf(fin,"%s",data_out);
fscanf(fin,"%d",&noise);


if ((fin1pp=fopen(data_pp, "r"))==NULL)
   { printf("ERROR opening file %s !\n",data_pp); exit(1); }

if ((fin2temp=fopen(data_temp,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_temp); exit(1); }


if ((fin3datatemp=fopen(data_2rrtemp,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_2rrte mp); exit(1); }


   fseek(fin1pp,0,0);
for (k=1; k < cntpeak+1; k++){
   fseek(fin3datatemp,0,0);
   fscanf(fin1pp,"%s %d %d",peaknam, &peakpnt, &max);
    printf("i am here baby %s %d !\n",peaknam, peakpnt );

   strcpy(s1,data_out);
   strcat(s1,"_");
   strcat(s1,peaknam);
```

```
   strcat(s1,"_");
   strcat(s1,"temp");
   fout1=fopen(s1,"w");

fprintf(fout1,"#temp   relative concentration\n");


   /* Create Output File for REF at specific Peak Position */

   fseek(fin3datatemp,0,0);
   fseek(fin2temp,0,0);

   for (i=1; i < 2; i++){
      refint=0;
      refnoise0=0;
      fread(&re1[0],4,si2,fin3datatemp);
      for (j=1;j < 51; j++){
       refnoise = re1[11500+j];
       if (refnoise<0) {
             refnoise=-refnoise;
               }
       refnoise0=refnoise+refnoise0;
          }
       refnoise2=refnoise0/50;

      for (j=1;j < max;j++){
       refint0 = re1[peakpnt+j];
          if (refint0<-refnoise2) {
             refint0=-refnoise2;
                  }
          refint0=refint0+refnoise2;
          refint=refint+refint0;
            }
            }
       fseek(fin3datatemp,0,0);
      if (refint==0){
      for (l=1; l < cnttemp+1; l++){
         refint6=0;
          fread(&re1[0],4,si2,fin3datatemp);
          for (j=1;j < max; j++){
           refint0 = re1[peakpnt+j];
           if (refint0<-refnoise2) {
              refint0=-refnoise2;
                   }
          refint0=refint0+refnoise2;
          refint6=refint6+refint0;
            }
        if  (refint<refint6) {
          refint= refint6;
              }
          }
```

174

```c
        printf("i am here baby %f %f !\n",refint0, refint );
        refint0=0;
        }
        fseek(fin2temp,0,0);
        fseek(fin3datatemp,0,0);

    for (i=1; i < cnttemp+1; i++){
        refnoise1=0;
        refint2 =0;
        fread(&re1[0],4,si2,fin3datatemp);
        fscanf(fin2temp,"%f  %s ",&temp, s3);
        for (j=1;j < 51; j++){
        refnoise = re1[11500+j];
        if (refnoise<0) {
                refnoise=-refnoise;
                }
        refnoise1=refnoise+refnoise1;
                }
         refnoise3=refnoise1/50;
        for (j=1;j < max;j++){
        refint1 = re1[peakpnt+j];
            if (refint1<-refnoise3) {
                refint1=-refnoise3;
                    }
            refint1=refint1+refnoise3;
            refint2=refint1+refint2;
                }
        concentration = refint2/refint;
        printf("i am here baby  %f %f %f!\n",refint1, refint2 , concentration);
        fprintf(fout1," %f %f \n", temp, concentration);
      printf("i = %2d, temp = %7.1f, concentration = %7.3f, refint0 = %7.1f,  \n", i, temp,
concentration, refint0);

    }
 fclose(fout1);
}
 fclose(fin3datatemp);
 fclose(fin2temp);
 fclose(fin1pp);
printf("\n ==> FINISHED !\n");
}
```

-- -----------
Programme #3   Converts Bruker format to binary format
                                Reads out from 2rr NMR file

```c
/*  USAGE: convert2rr  (size indirect dimension) (size direct dimension) <2rrprocessed file>
<2rr resorted output name>*/

/* AG BERND REIF */
/* source code from BERND */

#include <stdio.h>
#include <string.h>

#define MAXVEC 131072

FILE *fout, *fin, *fout1, *fout2, *fout3, *fout4;
FILE *dum1, *dum2, *dum3, *dum4;
int test, i, j, a, count, k, si1, si2, td, n[4096];
int intens;
int  re1[MAXVEC];
char ser[80],serin[80],sout[80], s2[80], istring[10];

main(argc,argv)
int argc;
char *argv[];


{

 if (argc != 5)
   { printf("\n convert2rr: wrong number of arguments\n\n");
    printf("USAGE: convert2rr si1 si2 <binary 2rr> <resorted binary 2rr>\n");
    exit(1);
   }

 sscanf(argv[1],"%d",&si1);
 sscanf(argv[2],"%d",&si2);

 if ((fin=fopen(argv[3],"r"))==NULL)
   { printf("ERROR opening file %s !\n",argv[3]); exit(1); }

 if ((fout=fopen(argv[4],"w"))==NULL)
   { printf("ERROR opening file %s !\n",argv[4]); exit(1); }


 if ((fout1=fopen("dum1","w"))==NULL)
   { printf("ERROR opening file %s !\n","dum1"); exit(1); }

 if ((fout2=fopen("dum2","w"))==NULL)
   { printf("ERROR opening file %s !\n","dum2"); exit(1); }

 if ((fout3=fopen("dum3","w"))==NULL)
```

```
    { printf("ERROR opening file %s !\n","dum3"); exit(1); }

if ((fout4=fopen("dum4","w"))==NULL)
    { printf("ERROR opening file %s !\n","dum4"); exit(1); }


for(i=1; i<(0.25*si1+1); i++) {
    for(test=1; test<9; test++){

        fread(&re1[0],4,4096,fin);
        fwrite(&re1[0],4,4096,fout1);

        fread(&re1[0],4,4096,fin);
        fwrite(&re1[0],4,4096,fout2);

        fread(&re1[0],4,4096,fin);
        fwrite(&re1[0],4,4096,fout3);

        fread(&re1[0],4,4096,fin);
        fwrite(&re1[0],4,4096,fout4);

    }
  }

  fclose(fin);
  fclose(fout1);
  fclose(fout2);
  fclose(fout3);
  fclose(fout4);
  fout1=fopen("dum1","r");
  fout2=fopen("dum2","r");
  fout3=fopen("dum3","r");
  fout4=fopen("dum4","r");


  for(k=1; k<(0.25*si1 +1); k++) {

      fread(&re1[0],4,si2,fout1);
      fwrite(&re1[0],4,si2,fout);

      fread(&re1[0],4,si2,fout2);
      fwrite(&re1[0],4,si2,fout);

      fread(&re1[0],4,si2,fout3);
      fwrite(&re1[0],4,si2,fout);

      fread(&re1[0],4,si2,fout4);
      fwrite(&re1[0],4,si2,fout);

  }
```

```
    fclose(fout1);
    fclose(fout2);
    fclose(fout3);
    fclose(fout4);
    fclose(fout);


}


------------------------------------------------------------------------------------------------------ --------------
Programme #4        construct time file in mi nutes  for the entire 1D kinetics
------------------------------------------------------------------------------------------------------------------
/*  Written by Saravanakumar Narayanan, AG Bernd Reif and AK Horst Kessler */

/*  Usage after compilation: a.out   <timeconversion.input>
<timeconversion.input> contains   number of files, number of points, datafiletoindicate file
name
simple usage: go to unix shell, go to the directory of acquisition: type grep PULPROG */fid
get fid and its date and time it records, copy everything and paste into a file. Instead of
subtracting one by one use this programme to extract the time in min. after starting the
experiment*/


/* timeconversion.c */


#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <math.h>

FILE *fout1,  *fin0,*fin1, *fin1data ;
int    a, j,no2, cntpnts, cntfile, k,no5 ;
float date, minutes, minutes1;
double time, frac, intr;
char peaknam[140],data_file[140],dat a_file[140],data _2rrtemp[140];
char data_out[140],s1[140] , filename[140], s2[140];
char no7[140],no1[140], no3[140], no4[140],  no6[140],mon[140];
main(argc,argv)
int argc;
char *argv[];

{

 if (argc != 2)
   { printf("\n std_from_2rr_fi  nal: wrong number of arguments\n\n");
    printf("USAGE: timec onversion <timeconversion.in>\n");
    printf("Synthax <timeconversion.in>: \n");
    printf("<number_of_files  >  <number_of_points> \n");
    printf(" <data_file_t  o_indicate_the_file_name> \n");
    exit(1);
```

178

```c
    }


  if ((fin1=fopen(argv[1],"r"))==NULL)
     { printf("ERROR opening file %s !\n",argv[1]); exit(1); }



 fscanf(fin1,"%d",&cntfile);
 fscanf(fin1,"%d",&cntpnts);
 fscanf(fin1,"%s",data_file);



 if ((fin1data=fopen(data_file, "r"))==NULL)
    { printf("ERROR opening file %s !\n",data_file); exit(1); }



    fseek(fin1data,0,0);

for (k=1; k < cntfile+1; k++){
    fscanf(fin1data,"%s",filename);
    strcpy(s1,filename);
    strcpy(s2,s1);
    strcat(s2,".input");
    fout1=fopen(s2,"w");
   if ((fin0=fopen(s1,"r"))==NULL)
   {printf ("Error in creating file %s \n", s1); exit(1);}
   if ((fout1=fopen(s2,"w"))==NULL)
   {printf ("Error in creating file %s \n", s2); exit(1);}

    fseek(fin0,0,0);
 for (j=1; j < 2; j++)
    {
    fscanf(fin0,"%s %d %s %s %d %s %f %lg %s", no1,&no2,no3, no4, &no5,  mon, &date,
&time, no7);
    printf ("excuse me! i am here!  %lg \n",time);
    frac = modf(time, &intr);
    frac=frac*100;
    minutes1= date*24*60+intr*60+frac;
    printf ("excuse me! i am here!  %f \n",minutes1);
    }
    fseek(fin1data,0,0);
    fseek(fin0,0,0);
 for (j=1; j < cntpnts+1; j++)
    {
    fscanf(fin0,"%s %d %s %s %d  %s %f %lg %s", no1,&no2,no3, no4, &no5,  mon, &date,
&time, no7);
    frac = modf(time, &intr);
    frac=frac*100;
    minutes= (24*date*60+intr*60+frac)-minutes1;
    printf ("excuse me! i am here!  %f \n",minutes);
    printf ("excuse me! i am here!  %s \n",no7);
```

```
   fprintf(fout1," %5.3f %10s \n", minutes, no7);
   }
 fclose(fout1);
 fclose(fin0);
  }
 fclose(fin1data);
 fclose(fin1);
printf("\n ==> enjoy baby ..its FINISHED !\n");
}
```

-- -------------
     Programme #5   STD-NMR Quantitative anal ysis for 1D STD experiments with interleaved 1D for aggregation kinetics

----
```
/* Written by Saravanakumar Narayanan, AG Bernd Reif and AK Horst Kessler */

/* Usage after compilation: a.out   <input>
<input> contains (how_many_peaks_in_pp_list) (   number of time pts in ref)  (time pts in std)
<peaklist_for_stdfrom2rr>       <file_time_1dexpts>     <file_time_stdexpt>     <si>    <2rr_ref>
<2rr_std> <ns_std> <ns_ref> <conc_li  g_mM> <molarexcess_lig_to_protein> \n");
    printf("  <noise near peak of interest> <file_out>\n");

useful 3 outputs:    aggregation  kinetics,  time  dependent  protein-ligand  interaction,
amplification factor with error propagaion   */

#include <stdio.h>
#include <string.h>
#include <math.h>
#define MAXVEC 32768
#define max  1


FILE *fout1, *fout2, *fout3, *fout4, *fout5, *fout6, *fout7, *fin, *fin2, *fin1pp, *fin3dataref,
*fin3datastd, *fin2timeref, *fin2timestd;
int i, j, a, cntref, cntstd, cntpeak, k, si2, peakpnt, noise,nsstd, nsref;
float concold, molarexcess;
int  re1[MAXVEC];
float    slope, noiseref4,noiseref,noiseint0,noiseint2,noiseint1, noiseint, noisestd, noisestd,
noisestd1, noisestd0, noiseref0, noiseref1;
float  refint, refint3,refint4,refint5, stdint, Ampfactor, conc, stdint5,stdint0;
float    refnoise,concnoise,noisenew,refnoise0, refnoise3, refnoise1, refnoise2, stdnoise0,
stdnoise2, stdnoise1,stdnoise3;
float
stdnoise,noiseref3,scalenoise,noisestdscale,timestd0,timestd,timeref0,timeref,timeref1,timeref
2, timerefnew;
float                                                  errorstd,errorref,errorref0,noiserefmean,
noisestdmean,noiseref1,noiseref2,errorstd0,errorstd1,errorstd2,noiserefmod,noisefinal,errorpr
opagation;
float scale,ref,intstdscale,intstdpercent, intensitynoisemod,Intensityrefmod;
float intstd0, snratioref, snratiostd, snratioref0, snratiostd0, snratiorefnew, snratioall,
snratioref1, snratioref2, intstd;
```

```c
float intref,intref0,refint0,intref1,intref2, timerefnew1, timerefnew2, timestdnew, timestdnew1,
timerefnew;
char peaknam[140],data_pp[140],data_t imeref[140],data_2rrref[140];
char data_out[140],s1[140],s2[140],s3[ 140],s4[140], s5[140], s6[140], s7[140];
char data_timestd[140] ,data_2rrstd[140] ;


main(argc,argv)
int argc;
char *argv[];

{

if (argc != 2)
   { printf("\n std_from_2rr_fi  nal: wrong number of arguments\n\n");
     printf("USAGE: std_from_2  rr_final <std_from_2rr_final.in>\n");
     printf("Synthax <std_from_2rr_final.in>: \n");
     printf("    <how_many_peaks_in_pp_lis  t> <time  pts  in  ref>  <time  pts  in  std>
<peaklist_for_stdfrom2rr> \n");
     printf("   <file_time_ref> <file_time_st   d> <si> <2rr_ref> <2rr_std>  <ns_std> <ns_ref>
<conc_lig_mM> <molarexce ss_lig_to_protein> \n");
     printf("   <noise near peak of interest> <file_out>\n");
     exit(1);
   }


 if ((fin=fopen(argv[1],"r"))==NULL)
    { printf("ERROR opening file %s !\n",argv[1]); exit(1); }


fscanf(fin,"%d",&cntpeak);
fscanf(fin,"%d",&cntref);
fscanf(fin,"%d",&cntstd);
fscanf(fin,"%s",data_pp);
fscanf(fin,"%s",data_timeref);
fscanf(fin,"%s",data_timestd);
fscanf(fin,"%d",&si2);
fscanf(fin,"%s",data_2rrref);
fscanf(fin,"%s",data_2rrstd);
fscanf(fin,"%d",&nsstd);
fscanf(fin,"%d",&nsref);
fscanf(fin,"%f",&concold);
fscanf(fin,"%f",&molarexcess);
fscanf(fin,"%d",&noise); /* noise po int for the peak of interest */
fscanf(fin,"%s",data_out);

if ((fin1pp=fopen(data_pp,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_pp); exit(1); }

if ((fin2timeref=fopen(data_timeref,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_timer ef); exit(1); }
```

```c
if ((fin2timestd=fopen(data_timestd,"r"))==NULL)
  { printf("ERROR opening file %s !\n",data_timestd); exit(1); }

if ((fin3dataref=fopen(data_2rrref,"r"))==NULL)
  { printf("ERROR opening file %s !\n",data_2rr ref); exit(1); }

if ((fin3datastd=fopen(data_2rrstd,"r"))==NULL)
  { printf("ERROR opening file %s !\n",data_2rrstd); exit(1); }


for (k=1; k < cntpeak+1; k++){
  fscanf(fin1pp,"%s %i",peaknam, &peakpnt);

  strcpy(s2,data_out);
  strcat(s2,"_");
  strcat(s2,peaknam);
  strcat(s2,"_");
  strcpy(s1,s2);
  strcpy(s3,s2);
  strcpy(s4,s2);
  strcpy(s5,s2);
  strcpy(s6,s2);
  strcat(s1,"ref");
  strcat(s2,"std");
  strcat(s3,"all");
  strcat(s4,"noise1");
  strcat(s5,"amp");
  strcat(s6,"noise2");
 if ((fout1=fopen(s1,"w"))==NULL)
  {printf ("Error in creating file %s \n", s1); exit(1);}
 if ((fout2=fopen(s2,"w"))==NULL)
  {printf ("Error in creating file %s \n", s2); exit(1);}
 if ((fout3=fopen(s3,"w"))==NULL)
  {printf ("Error in creating file %s \n", s3); exit(1);}
 if ((fout4=fopen(s4,"w"))==NULL)
  {printf ("Error in creating file %s \n", s4); exit(1);}
 if ((fout5=fopen(s5,"w"))==NULL)
  {printf ("Error in creating file %s \n", s5); exit(1);}
 if ((fout6=fopen(s6,"w"))==NULL)
  {printf ("Error in creating file %s \n", s5); exit(1);}


fprintf(fout3,"#STD data\n");
fprintf(fout3,"#t(min)   RefInt      STDI nt(unscaled)    norm_STD              STD-
Amplification\n");
fprintf(fout5,"#timestd   STD-Amp lification  STD-Norm-intensity\n");

  /* Create Output File for REF at specific Peak Position */

  fseek(fin2timeref,0,0);
```

```
fseek(fin3dataref,0,0);
for (i=1; i < 2; i++) {
    fscanf(fin2timeref,"%f",&timeref);
    fread(&re1[0],4,si2,fin3dataref);
   refnoise0=0;
   for (j=1;j < 101;j++){
    refnoise = re1[noise+j];
    refnoise0+=refnoise;
          }
    refnoise0=refnoise0/100;
   refint0=0;
   for (j=1;j < max+1;j++){
    refint4 = re1[peakpnt+j];
    refint0+=refint4;
          }
    refint0=refint0/max;
    }
fseek(fin2timeref,0,0);
fseek(fin3dataref,0,0);
   refnoise2=0;
   refint4=0;
for (i=1; i < cntref+1; i++){
    fscanf(fin2timeref,"%f",&timeref);
    fread(&re1[0],4,si2,fin3dataref);
   for (j=1;j < 101;j++){
    refnoise = re1[noise+j];
    refnoise2+=refnoise;
          }
    refnoise2=refnoise2/100;
   for (j=1;j < max+1;j++){
    refint = re1[peakpnt+j];
    refint4+=refint;
          }
    refint4=refint4/max;

   noiseref0 = 0;
   for (j=1; j <500+1; j++) {
   noiseref1 = re1[j+1000];
   if (noiseref1<0){
     noiseref1=-noiseref1;
       }
   noiseref0 = noiseref0 + noiseref1;
    }
   noiserefmean = noiseref0/500;
   conc = refint4/refint0*concold;
   concnoise = refnoise2/refnoise0*concold;

   printf("%7.4f %7.4f %7.4f %7.4f\n", noiserefmean,refint4,conc,concnoise);
   fprintf(fout4,"%6.4f \n",  noiserefmean);
    }
   fclose(fout4);
```

```
fout4=fopen(s4,"r");
fseek(fin2timeref,0,0);
fseek(fin3dataref,0,0);
fseek(fout4,0,0);
for (i=1; i < cntref+1; i++){
    fscanf(fin2timeref,"%f",&timeref);
    fscanf(fout4,"%f", &noiserefmean);
    /* noiserefmean=2000000.0; */
    fread(&re1[0],4,si2,fin3dataref);
    refint5=0;
    refnoise3=0;
  for (j=1;j < 101;j++){
   refnoise = re1[noise+j];
   refnoise3+=refnoise;
         }
   refnoise3=refnoise3/100;
  for (j=1;j < max+1;j++){
   refint = re1[peakpnt+j];
   refint5+=refint;
           }
   refint5=refint5/max;
   noisestd0 = 0;
   noiseref0=0;
  for (j=1; j < 500; j++) {
   noiseref1 = re1[j+1000];
   if (noiseref1<0){
      noiseref1=-noiseref1;
        }
   noiseref = (noiseref1-noiserefmean)*(noiseref1-noiserefmean) ;
    /*   printf("noiserefmean=   %7.4f   noiseint=   %7.4f   rmsd=   %7.4f   \n",
noiserefmean,noiseref1,noiseref); */
   noiseref0=noiseref+noiseref0;
   }
   noiseref= sqrt(noiseref0/500); /*adjustment factor 15*/
  /*noiseref=sqrt(noiseref0)/500;*/
   printf("%7.4f  %7.4f\n", noiseref0,noiseref);
   fprintf(fout1,"%7.1f %7.4f %7.4f %12.4f\n", timeref, refint5, refnoise3,noiseref);
   }

/* Create Output File for STD at specific Peak Position */
fseek(fin2timestd,0,0);
fseek(fin3datastd,0,0);
for (i=1; i < cntstd+1; i++){
    fscanf(fin2timestd,"%f",&timestd);
    fread(&re1[0],4,si2,fin3datastd);
     stdnoise0=0;
     stdint0=0;
   for (j=1;j < 101;j++){
    stdnoise = re1[noise+j];
    stdnoise0+=stdnoise;
```

```
      }
    stdnoise0=stdnoise0/100;
   for (j=1;j < max+1;j++){
    stdint = re1[peakpnt+j];
    stdint0+=stdint;
          }
    stdint0=stdint0/max;
   noisestd0 = 0;
   for (j=1; j < 500; j++) {
   noisestd1 = re1[j+1000];
   if (noisestd1<0){
      noisestd1=-noisestd1;
       }
   noisestd0 = noisestd0 + noisestd1 ;
    }
    noisestdmean=noisestd0/500;
    fprintf(fout6," %6.1f \n",  noisestdmean);
    printf("%7.4f \n",noisestdmean);
    }
    fclose(fout6);
fout6=fopen(s6,"r");
fseek(fin2timestd,0,0);
fseek(fin3datastd,0,0);
fseek(fout6,0,0);
for (i=1; i < cntstd+1; i++){
    fscanf(fin2timestd,"%f",&timestd);
    fscanf(fout6,"%f",  &noisestdmean);
    /* noisestdmean=6000000.0; */
    fread(&re1[0],4,si2,fin3datastd);
    stdnoise1=0;
    stdint5=0;
   for (j=1;j < 101;j++){
    stdnoise = re1[noise+j];
    stdnoise1+=stdnoise;
        }
    stdnoise1=stdnoise1/100;
   for (j=1;j < max+1;j++){
    stdint = re1[peakpnt+j];
    stdint5+=stdint;
         }
    stdint5=stdint5/max;
   noisestd0 = 0;
   for (j=1; j < 501; j++) {
   noisestd1 = re1[j+1000];
   if (noisestd1<0){
      noisestd1=-noisestd1;
       }
   noisestd = (noisestd1-noisestdmean)*(noisestd1-noisestdmean) ;
   noisestd0 = noisestd +noisestd0;
    }
    noisestd= sqrt(noisestd0/500);/*adjustment factor 15*/
```

```
      /*noisestd=sqrt(noisestd0)/500;*/
      errorstd=noisestd/stdint5;
      errorstd=errorstd*errorstd;
      fprintf(fout2,"%7.1f %10.1f %10.1f %12.4f\n", timestd, stdint5, stdnoise1,errorstd);
      }
   fclose(fout2);
   fclose(fout4);
   fclose(fout1);
   fclose(fout6);



   /* Start STD Calculations */
   fout1=fopen(s1,"r");
   fout2=fopen(s2,"r");

/*
   fseek(fout1,0,0);
   fseek(fout2,0,0);
*/

   fscanf(fout2,"%f %f %f %f", &timestd0, &intstd0,&noisestd0,  &errorstd0);
   fscanf(fout1,"%f %f %f %f",&timeref0, &intref0, &noiseref4, &noiseref0);

   fseek(fout2,0,0);
   for (i=1; i < cntstd+1; i++){
      fscanf(fout2,"%f %f %f %f",& timestd, &intstd, &noisestd, &errorstd);
      timeref1=-0.2;
      fseek(fout1,0,0);
       while (timeref1<timestd){
          fscanf(fout1,"%f %f %f %f", &timeref1, &intref1, &noiseref, &noiseref1);
          }
      fscanf(fout1,"%f %f %f %f",&timeref2, &intref2, &noiseref3, &noiseref2);
      timerefnew1= 1/timeref1;
      timerefnew2= 1/timeref2;
      timestdnew= 1/timestd;
      ref= intref1+ ((intref2-intref1)*(timestd-timeref1)/(timeref2-timeref1));
      noisenew= noiseref+ ((noiseref3-noiseref)*(timestd-timeref1)/(timeref2-timeref1));
      scale= ref/intref0;
      intstdscale= intstd/scale;
      intstdpercent= intstd*100;
      Intensityrefmod= ref*nsstd/nsref;
      intensitynoisemod= 1.414*noisenew*(nsstd/nsref);
      errorref=noiseref1;
      /*errorref = noiseref1/ref;*/
      errorref = errorref*(sqrt(nsstd)/sqrt(nsref))/(Intensityrefmod-intensitynoisemod);
      errorref = 2*errorref*errorref;
      errorpropagation = sqrt(errorref+errorstd);
      Ampfactor= molarexcess*(intstd-noisestd)/(Intensityrefmod-intensitynoisemod);
      printf("printed    finally   %7.3f    %7.3f,    %7.3f\n",  Ampfactor, (intstd-noisestd),
(Intensityrefmod-intensitynoisemod));
      if (ref<0)
```

```
    {
     Ampfactor=0;
     intstd=0;
     Intensityrefmod=1;
     i=cntstd+1;
    }
    errorpropagation=errorpropagation*Ampfactor;
    printf("%7.4f %7.4f %7.4f \n", errorref,errorstd,errorpropagation);

    printf("i = %2d, timeref1 = %7.1f, timeref2 = %7.1f, timestd = %7.1f, ref = %10.1f,
scale = %4.3f, Ampfactor = %5.3f, errorpropagation = %6.4f \n",
         i, timeref1, timeref2, timestd, ref, scale, Ampfactor, errorpropagation );
    fprintf(fout3,"%8.1f %8.4f %12.4f %12.4f %12.4f\n", timestd, scale, intstd/intstd0,
intstd/Intensityrefmod, Ampfactor);
    fprintf(fout5,"%6.4f %6.4f %6.4f  \n", timestd, Ampfactor, errorpropagation);
    }

  }

 fclose(fin1pp);
 fclose(fin2timeref);
 fclose(fin2timestd);
 fclose(fin3dataref);
 fclose(fin3datastd);
 fclose(fout1);
 fclose(fout2);
 fclose(fout3);
 fclose(fout4);
 fclose(fout5);

 printf("\n ==> FINISHED !\n");

}


------------------------------------------------------------------------------------------------ --------------
      Programme #6          Plots gradie nt strength vs concentration/intensity from DOSY

-----------------------------------------------------------------------------------------------------------------

/* Written by Saravanakumar Narayanan,AG Bernd Reif and AK Horst Kessler*/

/* Usage after compilation: a.out   <concmap.input>
<concmap.input> contains  number of peaks, number of gradient points, <peakpick file>
<gradient file> size_usually 32k, <2rrfile> <outputfilename> <noise region> */

#include <stdio.h>
#include <string.h>
#define MAXVEC 32768
```

```
FILE *fout1, *fout2, *fout3, *fin, *fin2, *fin1pp, *fin2temp, *fin3datatemp ;
int l, i, a, cnttemp, cntgrad, j, cntpeak, k, si2, peakpnt, peakpnt2, noise;
int  re1[MAXVEC], ns1, ns2,  s2, difftime1,  max;
float     slope1,conc1,  conc2,  slope,  refnoise,gradient,  refnoise0,  refnoise3,refnoise2,
refnoise1,refint, refint6, refint0, refint2, refint1, concentration ;
float  timeref0,timeref,timeref1,timeref2;
float intref,intref0,intref1,intref2;
char peaknam[140],data_pp[140],data_temp[140],data_2rrtemp[140];
char data_out[140],s1[140], s3[140];

main(argc,argv)
int argc;
char *argv[];

{

if (argc != 2)
   { printf("\n concentratindiff: wrong number of arguments\n\n");
    printf("USAGE: concentrationdiff <concmap.in>\n");
    printf("Synthax <concmap.in>: \n");
    printf("<number_of_peaks>   <number_of_temppoints> \n");
    printf(" <peak_pick_file>  <temp_file>  <size>  <2rr_tem p>  <dataout>  <noise> \n");
    exit(1);
   }


 if ((fin=fopen(argv[1],"r"))==NULL)
    { printf("ERROR opening file %s !\n",argv[1]); exit(1); }


fscanf(fin,"%d",&cntpeak);
fscanf(fin,"%d",&cnttemp);
fscanf(fin,"%s",data_pp);
fscanf(fin,"%s",data_temp);
fscanf(fin,"%d",&si2);
fscanf(fin,"%s",data_2rrtemp);
fscanf(fin,"%s",data_out);
fscanf(fin,"%d",&noise);


if ((fin1pp=fopen(data_pp, "r"))==NULL)
   { printf("ERROR opening file %s !\n",data_pp); exit(1); }

if ((fin2temp=fopen(data_temp,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_temp); exit(1); }


if ((fin3datatemp=fopen(data_2rrtemp,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_2rrte mp); exit(1); }
```

```c
    fseek(fin1pp,0,0);
for (k=1; k < cntpeak+1; k++){
  fseek(fin3datatemp,0,0);
  fscanf(fin1pp,"%s %d %d",peaknam, &peakpnt, &max);
   printf("i am here baby %s %d !\n",peaknam, peakpnt );

  strcpy(s1,data_out);
  strcat(s1,"_");
  strcat(s1,peaknam);
  strcat(s1,"_");
  strcat(s1,"gradient");
  fout1=fopen(s1,"w");

fprintf(fout1,"#gradient     relative concentration\n");


  /* Create Output File for REF at specific Peak Position */

  fseek(fin3datatemp,0,0);
  fseek(fin2temp,0,0);

  for (i=1; i < 2; i++){
     refint=0;
     refnoise0=0;
     fread(&re1[0],4,si2,fin3datatemp);
     for (j=1;j < max+1; j++){
      refnoise = re1[1000+j];
      if (refnoise<0) {
           refnoise=-refnoise;
             }
      refnoise0=refnoise+refnoise0;
         }
      refnoise2=refnoise0/max;

     for (j=1;j < max+1;j++){
      refint0 = re1[peakpnt+j];
         if (refint0<-refnoise2) {
            refint0=-refnoise2;
                }
        refint0=refint0+refnoise2;
        refint=refint+refint0;
          }
          }

     fseek(fin2temp,0,0);
     fseek(fin3datatemp,0,0);

  for (i=1; i < cnttemp+1; i++){
     refnoise1=0;
     refint2 =0;
     fread(&re1[0],4,si2,fin3datatemp);
```

```c
    fscanf(fin2temp,"%f ",&gradient);
   for (j=1;j < max+1; j++){
   refnoise = re1[1000+j];
   if (refnoise<0) {
        refnoise=-refnoise;
          }
   refnoise1=refnoise+refnoise1;
        }
    refnoise3=refnoise1/max;
   for (j=1;j < max;j++){
   refint1 = re1[peakpnt+j];
      if (refint1<-refnoise3) {
        refint1=-refnoise3;
            }
      refint1=refint1+refnoise3;
      refint2=refint1+refint2;
        }

   concentration = refint2/refint;
   printf("i am here baby %f %f %f!\n",refint1, refint2 , concentration);
   fprintf(fout1," %f %f \n", gradient, concentration);
  printf("i = %2d, gradient = %7.1f, concentration = %7.3f, refint0 = %7.1f,  \n", i,
gradient, concentration, refint0);


   }
 fclose(fout1);
}
 fclose(fin3datatemp);
 fclose(fin2temp);
 fclose(fin1pp);
printf("\n ==> FINISHED !\n");
}
```

-- --------------
        Programme #7   DOSYNMR analysis fo r Internal standard method

----

```c
/* Written by Saravanakumar Narayanan,AG Bernd Reif and AK Horst Kessler*/

/*   reads out intensities directly from 2rr files and calculating molecular weight,
hydrodynamic radii, slopes  diffusion coeff*/
/*  good for the titrations and time dependent studies and designed for diffusion analysis
internal reference method*/
/*  this programme basically integrates the set of points of interest and using weighted least
square method to fit the slopes */
/* slope should be given by calculating the decay rate and it will automatically calculate how
much remains*/

#include <stdio.h>
```

```c
#include <string.h>
#include <unistd.h>
#include "math.h"

#define MAXVEC 32768
#define diffcoeffbuf 1.192   /*(exp(-05)) cm2 per sec */
#define radhydbuf 2.48      /*(exp(-10 ))  in angstrom  from pdb file*/
#define totaltime  187         /* fill the total time of the expt in min. */
#define radhydglobular 11.67            /* fill the  value (4.75+/-1.11 times N^0.29 +/- 0.02)
where N is number of AA residues*/
#define radhyddenature 13.3         /* fill the  value (2.21+/-1.07 times N^0.57 +/- 0.02)*/
#define Molecularweightbuf 192

FILE    *fout1,   *fout2,   *fout3,*fout4,   *fout5,*fin,*fout6,*fin2,   *fin1pp,   *fin2time,
*fin3datadiff, *fin4file;
int line2,line, l, i,  j,  k, si2, peakpnt, noise, noise2 ;
int  cntfiles,cntgrad, re1[MAXVEC],max2,max,bufferpnt, difftime0;

double  gradient0, gradient3, gradient4,finalgrad,gradient,difftime2;
double    slope2,slope1, refnoise,time,  refnoise0,  refnoise3,refnoise2,  refnoise1,refnoise4,
refnoise5;
double refnoise6a,refbuf, refbuf1, refbuf0,refbuf2,scale1,scale2,finalgrad;
double refnoise6, refnoise7, refnoise8, refint,  refint0, refint2, refint1;
double wp1,wb1,wp,wb,covarpep,covarbuf,varpep,varbuf,corrpep,corrbuf;
double diffusionbuf0,diffusionpep0,diffusionpep1,diffusionpep, diffusionbuf, diffusionbuf1 ;
double
sx,y2,slopepep,slopebuf,interpep,interbuf,diffcoeffpep,radhydpep,diffcoeffnorm,diffcoeffref,
Molecularweightnorm,Molecularweight;
double Molecularweightref,Molecularweightnorm,cref,cpep,cnorm,radhydref,radhydnorm;
double xy1,xy2,x1bar,x2bar,y1bar,y2bar,deltapep,deltabuf,y0p,x1,x2;
double sxx,x, corrpep, corrbuf, slopepep, slopebuf,interpep, interbuf,xbar,y1bar,y2bar;
double sx1,sx2,sy1,sy2,sxx1,sxx2,sxy2,sxy1,syy1,syy2,delta;
double y0d,y2d;

char inbuf0[500], buffernam[140],peaknam[140],data_2rrtime [140],data_pp[140];
char               data_2rr[140],data_time[140],da ta_2rr_all[140],data_out[140],s1[140],
s2[140],s3[140],s4[140],s5[140],s6[140],s0[140];


main(argc,argv)
int argc;
char *argv[];

{

 if (argc != 2)
   { printf("\n diffmolwt: use cc -lm to link the <math.h> library \n");
    printf("\n diffmolwt: wrong number of arguments\n\n");
    printf("USAGE: diffmolwt <diffusion.in>");
    printf("Synthax <diffusion.in>: \n");
    printf("<number_of_files >   <no_of_gradientsteps>\n");
```

```
    printf("        <strength_of_final_gradi ent>        <peak_pick_file>        <time_file>        <size>
<2rr_files_containing_all_2   rr_diffusion_file>\n");
    printf(" <data_out> <noise_near _peak_of_interest> <noi  se_near_buffer>\n");
    printf("Synthax <peakpickfile.in>: peakname, buffername, peakpnt, bufferpnt, max \n");
    printf("Synthax <timefile.in>: time_in_minutes\n");
    printf("Synthax                                                           <2rrfiles.in>:
2rr_files_processed_with_32k_pnts_   in_direct_dimension\n");
    printf("Synthax <2rrfiles>: 2rrfile, slope1(ref), slope2(buffer)\n");
    exit(1);
  }


 if ((fin=fopen(argv[1],"r"))==NULL)
    { printf("ERROR opening file %s !\n",argv[1]); exit(1); }


 fscanf(fin,"%d",&cntfiles);
 fscanf(fin,"%d",&cntgrad);
 fscanf(fin,"%lg",&finalgrad);
 fscanf(fin,"%s",data_pp);
 fscanf(fin,"%s",data_time);
 fscanf(fin,"%d",&si2);
 fscanf(fin,"%s",data_2rr_all);
 fscanf(fin,"%s",data_out);
 fscanf(fin,"%d",&noise); /* noise po int for the peak of interest */
 fscanf(fin,"%d",&noise2); /*  noise point for the buffer*/


 if ((fin1pp=fopen(data_pp, "r"))==NULL)
   { printf("ERROR opening file %s !\n",data_pp); exit(1); }

 if ((fin2time=fopen(data_time,"r"))==NULL)
   { printf("ERROR opening file %s !\n",data_time); exit(1); }

 if ((fin4file=fopen(data_2rr_all, "r"))==NULL)
   { printf("ERROR opening file %s !\n",data_2rr_a ll); exit(1); }

   fseek(fin4file,0,0);
   fseek(fin2time,0,0);

   strcpy(s0,data_out);
   strcat(s0,"_");
   strcpy(s3,s0);
   strcpy(s4,s0);
   strcpy(s5,s0);
   strcpy(s6,s0);
   strcat(s3,"slopes");
   strcat(s4,"diffusioncoeff");
   strcat(s5,"Molecularweight");
   strcat(s6,"radhyd");
   fout3=fopen(s3,"w");
```

```c
    fout4=fopen(s4,"w");
    fout5=fopen(s5,"w");
    fout6=fopen(s6,"w");

 for (k=1; k < cntfiles+1; k++){
    fscanf(fin4file,"%s %lg %lg",data_2rr,&slope1,&slope2);
    strcpy(s1,s0);
    strcat(s1,data_2rr);
    strcat(s1,"_");
    strcpy(s2,s1);
    strcat(s1,"intensityref");
    strcat(s2,"intensitybuf");
    fout1=fopen(s1,"w");
    fout2=fopen(s2,"w");
 /* slope = 0 for no concentration flux during the time lapse  otherwise negative slope  enter
only the difference in the normalized intensity */
    fscanf(fin2time," %lg", &time);

    if ((fin3datadiff=fopen(data_2rr,"r"))==NULL)
    { printf("ERROR opening file %s !\n",data_2 rr); exit(1); }


    /* Create Output File for REF at specific Peak Position */

    fseek(fin1pp,0,0);
    fscanf(fin1pp,"%s %s %d %d %d %d", peaknam, buffernam, &peakpnt, &bufferpnt,
&max, &max2);

    fseek(fin3datadiff,0,0);

/*fprintf(fout1,"#gradient strength  norm. int.peptide        norm.intensity.buffer\n");
fprintf(fout2,"#time            slope(pep)    slope(ref )  correlationpep correlationref\n");
fprintf(fout3,"#time            diffusioncoeff(norm) diffusion  coefficient(pep)     c   value
Molecularweight \n"); */

    for (i=1; i < 2; i++){
        refint=0;
        refbuf=0;
        refnoise0=0;
        refnoise5=0;
        fread(&re1[0],4,si2,fin3datadiff);
        for (j=1;j < 501; j++){
         refnoise = re1[noise+j];
         refnoise4 = re1[noise2+j];
         refnoise = (refnoise<0) ? -refnoise : refnoise;
         refnoise4 = (refnoise4<0) ? -refnoise4 : refnoise4;
         refnoise0+=refnoise;
         refnoise5+=refnoise4;
             }
        refnoise2=refnoise0/500;
        refnoise6=refnoise5/500;
```

193

```
    for (j=1;j < max+1;j++){
    refint0 = re1[peakpnt+j];
       refint0=(refint0/(slope1*(totaltime/cntgrad)+1));
        refint0= (refint0 <-refnoise2) ? -refnoise2: refint0;
       refint0+=refnoise2;
       refint+=refint0;
          }
     for (j=1;j < max2+1;j++){
      refbuf0 = re1[bufferpnt+j];
        refbuf0=(refbuf0/(slope2*(totaltime/cntgrad)+1));
          refbuf0 = (refbuf0<-refnoise6) ? -refnoise6 : refbuf0;
       refbuf0+=refnoise6;
       refbuf+=refbuf0;
          }
          }
     fseek(fin3datadiff,0,0);
    difftime2=totaltime/cntgrad;
    gradient=finalgrad/cntgrad;
    gradient0 =0;
    difftime0=0;


 for (i=1; i < cntgrad+1; i++){
     difftime0 += difftime2;
     gradient0 += gradient;
     refbuf2=0;
     refnoise7=0;
     refnoise1=0;
     refint2 =0;
     fread(&re1[0],4,si2,fin3datadiff);
     for (j=1;j < 501; j++){
     refnoise = re1[noise+j];
      refnoise6a = re1[noise2+j];
       refnoise= (refnoise<0) ? -refnoise: refnoise;
       refnoise6a= (refnoise6a<0) ? -refnoise6a: refnoise6a;
       refnoise1+=refnoise;
      refnoise7+=refnoise6a;
           }
      refnoise3=refnoise1/500;
      refnoise8=refnoise7/500;

     scale1 = (slope1*difftime0)+1;
     scale2 = (slope2*difftime0)+1;
     for (j=1;j < max+1;j++){
     refint1 = re1[peakpnt+j];
         refint1=refint1/scale1;
         refint1= (refint1<-refnoise3) ? -refnoise3 : refint1;
         refint1+=refnoise3;
         refint2+=refint1;
         }
```

```
    for (j=1;j < max2+1;j++){
     refbuf1 = re1[bufferpnt+j];
          refbuf1=refbuf1/scale2;
          refbuf1=(refbuf1<-refnoise8) ? -refnoise8 : refbuf1;
          refbuf1+=refnoise8;
          refbuf2+=refbuf1;
          }
     diffusionpep = (refint2/refint); /*because ln(1+x)= x-x*x/2 and so on...)*/
     printf(" i am here baby  %f %f %f!\n", diffusionpep,refint2, refint);
     diffusionpep1= log(diffusionpep);
     diffusionbuf = (refbuf2/refbuf);
     diffusionbuf1= log(diffusionbuf);
     if (i>0 && i<30){
    fprintf(fout1," %15.9lg % 15.9lg \n", gradient0*gradient0,diffusionpep1);
         }
     if (i>0 && i<15){
     fprintf(fout2," %15.9lg % 15.9lg \n", gradient0*gradient0, diffusionbuf1);
      }
      }
     fclose(fout1);
     fclose(fout2);
     fclose(fin3datadiff);
   /*weighted least square analysis is getting started to find the slope of the individual
curves */

     fout1=fopen(s1,"r");
     line = 0;     /* number of line */
     fseek(fout1,0,0);
     while (fscanf(fout1,"%s",inbuf0) != EOF){       /* cutting down the number of points
below a certain thershold */
     fscanf(fout1,"%s",inbuf0);                      /* because below that point all points are
wandering at 0 */
          line++;
     }

     fout2=fopen(s2,"r");
     fseek(fout2, 0,0);
     fout2=fopen(s2,"r");
     line2 = 0;     /* number of line */
     fseek(fout2,0,0);
     while (fscanf(fout2,"%s",inbuf0) != EOF){
     fscanf(fout2,"%s",inbuf0);
          line2++;
     }

     fseek(fout1, 0,0);
     fseek(fout2, 0,0);

     sy1=sy2=y0d=syy1=y2d=syy2=0; /*resetting the sum*/
   for (i=1; i < line+1; i++){
     fscanf(fout1,"%lg %lg ", &x1, &y0p);
```

195

```
      sy1+=y0p;
      y0d=y0p*y0p;
      syy1+=(y0d);
   }

  for (i=1; i < line2+1; i++){
      fscanf(fout2,"%lg %lg ", &x1, &y2);
      sy2+=y2;
      y2d=y2*y2;
      syy2+=y2d;
   }
 varpep= ((syy1-(sy1*sy1/line))/(line-1));  /*unbiased */
 varbuf= ((syy2-(sy2*sy2/line2))/(line2-1));
 wp1= 1/varpep; /*weightage factor */
 wb1= 1/varbuf;
 y0d=syy1=y2d=syy2=0; /*resetting the sum*/
 sx1=sx2=sy1=sy2=sxx1=sxx2=sxy1=sxy2=wp=wb=0; /*resetting the sum*/
  fseek(fout1, 0,0);
  fseek(fout2, 0,0);
  for (i=1; i < line+1; i++){
      fscanf(fout1,"%lg %lg ", &x1, &y0p);
      wp+=wp1;
      sx1+=(wp1*x1);
      sy1+=(wp1*y0p);
      sxx1+=(wp1*(x1*x1));
      syy1+=(wp1*(y0p*y0p));
      sxy1+=(wp1*(x1*y0p));
   }
  for (i=1; i < line2+1; i++){
      fscanf(fout2,"%lg %lg ", &x2, &y2);
      wb+=wb1;
      sx2+=(wb1*x2);
      sy2+=(wb1*y2);
      sxx2+=(wb1*(x2*x2));
      syy2+=(wb1*(y2*y2));
      sxy2+=(wb1*(x2*y2));
      }
 x1bar=sx1/line;
 x2bar=sx2/line2;
 y1bar=sy1/line;
 y2bar=sy2/line2;
 printf(" i am here baby  %d %d!\n", line, line2);
 deltapep = (sxx1*wp)-(sx1*sx1);
 deltabuf = (sxx2*wb)-(sx2*sx2);
 slopepep= ((sxy1*wp)-(sx1*sy1))/deltapep;
 slopebuf= ((sxy2*wb)-(sx2*sy2))/deltabuf;
 interpep= ((sxx1*sy1)-(sx1*sxy1))/deltapep;
 interbuf= ((sxx2*sy2)-(sx2*sxy2))/deltabuf;
 covarpep=(wp*sxy1)-(sx1*sy1/line);
 covarbuf=(wb*sxy2)-(sx2*sy2/line2);
 corrpep=covarpep/(sqrt((((wp*sxx1)-(sx1*sx1/line))*((wp*syy1)-(sy1*sy1/line)))));
```

```c
corrbuf=covarbuf/(sqrt(((wb*sxx2)-(sx2*sx2/line2))*((wb*syy2)-(sy2*sy2/line2))));
fprintf(fout3," %15.8lg %15.8lg %15.8lg %15.8lg %15.8lg \n", time, slopepep, slopebuf,
corrpep, corrbuf);
}
  fclose(fin3datadiff);
  fclose(fin2time);
  fclose(fin1pp);
  fclose(fout2);
  fclose(fout1);
  fclose(fout3);

 /* diffusion calculation starts from here */

  fout3=fopen(s3,"r");

 fseek(fout3,0,0);
 fseek(fin4file, 0,0);

 for (i=1; i < cntfiles+1; i++){
 fscanf(fout3,"%lg %lg %lg %lg %lg ", &time, &slopepep, &slopebuf, &corrpep, &corrbuf);
 diffcoeffpep= (slopepep/slopebuf)*(diffcoeffbuf);
 radhydpep= (slopebuf/slopepep)*radhydbuf;
 if (i==1) {
 radhydref = radhydpep;
       }
 radhydnorm= radhydpep/radhydref;
 /* cpep=((radhyddenature-radhydpep)/(radhyddenature-radhydglobular));   if the c is 1 then
the pep is denatured */
 if (i==1) {
 diffcoeffref = diffcoeffpep;
       }
 diffcoeffnorm=diffcoeffpep/diffcoeffref;

Molecularweight=((slopebuf*slopebuf*slopebuf)/(slopepep*slopepep*slopepep))*Molecular
weightbuf;
 if (i==1) {
 Molecularweightref =  Molecularweight;
       }
 Molecularweightnorm=Molecularweight/Molecularweightref;
 fprintf(fout4," %15.8lg %15.8lg %15.8lg  \n", time, diffcoeffnorm, diffcoeffpep);
 fprintf(fout5,"   %15.8lg   %15.8lg   %18.8lg           \n", time, Molecularweightnorm,
Molecularweight);
 fprintf(fout6," %15.8lg %15.8lg %15.8lg  \n", time, radhydpep, radhydnorm);
   }
 fclose(fin4file);
 fclose(fout3);
 fclose(fout4);
 fclose(fout5);
 fclose(fout6);
printf("\n ==> enjoy baby!.. its.. FINISHED !\n");
}
```

-- --------------
Programme #8   Analysis of TMREV experiment (solid state NMR)
Plots increment in mixing time vs concentration

/* Written by Saravanakumar Narayanan,AG Bernd Reif and AK Horst Kessler*/

```c
#include <stdio.h>
#include <string.h>
#define MAXVEC 32768


FILE *fout1, *fout2, *fout3, *fin, *fin2, *fin1pp, *fin3datatemp ;
int i,j, k, cnttemp, cntpeak,  si2, peakpnt ;
int  re1[MAXVEC], max;
float  increment, incrementnew, initial;
float  conc1, refint, refint6, refint10, refint0, refint2, refint1, concentration ;
char peaknam[140],data_pp[140],data_2rr[140];
char data_out[140],s1[140], s3[140];

main(argc,argv)
int argc;
char *argv[];

{

 if (argc != 2)
   { printf("\n concentratindiff: wrong number of arguments\n\n");
    printf("USAGE: concentrationdiff <concmap.in>\n");
    printf("Synthax <tmrevconc.in>: \n");
    printf("<number_of_peaks>        <ini tial  value  for  the  increment>  <increment>
<number_of_increments> \n");
    printf(" <peak_pick_file>     <size> <2rr> <dataout>  \n");
    exit(1);
   }


 if ((fin=fopen(argv[1],"r"))==NULL)
    { printf("ERROR opening file %s !\n",argv[1]); exit(1); }


fscanf(fin,"%d",&cntpeak); /* number of peaks to be analyzed*/
fscanf(fin,"%f",&initial);    /* initial time in ms*/
fscanf(fin,"%f",&increment);        /* size of the increment*/
fscanf(fin,"%d",&cnttemp); /* number of increments*/
fscanf(fin,"%s",data_pp);     /* peakpicking file*/
fscanf(fin,"%d",&si2);              /* size as per 2s si */
fscanf(fin,"%s",data_2rr);  /* after convert ing BRUKER 2rr file into binary using
convert2rr.c */
```

```c
fscanf(fin,"%s",data_out);    /*  desired output file*/


if ((fin1pp=fopen(data_pp, "r"))==NULL)
  { printf("ERROR opening file %s !\n",data_pp); exit(1); }


if ((fin3datatemp=fopen(data_2rr,"r"))==NULL)
  { printf("ERROR opening file %s !\n",data_2 rr); exit(1); }


  fseek(fin1pp,0,0);
for (k=1; k < cntpeak+1; k++){
  fseek(fin3datatemp,0,0);
  fscanf(fin1pp,"%s %d %d",peaknam, &peakpnt, &max);
  printf("i am here baby %s !\n",peaknam);

  strcpy(s1,data_out);
  strcat(s1,"_");
  strcat(s1,peaknam);
  strcat(s1,"_");
  strcat(s1,"tmrevconc");
  fout1=fopen(s1,"w");



  /* Create Output File for REF at specific Peak Position */

  fseek(fin3datatemp,0,0);

    refint10=0;
  for (i=1; i <2; i++){
    refint=0;
    fread(&re1[0],4,si2,fin3datatemp);
     for (j=1;j <max+1;j++){
     refint0 = re1[peakpnt+j];
       refint=refint+refint0;
         }
       refint10=refint+refint10;
        }
       refint0=0;

    fseek(fin3datatemp,0,0);

  for (i=1; i < cnttemp+1; i++){
     refint2 =0;
     fread(&re1[0],4,si2,fin3datatemp);
      incrementnew=initial+(i-1)*increment;
     for (j=1;j < max+1;j++){
     refint1 = re1[peakpnt+j];
         refint2=refint1+refint2;
```

```
          }
            concentration= (refint2/refint10)*100;
        printf("i am here baby %f %f !\n",refint1, refint2 );
        fprintf(fout1," %f %f \n", incrementnew,concentration);


    }
 fclose(fout1);
}
 fclose(fin3datatemp);
 fclose(fin1pp);
 printf("\n ==> FINISHED !\n");
}
```