UNIVERSITA' DEGLI STUDI DI PADOVA

Centro di Ricerca Interdipartimentale per le Biotecnologie Innovative

- CRIBI -

Biopolymer Mass Spectrometry Laboratory, Imperial College London (UK)

SCUOLA DI DOTTORATO DI RICERCA IN BIOSCIENZE E BIOTECNOLOGIE

INDIRIZZO DI BIOTECNOLOGIE

XXIV CICLO

Insight into the aggregation process of α-synuclein

Structural study of α-synuclein covalent dimers

Direttore della scuola: Ch.mo Prof. Giuseppe ZANOTTI

Coordinatore di Indirizzo: Ch.mo Prof. Giorgio VALLE

Supervisore: Prof. Patrizia POLVERINO DE LAURETO

Dottoranda: Micaela PIVATO

A. A. 2011/2012

THESIS CONTENTS

List of tables

List of figures

Abbreviations

aS	α-Synuclein
NN	dimer produced by the link of N-terminal regions of aS
CC	dimer produced by the link of C-terminal regions of aS
NC	dimer produced by the link of N- and C-terminal regions of aS
DC	*double core* dimer, composed by 1-104 and 29-140 residues of aS

AFM	Atomic force microscopy
BSA	Bovine serum albumine
CD	Circular dichroism
CNS	Central nervous system
DLB	Dementia with Lewy bodies
FT-IR	Fourier transform infrared spectroscopy
FPLC	Fast protein liquid chromatography
LBs	Lewy bodies
MS	Mass spectrometry
MW	Molecular weight
SDS	Sodium dodecyl sulphate
SN	Substantia nigra pars compacta
PD	Parkinson's disease
RP-HPLC	Reverse phase high pressure liquid chromatography
TEM	Transmission electron microscopy
ThT	Thioflavin T
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
Glu	E	Glutamic acid
Gln	Q	Glutamine
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine

Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

Summary

My PhD thesis is composed of two parts. A part deals with the characterization of α-synuclein (aS) dimers aggregation properties in respect to those of aS. The experimental work was conducted at CRIBI laboratory, at University of Padua, and constitutes the main project in which I was involved. During the third year of my PhD I spent six months at the Biopolymer Mass Spectrometry Laboratory of Imperial College in London. I conducted a glycomic analysis of mice tissues and a pilot study on expression and biosynthesis of mixed linked glucans emicellulose.

Parkinson's disease (PD) is a progressive, neurodegenerative disorder characterized by the loss of dopaminergic neurons in *substantia nigra*. The histological hallmarks of PD are intracellular inclusions, known as Lewy bodies (LBs), composed by filamentous and aggregated protein. The pathogenesis of the disease is still unclear, but a key step in the onset of PD is the aggregation of aS into amyloid fibrils, that deposit within LBs as the major component. Despite its importance in neurodegeneration, little is known about aS function, native physiological state and mechanism of aggregation. aS was recently described as a folded tetramer, but was generally considered a natively unfolded protein. aS is able to acquire α-helix conformation upon interaction with lipids and to convert to β-structure in pathological processes. During the aggregation process, aS forms soluble oligomers, transient β-structured intermediate between the physiological form of aS and amyloid fibrils.

Dimerization of aS could represent a critical, rate-limiting step in the aggregation and amyloid formation of the protein. Therefore, we decided to study the aggregation of several different dimers of aS, produced through molecular biology techniques. A cysteine residue has been added at the N-terminal or at the C-terminal of aS, therefore producing a dimer N-N or C-C linked through a disulfide bond. A N-C dimer, formed two consecutive aS molecules, was obtained as a single polypeptide chain. During the project, another dimer, called DC dimer, was produced in order to further draw up the hydrophobic regions, and avoid the interferences of side chains within the molecule. DC is constituted by two consecutive central, highly amyloidogenic regions, containing aS residues from 1 to104 joined to residues from 29 to 140. The dimers represent a suitable tool for the study of intramolecular aS interaction pathway. Some remarkable differences define and limit the mobility freedom of the dimers respect to aS, hypothetically differentiating the fibrillation process of the four protein structures.

The characterization of the dimers was performed using chemical and biophysical techniques in order to define their behaviour in solution as monomer. CD, IR and NMR spectroscopy studies show that all the dimers are unfolded. They undergo α-helical transition upon interaction with the detergent SDS. These results evidence that dimers strongly resemble aS conformational features. All the dimers were tested for the ability to form fibrils, by incubating the molecules under physiological buffer and at a protein concentration of 1 mg/ml. They show to be able to form fibrils, that are positive to Thioflavin T binding assay. Moreover, analysis of the structure of fibrils, conducted using circular dichroism (CD) and Fourier Transformed-IR (FTIR) spectroscopy, detects the structural transition from random to β-sheet structure as attended for typical amyloid structure. Fibrils morphology was investigated by transmission electron microscopy (TEM) and atomic force microscopy (AFM) imaging. Fibrils derived from aS dimers are quite long, unbranched and formed by a single filament, a peculiar difference with aS fibril morphology. To identify which amino acids in the respective types of fibrils belong to the fibril core, proteolysis was performed. The rationale of this experiment reside in the fact that disordered regions of proteins are generally site of enzymatic attack and hydrolysis occurs at flexible chain region devoid of hydrogen-bonded secondary structure. Therefore, the prospects are to remove the flexible parts or tail from the amyloid core. Results showed that the core structures of the fibrils of the different molecules seems to be constituted by the same amino acidic region, which encompasses the segment 35-96, in analogy with previous studies. The kinetic of the process was analyzed by fluorescence techniques (ThT binding assay) and by evaluating the amount of protein present in fibrils on time. This calculation was indirectly performed measuring the absorbance of the supernatant obtained after centrifugation of each aliquot. NN and NC dimers show a slower kinetic of fibrillation than aS, while the rate of fibril formation of CC and DC dimers is faster than aS. Moreover, aggregation experiments on mixtures of aS in the presence of small amount of dimers were also conducted in order to check if the presence of dimer influence aS kinetic. Results evidenced the ability of CC dimer to affect the aggregation of aS. On the base of collected results, models of the dimer conformation within the fibrils are proposed.

The research experience performed at Imperial College London gave me the possibility to learn and apply advanced techniques in mass spectrometry analysis of small organic compound, using GC-MS and MALDI-TOF spectrometers.

N-acetylglucosaminyltransferase V (GlcNacT-V), encoded by the *Mgat5* gene, is a medial Golgi enzyme which catalyzes the addiction of a β-1,6-linked GlcNAc to the α-1,6 mannose of the trimannosyl N-glycan core. GlcNacT-V plays a pivotal role in the formation of tri- and tetra-antennary N-glycans on newly synthesized glycoprotein. This branch provides the preferred substrate for the enzymatic subsequent synthesis of polylactosamine chains and terminal modification including the Lewis antigens. In my

study, glycomic analyses were performed to investigate possible changes in protein N-glycosylation in wild type conditions and in the absence of *Mgat5* gene in C57B5 mice kidneys. In parallel, N-glycan profile of kidneys and spleens coming from mice treated with high fat diet GlcNAc supplementation were analyzed. Previous results demonstrate that the effects of GlcNAc salvage appear to increase flux to UDP-GlcNAc. Therefore we were interested to know whether this implementation affects N-glycan branching.

Results show that *Mgat5* deficient mouse kidney display less amount of tri-antennary and tetra-antennary structure compared to controls. However, GlcNAc dietary salvage has no apparent effect on N-linked glycosylation in the kidney and spleen, even if the experiments conducted on cell lines demonstrate that increased influx of UDP-GlcNAc resulted on increased N-glycan branching. Moreover, the performance of optimized glycome procedure allowed the identification of more tri-antennary glycan structures than the one reported on CFG (Consortium of Functional Glycomics) database.

Riassunto

La mia tesi di dottorato è composta di due sezioni. Una sezione riguarda la caratterizzazione di dimeri di α-sinucleina (aS) in confronto con le proprietà di aS, sia in soluzione che in esperimenti di aggregazione. Il lavoro sperimentale è stato condotto nel laboratorio di Chimica delle Proteine (CRIBI Biotechnology Center), presso l'Università degli studi di Padova, e costituisce il progetto principale nel quale sono stata coinvolta. Durante il mio terzo anno di dottorato ho trascorso 6 mesi al laboratorio Biopolymer Mass Spectrometry Laboratory presso l'Imperial College a Londra. In questo laboratorio sono stata coinvolta in due progetti: uno studio di analisi glicomica di tessuti murini e un progetto pilota sulla biosintesi di emicellulosa mixed linked glucans (MLG).

Il morbo di Parkinson è una malattia neurodegenerativa progressiva caratterizzata dalla perdita di neuroni dopaminergici nella *substantia nigra*. La principale caratteristica istologica della malattia è la presenza di inclusioni intracellulari, conosciute come corpi di Lewy, composti da aggregati proteici filamentosi. La patogenesi della malattia è ancora poco chiara, ma un passaggio chiave nello sviluppo della malattia è l'aggregazione di α-synuclein (aS) in fibrille amiloidi, che si accumulano dei corpi di Lewy e ne costituiscono il componente principale. Nonostante la sua importanza nella neurodegenerazione, si conoscono poco la funzione di aS, il suo stato nativo fisiologico e il meccanismo di aggregazione. aS è stata di recente descritta come un tetramero di proteine in α-elica, ma aS è stata generalmente descritta come una proteina *natively unfolded*. aS assume conformazione ad α-elica a seguito di interazione con lipidi e converte a struttura β durante i processi patologici. Durante il processo di aggregazione, aS forma oligomeri solubili di struttura β, transienti intermedi tra la forma fisiologica di aS e le fibrille amiloidi.

La dimerizzazione di aS può rappresentare un fattore limitante nell'aggregazione e nella formazione di struttura amiloide. Pertanto, abbiamo deciso di studiare l'aggregazione di diversi dimeri di aS, prodotti mediante biologia molecolare. E' stato aggiunto un residuo di cisteina all' N- o al C- terminale di aS, producendo quindi dimeri NN o CC, legati attraverso un legame disolfuro. Un dimero NC, formato da due molecole consecutive di aS, è stato ottenuto come singola catena polipeptidica. Durante il progetto è stato prodotto un altro dimero, chiamato DC, disegnato in modo da avvicinare ulteriormente le regioni idrofobiche di aS, ed evitare le interferenze provocate dalle catene laterali, che vengono a trovarsi all'interno della molecola nei dimeri NN, CC ed NC. Il dimero DC contiene i residui 1-104 uniti al segmento 29-140 di aS, ed è quindi costituito da due regioni centrali di aS, altamente amilodoigeniche, disposte in modo consecutivo. I dimeri rappresentano uno strumento adatto per lo studio delle interazioni intramolecolari di aS. Alcune differenze sostanziali definiscono e

limitano la libertà di movimento dei dimeri rispetto ad aS, ipoteticamente differenziando il processo di fibrillazione delle cinque strutture proteiche.

La caratterizzazione dei dimeri è stata effettuata utilizzando tecniche biofisiche e chimiche al fine di definire il loro comportamento in soluzione come monomero. Studi di dicroismo circolare (CD), spettroscopia IR ed NMR hanno dimostrato che tutti i dimeri sono *unfolded*. Tutti effettuano transizione ad α-elica a seguito dell'interazione con il detergente SDS. Questi risultati provano che i dimeri hanno caratteristiche conformazionali simili ad aS.

Successivamente, è stata esaminata la capacità dei dimeri di formare fibrille, incubando le molecole in tampone fisiologico alla concentrazione di 1 mg/ml. Tiutti sono in grado di formare fibrille, che sono positive al saggio di legame alla Tioflavina T (ThT), generalmente utilizzato per determinare la presenza di struttura amiloide. Inoltre, le analisi della struttura delle fibrille, condotte usando CD e spettroscopia IR in trasformata di Fourier (FT-IR), rilevano la presenza di transizione strutturale da *random* a struttura β, come ci si aspetta per fibrille amiloidi. La morfologia delle fibrille è stata studiata mediante microscopia elettronica a trasmissione (TEM) e microscopia di forza atomica (AFM). Le fibrille derivate dai dimeri di aS sono abbastanza lunghe, non ramificate e a singolo filamento, una differenza peculiare rispetto alle fibrille di aS, che si presentano *twisted* e formate da più filamenti. Per identificare quali amminoacidi di ciascun dimero fosse coinvolto nel *core* fibrillare sono sati eseguiti esperimenti di proteolisi. Il razionale di questo esperimento risiede nel fatto che le regioni non strutturate delle proteine sono in genere sito di attacco enzimatico, e l'idrolisi si verifica quindi in regioni flessibili, sprovviste di legani idrogeno intermolecolari che stabilizzano una struttura secondaria. Quindi lo scopo dell'esperimento è di rimuovere le parti flessibili dal *core* amyloide. I risultati hanno mostrato come le strutture *core* delle fibrille dei diversi dimeri sembrino essere costituite dalla stessa regione amminoacidica, che comprende il segmento 35-96, in analogia con studi precedenti su aS. La cinetica del processo è stata analizzata con tecniche di fluorescenza (saggio ThT) e valutando la quantità di proteine presenti nel tempo. Questo calcolo è stato effettuato indirettamente misurando l'assorbanza del surnatante ottenuto dopo ultracentrifugazione delle aliquote prelevate da miscele di aggregazione a diversi tempi. I dimeri NN ed NC hanno mostrato una cinetica di aggregazione più lenta rispetto ad aS, mentre il tasso di formazione delle fibrille di CC e DC è più veloce. Inoltre, esperimenti di aggregazione su miscele di aS in presenza di piccole quantità di dimeri sono stati condotti al fine di verificare se la presenza del dimero influenzasse la cinetica di aS. I risultati hanno evidenziato la capacità del dimero CC di influenzare l'aggregazione di aS. Sulla base dei risultati ottenuti, sono stati proposti dei modelli sulla conformazione dei dimeri all'interno delle fibrille.

L'esperienza di ricerca svolta all'Imperial College London mi ha dato la possibilità di imparare e applicare tecniche avanzate di spettrometria di massa (MS) sull'analisi di

composti organici, utilizzando gas cromatografia accoppiata ad MS (GC-MS) e spettrometri MALDI-TOF.

L'enzima N-acetylglucosaminyltransferase V (GlcNAc-V), codificato dal gene *Mgat 5*, è un enzima del Golgi che catalizza l'addizione di un GlcNAc in posizione β-1,6 a un mannosio α-1,6 della struttura di base degli zuccheri legati a residui amminici (N-glicani). GlcNAc-V svolge un ruolo fondamentale nella formazione di N-glicani a tre- e quattro-antenne su una proteina appena glicosilata. Queste ramificazioni forniscono il substrato favorito per la successiva sintesi enzimatica di catene poli-lactosamminiche e per le modificazioni terminali, compresi gli antigeni di Lewis. Ho svolto analisi glicomiche su tessuti renali murini per studiare possibili cambiamenti nella N-glicosilazione in topi *wild type* e *knock out* per il gene *Mgat 5*. In parallelo, è stato analizzato il profilo glicomico di tessuti renali e di milza di topi alimentati con una dieta ricca di GlcNAc. Risultati precedenti avevano dimostrato un aumento nel flussio di UDP-GlcNAc (substrato di GlcNAc-V), perciò eravamo interessati a determinare se il maggione apporto di zucchero influenzasse le glicosilazioni proteiche.

I risultati hanno evidenziato come le glicoproteine dei topi *ko* per *Mgat 5* hanno meno strutture a tre- e quattro-antenne nelle glicosilazioni rispetto ai controlli. L'apporto di GlcNAc nella dieta non ha alcun affetto apparente sulla struttura e composizione delle glicosilazioni dei tessuti analizzati, nonostante precedenti esperimenti condotti su linee cellulari abbiano avuto un diverso esito. Inoltre, le analisi che ho condotto hanno permesso di identificare glicosilazioni non ancora registrate nel database CFG (Consortium of Functional Glycomics) per i tessuti analizzati.

1. Introduction

1.1 Parkinson's disease: historical remarks and general hallmarks

Parkinson's disease (PD) gets its name from the scientist that first formally described this disorder, in 1817. In his essay, James Parkinson systematically listed the symptoms and medical history of six individuals, referring to the disease as *Shaking palsy*, and encouraging further studies. Roughly fifty years later, the French Jean-Martin Charcot, father of modern neurology, investigated the pathology and assigned to the British scientist his place in medical history by giving the syndrome the name that is currently in use (Kempster et al., 2007).

PD is second only to Alzheimer's disease as common neurodegenerative disorder. The average age of onset is about 60, and incidence increase with advancing age, from 17.4 in 100,000 person years between 50 and 59 years of age to 93.1 in 100,000 person years between 70 and 79 years (Lees et al., 2009). The majority of cases of PD are idiopathic and sporadic, but there is a genetic component in about 10% of cases, showing both recessive and dominant modes of inheritance, and an earlier onset of the disease. Clinically, PD is characterized by motor symptoms such as unilateral rest tremor, slowness of movement (bradykinesia), muscular rigidity, inability to move (akinesia) and postural instability. PD patients often show reduced facial expressions and speak in a soft voice. The disease can be associated to depression, personality changes, dementia, sleep disturbances, speech impairments. The severity of Parkinson's symptoms tends to worsen over time.

A century after the publication of Parkinson's observations, a German neurologist, Friederick H. Lewy, gave his key contributes to the understanding of the disease, investigating the histological features of sixty PD patients' brains. Curiously, Dr. Lewy was working in the same University department of Dr. Alzheimer (who first described Alzheimer's pathology). In 1912, the scientist discovered the intraneural inclusion bodies and neurites that later came to bear his name (Rodrigues, 2010). Lewy bodies (LBs) are intracytoplasmatic inclusions identifiable by standard histological methods as large, spherical and highly eosinophilic inclusions (fig. 1.1). LBs represent the cardinal hallmark of PD pathology, and have been considered to be a marker for neuronal degeneration, because neuronal loss is found in the predilection sites for LBs (Wakabayashi, 2007; Lees, 2009).

Actually, since patients suffering from other neurological disorders can display parkinsonian features, a definitive diagnosis of Parkinson's disease can be confirmed only by *post mortem* histological identification of LBs presence in neurons, and by the examination of the *substantia nigra pars compacta* (SN) for the loss of pigmented neurons.

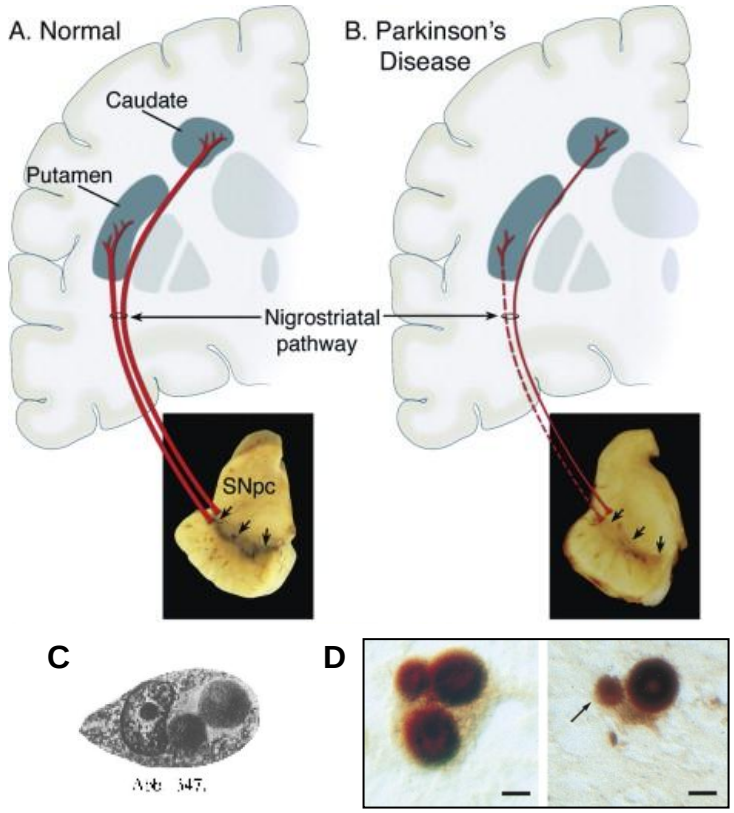


Fig. 1.1. Pathological hallmarks of PD. A schematic comparison of coronal brain slices from a control subject **(A)** and a patient with PD **(B)** illustrates the major neurodegenerative loss of dopamine-synthesizing neurons in the *substantia nigra pars compacta*, projecting to striatal nuclei (caudate and putamen) in the cerebrum (reprinted from Dauer et al. 2003). **(C)** Draw of LBs by Dr. Lewy (reprinted from Lewy 1923). **(D)** Nerve cell with three and two LBs those are double-stained for α-synuclein and ubiquitin. The halo of each LB is strongly immunoreactive for ubiquitin, whereas both the core and the halo of each Lewy body are immunoreactive for α-synuclein (Bar = 10 µm) (reprinted from Spillantini et al. 1998).

The SN constitutes a key regulatory nucleus of basal ganglia circuitry. Neuromelanin, the black pigment that gives its name to SN, is a product of the metabolic pathway for dopamine synthesis. When the symptoms of PD first become apparent, more than 70% of the dopamine-containing neurons have already been lost, hence turning the tissue less black (Irvine et al., 2008).

The role of dopamine (DA) as neurotransmitter was defined in the late 1950s by the Swedish scientist Arvid Carlsson, who deserved for this reason the Nobel Prize in 2000. His studies were rapidly converted into clinical investigations and resulted within a few years in the first clinical treatment of PD, a therapy that is still in wide use today (Andersen, 2009). DA neurons operate in a pathway that controls voluntary movement. This involves signals being relayed from the cerebral cortex through the basal ganglia back to the cortex and then on to muscles. Neurons from the SN project axons that release dopamine in synapses on interneurons in the striatum (composed by caudate nucleus and putamen). As the DA-containing neurons die, failure to complete this circuit results in inability to coordinate movement. Pharmacologic interventions to restore

striatal dopaminergic neurotransmission are accomplished by DA agonists, compounds that directly stimulate postsynaptic receptors. Levodopa (3,4-dihydroxy-L-phenylalanine), a naturally occurring amino acid, is an intermediate in the pathway of dopamine synthesis. After oral ingestion, levodopa is actively transported from the upper small intestine into the circulation by a mechanism specific for large, neutral L-amino acids. Because of ongoing metabolism and the distribution of levodopa throughout the body, only a small fraction of the drug reaches the brain after active transport across the blood–brain barrier. Once there, DA is rapidly formed from levodopa by aromatic L-amino acid decarboxylase. The coadministration of other drugs can improve the efficacy of Levodopa (Thomas and Beal, 2007).

Many other regions of the brain are affected in Parkinson's disease, and indeed in the early stages it may affect only a lower region of the brain stem called the *medulla oblongata*, spreading gradually upward through the basal ganglia into the cortical areas (Braak et al., 2003; Irvine et al., 2008).

1.2 Lewy bodies

LBs are hallmark lesions of degenerating neurons in PD and dementia with Lewy bodies (DLB) disease. DLB and PD are supposed to have common molecular pathologic pathways and differ for the location of LBs and the dementia appearance. When dementia develops prior to Parkinsonism or during the first year of disease, the diagnosed pathology is DLB (Johansen et al., 2010).

LBs and Lewy neurites have been described both in SN and in many other brain regions, such as the locus coeruleus, the dorsal motor nucleus of the vagus, the nucleus basalis of Meynert, thalamus, amygdale, olfactory system, periaqueductal grey matter and others (Lees, 2009). Cortical LBs have been correlated to dementia. Comparing LBs from different brain's area, they result heterogeneous both in morphology and in composition. In generally, brain-stem LBs appear as intracytoplasmic inclusions, 5-25 μm in diameter, with a dense eosinophilic core and a clearer surrounding halo, while cortical LBs and Lewy neurites are smaller and lack of halo. Ultrastructurally, they are composed of a dense core of filamentous and granular material, surrounded by looser radial fibrillar material. Ubiquitin was the first protein to be found in LBs. The presynaptic neuronal protein α-Synuclein (aS) was identified only recently and it is now regarded as the major protein constituent of LBs (Spillantini et al., 2007; Shults et al., 2006). aS antibodies strongly stain LBs and Lewy neurites. Because of this sensitivity, aS staining is now habitually used than eosin or ubiquitin staining for these structures. Biochemical analyses have shown that aS is part of the fibrillar structures of LBs. The deposited, pathological forms of aS are and insoluble aggregates, and may carry post-translational modifications such as truncation, nitration, ubiquitylation and phosphorylation (Cookson, 2009).

In addition to ubiquitin and aS, several studies identified more than 70 other molecules by mass spectrometry, present in smaller amount, and including several involved in protein folding, membrane trafficking and oxidative stress (Lees, 2009). Wakabayashi and colleagues listed and classified these proteins in ten groups: structural elements, among which aS; aS binding protein; synphilin-1-binding protein; components of ubiquitin-proteasome system; proteins implicated in cellular responses; proteins associated with phosphorylation and signal transduction; cytoskeletal proteins; cell cycle proteins; cytosolic proteins that passively diffuse into LBs; and others (see table 1.1). Furthermore, lipids represent a significant component of LBs. They are located in the core of brain-stem LBs and diffusely distributed in the cortical ones (Gay et al., 2000). Lipids have been proved to play a crucial role in the oxidation and aggregation of aS. The identity of the lipids in LBs is still poorly characterized, although it has been hypothesized that some could derived from degraded membranous organelles, including mitochondria (Issidorides et al., 1991).

LBs inclusions have been for long considered to be responsible for neuronal loss, because they were found predominantly in SN and locus ceruleus, the brain regions most affected by neuronal depletion. Moreover, patients with low amount of

Table 1.1 Protein components of LBs (Wakabayashi et al., 2007)

Group 1. Structural components of LBs fibrils α-synuclein (aS); neuronal filament protein.	**Group 2. α-Synuclein binding protein** Agrin, 14-3-3; microtubule-associated protein (MAP) 1B; synphilin-1; tau
Group 3. ynphilin-1-binding proteins aS; parkin, dorfin and SIAH-1 (ubiquitin ligases); NUB1 (down regulator of ubiquitin-like protein); prolyl-isomerase Pin 1.	**Group 4. Components of the ubiquity-proteosome system** ubiquitin, ubiquitin-activating enzyme (E1); ubiquitin-conjugating enzyme (E2); ubiquitin ligases (E3); proteosome subunits; proteosome activators; and ubiquitin-proteosome related proteins.
Group 5. Proteins implicated in cellular responses molecular chaperons; interferon induced protein ; proteins involved in glycosilation; oxidative stress and cell stress.	**Group 6. Proteins associated with phosphorilation and signal transduction** kinases; proteins and enzymes associated with signal tranduction.
Group 7.Cytoskeletal proteins microtubule-associated proteins; neurofilament; tubulin; tubulin polymerization promoting protein.	**Group 8. Cell cycle proteins** Cyclin B; retinoblastoma protein.
Group 9. Cytosolic proteins amyloid precursor protein; calbindin; chromogranin A; synaptophysin; tyrosine hydroxylase; vescicular monoamine trasporter 2.	**Group 10. Others** α2-macroglobulin; IgG; gelsolin-related amyloid protein Finnish type; synaptotagmin X; tissue transglutaminase; oxigenases.

LBs have less content of surviving neurons, compared to patients with moderate amount of LBs, indicating that neurons that contain LBs are dying.

However, that LBs are related to neuronal loss does not imply that the inclusions are the cause of the cell death (Cookson, 2009). Evidences that LBs could be present in brain regions not affected by cellular death (e.g. the neocortex), and the presence of LBs in elder individuals who have died without signs of PD, support other views. LBs alone are insufficient to cause neuronal death, and their formation could be peripheral to primary pathological process. Alternatively, they could represent a failed effort to protect the damaged neuron from toxic proteins species (Lees, 2009). Recent reports suggest that intermediate species of aggregated aS are cytotoxic, and that aS fibrils, of which LBs are constituted, could represent a cytoprotective mechanism in PD. From this view, LBs represent a response to sequester of the toxic proteins, and it would also explain the presence of chaperonins, ubiquitin and proteosome-associated proteins in LBs.

1.3 Pathogenesis

PD is a multifactor disease. Currently, most of the diagnosed cases are described as idiopathic (having no specific cause), but during the last decades several investigations proved that both environmental factors and genetic susceptibilities are associated with PD pathogenesis.

The hypothesis that environmental factors play a role in the aetiology of PD were done after the discovery of a group of intravenous drug users who unwittingly injected a synthetic analogue of Demerol contaminated with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) causing an acute, permanent parkinsonian state that was levodopa responsive (Langston et al, 1985). MPTP is able to enter the hematoencephalic membrane, and once inside the brain, is converted by astrocytes into an active metabolite, MPP^+ (1-methyl-4-phenylpyrdinium), which then can enter dopamine neurons and exert its toxicity. The ability of MPTP to reproduce so many of the features of PD intensified the search for potential environmental toxicants that might contribute to the development the disease (Hatcher et al., 2008).

Epidemiological studies suggested that exposure to pesticides may be a risk factor. The herbicide paraquat shows structural similarities to MPP^+ and may confer an increased risk for PD, especially if the exposure is combined with other pesticides, such as ziram and maneb (Wang et al., 2011). Rotenone, a plant derivative used as an insecticide, is an inhibitor of NADH dehydrogenase. Chronic infusion of either rotenone or MPTP in rodents results in Parkinsonism-like behaviour and pathology, including the formation of inclusion bodies, and indeed these are among the best animal models for the human disease (Dauer et al., 2003).

The majority of PD cases are sporadic, but the discovery of genes linked to rare familial forms of PD have confirmed the role of genetics in development of PD, and

provided important information in the understanding of molecular pathogenesis pathways implied in common sporadic disease (Thomas & Beal, 2007).

1.3.1 Genetic components in PD and related disorders

Till now, linkage studies have identified sixteen loci, named PARK, correlated with PD or related disorders (Parkinsonism syndromes and LBD) (table 1.2). The loci include two autosomal dominant genes, α-synuclein (SNCA) and leucine-rich repeat kinase 2 (LRRK2), and four autosomal recessive genes, parkin, DJ-1, PTEN-induced putative kinase 1 (PINK1) and a lysosome ATPase type (ATP13A2). Identification of other Mendelian forms of PD still remains a main challenge in PD research.

PARK1 and PARK4 were initially assigned to different regions on chromosome 4, but later ascribed to the same locus. Four loci are still lacking for accurate gene identification (PARK3, -10, -12 and -16) (PDGene database data). The significance for the disease of PARK5 locus has been debated by epidemiologic studies, and is know considerate to be not involved as PD genetic factor (discussed below in this section). The other PARK loci have been described only recently and require further epidemiologic studies. Moreover, other genes, coming from different loci, have been linked to PD and Parkinsonism, such as GBA (glucocerebroside), MAPT (microtubule associated protein tau), spatacsin, ataxin 3 and actasin 2 (Hardy, 2010).

Table 1.2 PARK loci associated with PK or related disorders (PDGene database data)

Chromosome	PARK Locus	Gene (protein)
4q21-q23	PARK1	SNCA (α-synuclein)
6q25.2-q27	PARK2	Parkin
2p13	PARK3¤	Unknown
4p14-16.3	PARK4	SNCA (α-synuclein)
4p14	PARK5¤	UCH-L1 (ubiquitin C-term esterase L1)
1p35-p36	PARK6	PINK1 (PTEN-induced putative kinase 1)
1p36.33 - p36.12	PARK7	DJ-1
12p11.23-q13.11	PARK8	LRRK2 (leucine-rich repeat kinase 2)
1p36	PARK9	ATP13A2 (ATPase type 13A2)
1p32	PARK10¤	Unknown
2q36-q37	PARK11¤	GIGYF2 (GRB10 interacting GYF protein 2)
Xq21-q25	PARK12¤	Unknown
2p12	PARK13¤	HtrA2 (HtrA peptidase2)
18q11	PARK14¤	PLA2G6 (phospholipase A2, group VI)
22q12-q13	PARK15¤	FBX07 (F-box protein 7)
1q32	PARK16¤	Unknown

Chromosomal location of PARK loci is annotated as defined in the ONIM (Online Mendelian Inheritance in Man) database. ¤ PD-associated loci with unknown relevance.

PARK1: α-Synuclein

The discovery that aS is the main component of LBs is subsequent to the recognition of a mutation linked to PD in the *SNCA* gene (Polymeropoulos et al., 1997). Three autosomal dominant point mutations of this gene have now been established to segregate with familial PD, and result in A53T, A30P and E46K substitutions (Krueger et al. 1998, Zarranz et al. 2003). A53T was the first mutation to be identified in large Greek kindred. PD patients with this mutation display earlier onset of the disease (about 7-10 years earlier), and much lower prevalence of tremor compared with patients with sporadic PD (Papapetropuolus et al., 2001). A curious aspect of the mutation is that threonine is already present at this position in rodents and other species (Hamilton, 2003). PD patients with A30P and E46K mutation display as well earlier onset of the disease. Moreover, also duplications and triplications of PARK1 locus cause autosomal dominant, early onset PD (Singleton et al., 2003; Chartier-Harlin et al., 2004). The age of onset and severity of the disease phenotype seems to correlate with SNCA copy number, suggesting a gene-dosage effect (Cookson, 2005). PD patients who carry duplications, which generate three copies of the gene, tend to have PD which develops slowly from the 40s. Locus triplications, which produce four copies of the SNCA gene, are responsible of earlier onset disease (mid 20-mid 30) (Farrer et al., 2004; Wood-Kaczmar et al. 2006).

aS belongs to synucleins family, which comprises also α-synuclein, α-synuclein (also named persyn), and synoretin (Clayton et al., 1998; Surguchov et al. 1999). Synucleins are small proteins (between 113-143 amino acids), natively unfolded, highly evolutionary conserved and abundantly expressed in nervous system of vertebrates. aS is expressed ubiquitously but heterogeneously in the brain, with highest levels of protein reported in deeper layers of the cerebral neocortex, the hippocampus and the SN. The protein is expressed within glia and neurons where it is particularly abundant at presynaptic terminals (Iwai et al., 1995; Mori et al., 2002). The precise physiological function of aS has yet to be established, although different studies evidenced several putative tasks of aS, such as play a role in synaptic plasticity and in regulation of dopamine neurotransmission. Other studies have evidenced that aS may function as a chaperone protein, based on its abundance in cytosol, its natively unfolded structure, and its prevention of protein aggregation. The physiological function of aS is also related with lipids and membrane since this protein seems to modulate presynaptic vesicle pool size and vesicle recycling.

The pathological role of aS in PD seems to be linked to its aggregation properties. Fibrils of aS are the principal constituent of LBs and Lewy neuritis in sporadic and in familial PD. Moreover, genomic multiplications of the SNCA gene were shown to increase the deposition of soluble aS into insoluble aggregates (Miller et al., 2004). The fibrils observed in LBs are structurally similar to those found in amyloid diseases and are linear rods of 5-10 nm diameters (Fink, 2006). The aggregation pathway of aS can be reproduced *in vitro*: when monomeric aS is incubated at 37 °C, pH 7.4, it forms fibrils

with a condition-dependent rate and agitation can significantly accelerate the process. *In vitro* produced aS fibrils are reminiscent of those observed in LBs (Fink, 2006; Wood-Kaczmar et al., 2006).

In its aggregation pathway, aS forms soluble intermediates, such as oligomers and protofibrils, that are considerate to exert toxicity within the neurons. Indeed, soluble oligomers and protofibrils of aS are toxic to some cell cultures, including the dopaminergic human neuroblastoma and neuroglioma cell line (El-Agnaf et al., 1998; Outerio et al. 2008). Both A53T and A30P mutants show an increased propensity to form these so called protofibrils, which can permeabilize vesicles in a way that is reminiscent of bacterial pore-forming toxins (Volles et al., 2001). Transgenic mice that overexpress wild-type or mutant aS have given conflicting results (Fleming et al., 2005). However, models in *Drosophilia melanogaster* demonstrated that high levels of aS cause abnormal protein aggregation and neurotoxicity in DA neurons.

A deeper discussion about structure, functions, aggregation pathways and toxicity of aS will be reported in section 2.

PARK2: Parkin

The PARK2 locus codes for *parkin*, the causative gene for some autosomal recessive juvenile form of PD (Kitada et al, 1998). The gene encodes for a protein of 465 amino acids, named as well Parkin, which has an ubiquitin-like domain at the N-terminal and a ring finger motif at the C-terminal. Parkin is an ubiquitin E3-ligase and acts as substrate-recognition molecules during the synthesis and attachment of polyubiquitin chains to proteins that are targeted for degradation by ubiquitin proteasome system (UPS) (Shimura et al., 2000).

Parkin mutations were found in 67% of cases with age of earlier than 20 years onset and in about 8% of cases with an age of onset between 30 and 45 years (West et al., 2004). At least twenty familiar associated Parkin mutations, including nonsense mutation and deletions, have been shown to disrupt differentially the solubility, localization, binding and ubiquitination properties of Parkin *in vitro* (Sriram et al., 2005; Hardy, 2007).

Loss of function of Parkin leads to dopaminergic cell death in PD, acting on two sides: the disposal of toxic protein and mitochondrial damage. One of the most accredited hypotheses of neuronal damage is that Parkin helps the degradation of proteins toxic to dopaminergic neurons through its UPS action. Putative substrates of parkin include synuclein, CDC-rel1 and -rel2 septins, cyclinE, p38 tRNA synthase, Paelr1, synphilin-1, synaptotagmin XI and Parkin itself (Cookson, 2005). Parkin could be linked to synaptic densities by an interaction with the PDZ protein Cask, and this explains why some of its substrates are synaptic protein.

Moreover, studies on knockout models for Parkin highlighted the protective role of the ligase within mitochondria. In addition to structural alteration in dopaminergic neurons, Parkin knockout *Drosophilae* display reduced lifespan, locomoter defects and male sterility. The earliest manifestation of muscle degeneration and spermatogenesis deficit is mitochondrial pathology, which subsequently leads to apoptosis (Green et al., 2002). To explain the mechanism of mitochondrial protection, studies on cell lines with depolarizing agent were performed, and proved that Parkin, which occurs mainly in the cell cytoplasm, re-localizes to damaged mitochondria, where it promotes autophagy of damaged mitochondria (Narendra et al., 2008). The role of Parkin in mitochondria is strongly connected to PINK1, the PTEN-induced kinase (discussed below in this section) (Clark et al., 2006; Abeliovic et al., 2010).

PARK5: UCH-L1

The ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) is one of the most abundant proteins in the brain. It is considerate to cleave polymeric ubiquitin to monomers and to hydrolyse bonds between ubiquitin molecules and small adducts such as glutathione and cellular amines (Leroy et al., 1998). Its involvement in PD has been supported by its presence in LBs, and by genetic studies that linked missense mutations to PD. However, several epidemiological studies have now confuted its role as PD susceptibility gene.

I93M mutation was described in a German family (Leroy et al., 1998). This particular mutation was shown to lead to partial loss of the catalytic activity of UCH-L1, which could lead to aberrations in the proteolytic pathway and aggregation of proteins. However, the following year a study harmed this hypothesis, suggesting that I93M could be a harmless substitution whose occurrence in the family reflected a change co-occurrence with S18Y allele mutation (Lincoln et al., 1999). A study which involved 1,970 individuals later supported S18Y as a mutation related to the disease (Maraganore et al., 2004). Conversely, an epidemiologic study on 3,023 persons confuted this data, strongly sustaining that UCH-L1 gene is not a PD susceptibility gene (Healy et al., 2006).

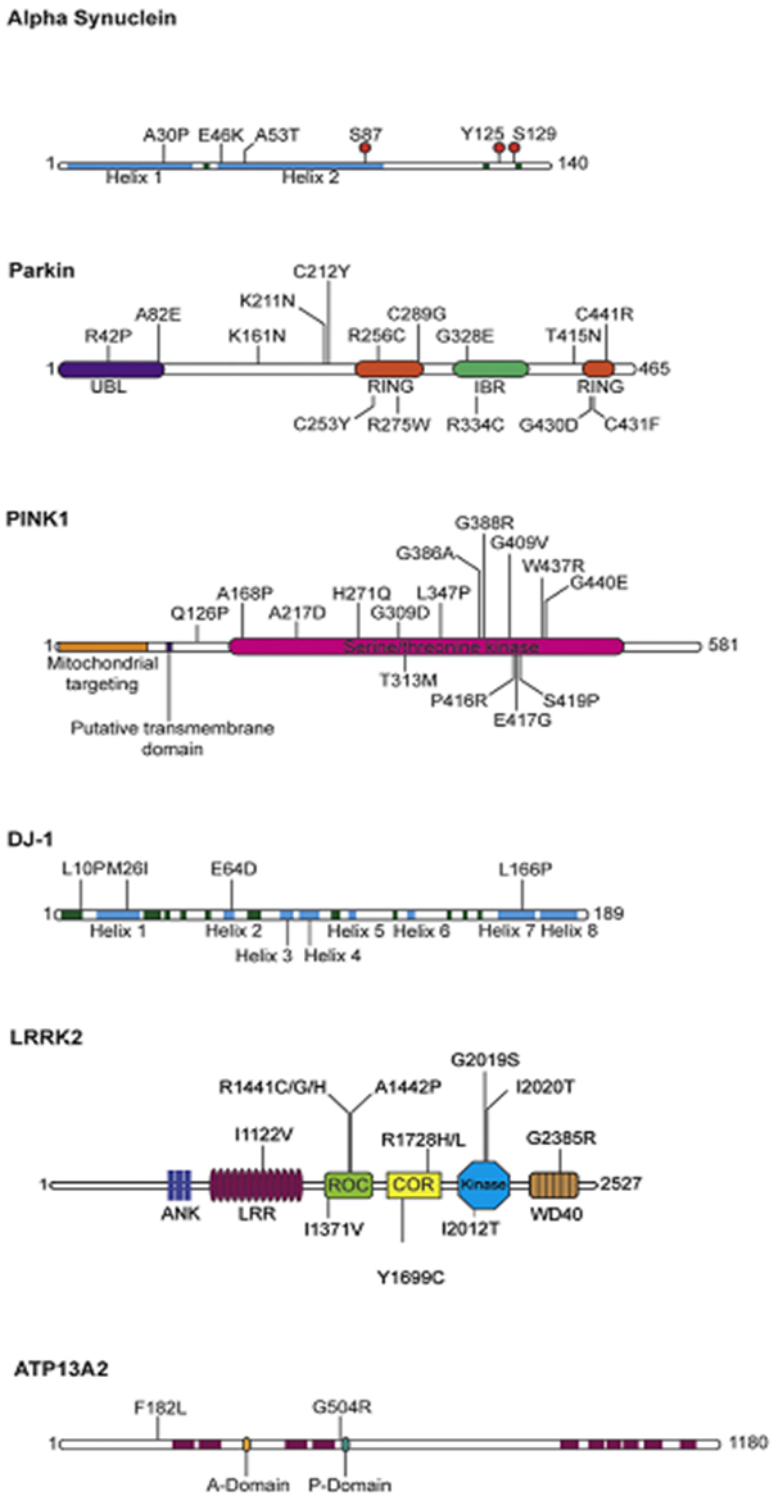


Fig 1.2 Main genes correlated with PD and related disorders. PARK loci 1, 2, 6, 7, 8 and 9, correspondent to alpha-synuclein, PARKIN, PINK1, DJ-1, LRRK2 and ATP13A2 genes are schematically represented. The mutation related to PD and related diseases are reported (reprinted from Hardy 2009).

PARK6: PINK1

PINK1 is the gene responsible for PARK6-associated autosomal recessive Parkinsonism. It codes for an enzyme, PTEN-induced kinase (PINK1), a 581 amino acid protein that contains an N-terminal mitochondrial targeting motif and a highly conserved kinase domain homologous to the serine/threonine kinases of the Ca^{2+}-calmodulin family. PINK1 has a ubiquitous and punctate expression pattern suggesting mitochondrial localization (Valente et al., 2004). Initially, a missense mutation (G309D) and a truncating mutation (W437X) were described, followed later by a number of further point mutations, frameshift and truncation mutations, that lead to the loss of PINK1 function (Ganhdi et al., 2006).

The discovery that mutations in the *PINK1* gene result in Parkinsonism provided a molecular link between mitochondria and neurodegeneration in PD. Mice with mutations in *PINK1* gene display mitochondrial dysfunction (Gautier et al., 2008), as well as *Drosophila* knockout models. In addition, studies on *Drosophilae* evidenced the link between PINK1 and Parkin: knockout model of Parkin display similar phenotypes to loss of PINK1 function (Green et al., 2002), and overexpression of Parkin rescues the mitochondrial dysfunction caused by PINK1 loss (Park et al., 2006), but did not rescue the sensitivity of PINK1 mutant flies to apoptosis induced by stressors. This suggested that PINK1 and Parkin interact to protect mitochondrial integrity, with PINK1 upstream of Parkin (Clark et al., 2006). Even if *in vitro* studies suggested the direct phosphorylation of Parkin by PINK1 (Kim et al., 2008), Parkin does not seem to be the target of PINK1 kinase activity (Narendra et al., 2010). Abeliovich and colleagues recently proposed the following mechanism of interaction in mitochondrial monitoring: in healthy mitochondria, PINK1 is maintained at low levels, regulated through the cleavage by an unidentified protease. During mitochondrial dysfunction, when membrane electrical-potential depolarizes, PINK1 is stabilized at the outer mitochondrial membrane with its kinase domain facing the cytoplasm. Directly or indirectly through an unknown protein, PINK1 then recruits Parkin to the mitochondrial surface, inducing disposal of the damaged organelle through autophagy mechanism (Abeliovic et al., 2010).

PARK7: DJ-1

Rare mutations in the *DJ-1* gene (PARK7 locus) also cause autosomal recessive early onset PD (Bonifati et al., 2003). *DJ-1* mutations account for 1–2% of all early onset PD, with a number of different pathogenic mutations, including exonic deletions, truncations and homozygous and heterozygous point mutations that result in DJ-1 loss of functions. It exhibits significant sequence homology with the PfpI family of intracellular proteases and with ThiJ family of bacterial protein involved in thiamine synthesis. DJ-1 is a highly conserved protein of 189 amino acids with multiple functions,

including antioxidant, transcriptional co-activator and chaperone activity (Olzmann et al., 2004; Thomas & Beal, 2007).

DJ-1 is a homodimeric protein, ubiquitary espressed in a variety of mammalian tissues including brain, and localized in cytosol and mitochondria (Zhang et al., 2005). The precise function of the protein is not clear, but it translocates from cytosol to mitochondria in response to oxidative stress and preventing this process sensitizes the cell to toxic cell damage (Hardy, 2009). Oxidative stress leads to an acidic shift in the DJ-1 isoelectric point by oxidation of Cys 106 which can be converted to cysteine sulfinic acid (Cys-SO_2H), and because of this ability, the protein could work as a scavenger of reactive oxygen species (ROS). DJ-1 is also able to play an indirect antioxidant action stabilizing an antioxidant transcriptional regulator, Nrf2 (nuclear factor erythroid 2-related factor), preventing its inhibition and ubiquitination). Furthermore, it can associates with Parkin during oxidative stress suggesting a common role in neuroprotection (Clements et al., 2006; Thomas & Beal, 2007).

Oxidative cytoplasmic environment also activate DJ-1 as chaperone, and its activity extends to aS. DJ-1 inhibits the generation of aS aggregates in dependence of its oxidation state. Although the oxidation of Cys 106 to sulfinic acid provides for DJ-1 the chaperon activity, further oxidation leads to loss of some secondary structure and to the loss of the ability to prevent aS fibrillation (Shendelman et al., 2004; Zhou et al, 2005).

Furthermore, DJ-1 plays a role in the regulation of protein expression, either as RNA binding protein, and interacting with proteins involved in RNA binding and transcription. It has been demonstrated that DJ-1 is able to bind RNA, and it is involved in the regulation of multiple pathways (van der Brug et al., 2008). Of particular significance to dopaminergic neuronal function is the ability of DJ-1 to transcriptionally upregulate the tyrosine hydroxylase (TH) expression, by inhibiting the sumoylation of a nuclear protein. TH is the enzyme responsible for catalyzing the conversion of tyrosine to DOPA, the precursor of dopamine. The transcriptional dysregulation caused by DJ-1 inactivation contributes therefore to the impairment of dopamine synthesis (Zhong et al., 2006).

PARK8: LRRK2

Mutations in the *LRRK2* gene cause autosomal dominant PD (Paisan-Ruiz, 2004; Zimprich et al., 2004). At least 20 mutations have been linked to Parkinsonism, that is clinically indistinguishable from typical, idiopathic, late onset PD (Mata et al., 2006). The gene contains 51 exons and code for an unusually large protein (2.527 amino acids), named LRRK2 or dardarin.

The protein is conserved among the vertebrate subphylum, and comprises several independent domains including a leucine rich repeat (LRR) domain, a Ras of complex protein (Roc) GTPase domain followed by its associated COR (C terminal of Roc) domain, a kinase domain of the tyrosine kinase-like subfamily, and a WD40 repeat (rich

in tryptophan and aspartic acid) domain. The presence of multifunctional domains and the incidence of diverse pathologies related to LRKK2 mutants suggest the involvement of the GTPase/kinase enzyme in a variety of cellular processes. The multiple functions and interactions pathways of LRRK2 are not yet clear, but the enzyme has been suggested to play a role in the control and maintenance of neurite length, in vesicle endocytosis and sorting between axons and dendrites, in activation of apoptosis through interaction with death adaptor Fas-associated protein, and in controlling protein translation. Several groups also suggest a role in cytoskeleton dynamics (Greggio et al., 2011).

PD-associated mutations can occur in several of the protein domains. *In vitro* experiments have shown that mutations in the kinase domain lead to toxicity through a gain of kinase function (Greggio et al., 2006). Noteworthy, as yet neither the complete panorama of neither the upstream activators nor the downstream targets of LRRK2 is known. Mutations in the GTPase domain are likely to bring to loss of protein interaction, and decreased ability to bind GTP (West et al., 2005). Three putatively pathogenic amino acid substitutions lie within the LRR domain, and two in the WD40 domain. Since the amino acids involved are located towards the surface, they are thought to interfere with protein binding (Mata et al., 2006).

Overexpression of wild type LRRK2, but not of LRRK2 mutant protect from mitochondrial toxins (Saha et al., 2009), suggesting a pathological mechanism that involves mitochondrial damage. LRRK2 knockout flies and mice gave variable results and do not display neurodegeneration. In parallel, some mice that express mutant LRRK2 are reported to display neuronal abnormalities (Cookson and Badmann, 2010).

PARK9: ATP13A2

The genetic cause of a peculiar type of hereditary Parkinsonism with dementia is the presence of mutations in *ATP13A2* gene, which code for a lysosomal ATPase (Ramirez et al., 2006). Even if the gene was assigned only recently, PARK9 locus was localized more than a decade ago, and related to Kufor-Rakeb syndrome, a disease characterized by Parkinsonism with pyramidal degeneration and dementia (Hampshire et al., 2001). Compared to Parkinson's disease, Kufor-Rakeb syndrome has additional distinctive symptoms; including supranuclear vertical gaze plasy, myoclonic jerks, pyramidal signs and cognitive impairment (Paisan-Ruiz et al, 2010). Loss of function of the ATPase coded by *ATP13A2* gene is related to an autosomal recessive form of early onset Parkinsonism.

ATP13A2 codes for a protein of the P-type ATPase superfamily, characterized, among the others, by ten-transmembrane domain topology. In their study, Ramirez and colleagues identified two mutations that lead to premature truncations. The protein function is unknown, as well as its substrate specificity. *ATP13A2* is ubiquitously

expressed, with strongest expression in brain, where it is mainly present in cortex, thalamus and SN. Moreover, expression analyses of SN dopaminergic neurons of patients with common, sporadic form of PD revealed tenfold higher *ATP13A2* mRNA levels compared to controls. This result suggests an activation of lysosome as cellular defence mechanism in PD.

The major pathways in cellular protein degradation are the proteosome and the lysosome system. Mutations of *Parkin* cause proteosomal dysfunction and are responsible for neurodegenerative disorders. In parallel, the lysosomal localisation of the ATPase suggests that dysfunctions of lysosome degradation pathway could play a role in the etiology of PD (Ramirez et al., 2006). This hypothesis is consistent with evidences that aS is degraded by both autophagy and the proteosome (Webb et al., 2006).

PARK11: GIGYF2

The *PARK11* locus was identified on chromosome 2q36-q37 by a whole-genome linkage analysis in 2002 (Pankratz et al., 2002), and the located gene was assigned only recently. The *GIGYF2* gene codes for the GYF protein-2. The protein interacts with Grb10 adaptor protein, and both the proteins are of interest for their involvement in insulin signalling (Lautier et al., 2008), which has been associated with PD (Craft and Watson, 2004).

GIGYF2 knockout animal models support the importance of the gene in neurodegeneration and PD pathogenesis. *Gigyf2* null mice do not survive the first 2 post-natal days, but heterozygous *Gigyf2*$^{+/-}$ mice survive to adulthood, exhibiting motor dysfunction and aS positive LB-like inclusions in spinal anterior horn motor neurons, but not in SN (Giovannone et al., 2009). Nevertheless, epidemiologic studies are actually harming its significance as PD susceptibility gene (Zimprich et al., 2009; Li et al. 2010).

PARK13: HtrA2

The interest for mice *Prss25* gene, encoding HtrA2 (also known as Omi) protease, as susceptibility gene for PD rose after the observation of neurodegeneration and parkinsonian phenotype in knockout mice. The HtrA family refers to a group of serine protease that contain at least one PDZ interaction domain. Mammals encode four HtrA proteases, named HtrA1 to HtrA4. After exression HtrA2 is targeted to the mitochondria, but it could be released in the cytosol, where it contributes to apoptosis though both caspase-dependent and -independent pathways (Walle et al., 2008).

HtrA knockout mice suffer loss of neurons in the striatum, resulting in a progressive movement disorder. They display a progressive akinetic, rigid syndrome, lack of coordination, decreased mobility, bended posture and tremor (Martins et al., 2004).

Mutations of the human gene encoding for HtrA2 that result in loss of function were found in German PD patients (Strauss et al., 2005). Nevertheless, recent epidemiologic studies deny their involvement in PD (Ross et al. 2008; Krueger et al., 2011). However, the enzyme could well take part of PD pathogenesis processes, since it is regulated by the PD-associated kinase PINK1 (PARK6). HtrA2 is phosphorilated through a PINK1 downstream pathway, and thus its proteolytic activity modulated by PINK1 in mitochondria stresses (Plun-Favreau et al., 2007).

PARK14: PLA2G6

Mutations in the gene encoding phospholipase A2 group VI (*PLA2G6*) are associated with neurodegenerative disorders. Enzyme defects lead to several phenotypes that include infantile neuroaxonal dystrophy, neurodegeneration with brain iron accumulation, Karak syndrome and dystonia-Parkinsonism (Gregory at al. 2008; Paisan-Ruiz et al., 2008). Affected individuals are L-dopa-responsive and share overlapping features with Kufor-Rakep syndrome, which is caused by *ATP13A2* mutations (Tan et al., 2009).

Phospholipases A2 are involved in reactions that result in the release of arachidonic acid and other fatty acids. PLA2G6 play an important role in cell cycle regulation through membrane phospholipids remodelling. Its activation in SN generates free radical and lipid peroxidation that are known to play a significant role in PD. This data is supported by the fact that PLA2 deficient mice are resistant to MPTP-induced neurotoxicity (Adibhatla and Hatcher, 2008).

PARK15: FBX07

In addiction to *ATP13A2* (PARK9) and *PLA2G6* (PARK14), also *FBXO7* gene (PARK15) is associated to a progressive form of Parkinsonism, which has clinical phenotype similar to PD. *FBXO7* codes for a member of the F-box family, a class of ubiquitary expressed proteins which mediates protein-protein interactions in many different biological processes. In particular, the Fbx7 protein is involved in substrate recognition being a component of a complex of proteins (SKP1-CUL1-F-box E3 ubiquitin-protein ligase) which mediates the phosphorylation-dependent ubiquitination and the subsequent proteosomal degradation of targeted proteins. It also acts as assembly factor for cyclin D-Cdk6 complexes (Larsen and Bendixen, 2011).

Four autosomal recessive mutations of *FBXO7* gene have been identified and results in rapidly progressive Parkinsonism, which resemble Parkin mutation associate phenotype (to note that the two proteins are both involved in the UPS). The disease has early onset and reduced life expenctancy, pyramidal signs and late cognitive problems. Initially it is responsive to Levodopa treatment, but later it is not (Paisan-Ruiz, 2010).

Other genes linked to PD and Parkinsonism syndromes

A main challenge in PD research is the identification of genetic risk factor for the disease. As the genetic studies continue, more and more genes that increase the risk of the disease are found. Recently, glucocerebrosidae, microtubule-associated protein tau, and spactasin were considered PD and Parkinsonism syndromes risk factors, whose relevance have to be confirmed by further epidemiologic studies.

Recent findings highlighted the importance of glucocerebrosidae (*GBA*) mutations as a risk factor for the development of Parkinsonism. Mutations of this enzyme are responsible for Gaucher's disease, a recessive lysosomal storage disorder. Liver damage characterizes the disease, but neurological problems could often be associated, and LBs were detected in some patients at autopsy. GBA catalyzes the breakdown of glucocerebrosides to ceramide and glucose. Starting from the observation that relatives of Gaucher's patients suffer for PD, several epidemiological studies confirm that heterozygous loss of function leads to an increase of risk of PD of more than 5-fold higher (Hardy et al., 2009).

The microtubule-associated protein tau, encoded by the *MAPT* gene, is associated with a range of neurodegenerative disease, called thautopaties, which include among the others Alzheimer disease (AD). Since PD shares some clinical features with the thautopaties, several studies have been conducted on the relationship between *MAPT* variants and PD, and two large case-control studies published in 2007 provides the evidence of a strong correlation. The mechanism by wich *MAPT* variations modulates the risk of PD is however still unclear (Bekris et al, 2010).

Mutations in *SPG11* gene, encoding spatacsin, are related to autosomal recessive spastic paraplegia with thin corpus callosum. The disease could sometimes manifest with complex Parkinsonism features such as resting tremor, akinesia and with either weak or no Levodopa response (Hardy et al., 2010; Paisan-Ruiz et al., 2010).

1.3.2 Intersecting pathways of pathogenesis

PD is a complex disorder with multiple etiological factors. It may depend on a complex set of circumstances that include genetic factors, environmental exposures and loss of cellular protective mechanisms (Wood-Kaczmar et al., 2006; Thomas & Beal, 2007). As seen analyzing its genetic components, PD shares lots of features with other neurological diseases, suggesting common pathogenesis mechanisms. The reasons why dopaminergic neurons are mainly affected remain unknown, and the inclusion formation process is still elusive. Despite that, some key pathogenesis mechanisms of both familial and sporadic PD can be identified with oxidative stress, mitochondrial dysfunction, unimpaired protein phosphorylation, protein misfolding and impairment of ubiquitin proteasome system (UPS) (Huang et al., 2003). These processes are tightly interconnected complement the main pathogenic event, which is aS aggregation.

Actually plenty of evidences underline the prominent role of oxidative stress as cellular damage mechanism in PD. To list, some of them are increased lipid peroxidation, reduced glutathione levels, high concentration of iron and reactive oxygen free radical generation via autocatalytic mechanisms in SN (Huang et al., 2003). Moreover, overexpression of aS and of some mutant form of Parkin induces increased oxidation of lipids, protein and DNA in cell lines (Lee et al., 2001; Hyun et al., 2002). Oxidative stress is closely linked to mitochondrial dysfunction, impairment of UPS, cytotoxicity and protein misfolding. However, it is difficult to determine whether oxidative stress leads to, or is a consequence of, there events (Jenner, 2003). For instance, UPS impairment may cause oxidative stress, but it could also be a result of oxidative damage. Instead MPTP, a Parkinsonism inducer poison, when converted to its metabolite MPP^+, induce the impairment of the mitochondrial respiratory chain and the consequent superoxide formation.

The formation of reactive oxygen and nitrogen species damages cellular component such as lipids, DNA and proteins. Oxidized proteins may not be adequately ubiquitinate or recognized by the proteosome and may accumulate (Jenner, 2003). Lipid peroxidation leads to the production toxic species, such as 4-hydroxynonenal (HNE), detected by immunocytochemistry in SN and cerebrospinal fluid of PD patients (Yoritaka et al., 1996). HNE can induce apopotosis though the activation of caspases, and, as results of its NF-kB inhibitor action, it cause the reduction of glutathione levels and the inhibition of complexes I and II of the mitochondrial respiratory chain (Camandola et al., 2000).

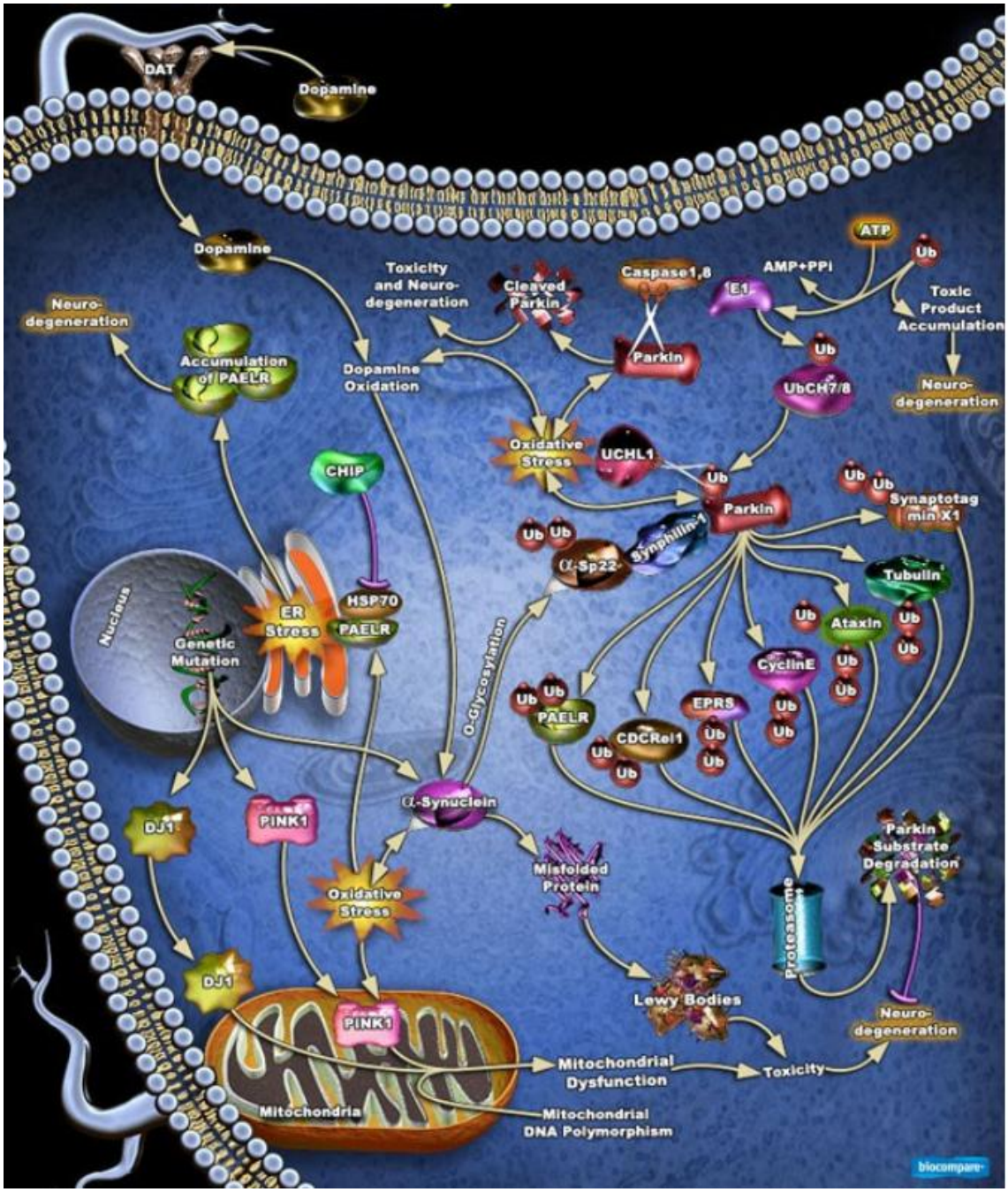


Fig. 1.3 Parkinson's disease pathways. Common intersecting pathways in PD pathogenesis.

To explain how the oxidative stress arises, it is important to note that the metabolism of dopamine (DA) might be responsible for the high basal levels of oxidative stress in SN, and could therefore explain the specific loss of dopaminergic neurons in PD. DA can undergo to autoxidation processes, formation of semiquinone species and polimerization with the production of radical species. Degradation of DA by monoamine oxidase produces hydrogen peroxide (H_2O_2) and increases the formation of oxidized glutathione. Together, these two events can explain the occurrence of oxidative stress and the impairment a major antioxidant system (Jenner, 2003). Moreover, oxidation, nitration, interaction with iron, as well as the formation of DA-aS adducts improve the aggregation of aS (Conway et al., 2001).

Strictly connected to oxidative stress, mitochondria dysfunction plays a central role in PD neuronal death. Mitochondria provide the energy that fuels maintenance, repair and turnover of cellular components. Impairments of the respiratory chain directly result in oxidative stress damage, though the production of reactive oxygen free radicals, which is a function both of the inefficiency of transfer of electrons and of the level of antioxidant defences in the cell (Huang et al., 2003).

Synuclein (aS), Parkin, PINK1, DJ-1, LRRK2 and Htra2 are products of six PARK loci genetically associated to PD that display a degree of localization to the mitochondria under certain conditions. Animal models which show dysfunctions of each of these protein display mitochondria abnormalities (Henchcliffe and Beal, 2008). PINK1 is a serine-threonine kinase which could localize to the outer mitochondrial membrane. In response to membrane depolarization, PINK1 recruits Parkin, an ubiquitin E3 ligase, to the mitochondrial surface, inducing disposal of the damaged organelle through autophagy mechanism (Abeliovic et al., 2010). Another function of Parkin is the promotion of mitochondrial biogenesis by the activation of mitochondrial transcription factor A (Thomas and Beal, 2007). Also DJ-1, which functions as a chaperone to block aS aggregation, relocalizes to the mitochondrial matrix and intermembrane space as a consequence of oxidative stress stimuli, where it exerts a protective function (Henchcliffe and Beal, 2008). Althought LRRK2 is mainly a cytoplasmic kinase, approximately 10% is associated with the outer mitochondrial membrane (West et al., 2005). LRRK2 hyperactive mutant forms cause abnormal protein phosphorylation, which induces mitochondria-dependent cell death (Thomas and Beal, 2007). HtrA2 is a mitochondrial serine protease, and its activity is controlled by PINK1 phosphorilation. Ultimately, aS contains an amino terminal mitochondrial targeting sequence, and acidification of the cytosol or overexpression of aS can cause a relocalization of the protein to mitochondria. aS can also induce oxidative damage and mitochondrial release of cytocrome C (Henchcliffe and Beal, 2008).

To cap it all, any impairment of mitochondrial function in PD may lead to a depletion of ATP and result in decreased processing of ubiquitinated proteins and decreased proteosomal activity (Mounsey and Teismann, 2011). Among the ATP-

dependent process, tubulin folding could be significant in PD. Indeed, mitochondria exert a metabolic control of cellular traffic by the involvement of microtubules. They are dynamic organelles that undergo continual cycles of fission and fusion. Tubulin is the main component of microtubules, and was shown to colocalize with aS in LBs. Microtubule depolymerization induces disruption in axonal transport, which leads to an accumulation of damaged organelles, vescicles and aggregated or misfolded proteins. Furthermore, microtubules regulate autophagosome-lysosome fusion, which can decrease the degradation of aS oligomers. Therefore, perturbation of the microtubule system results in stimulation of aS aggregation and toxicity (Esteves et al., 2011).

Because of the prominent role of mitochondria dysfunction, and the increasing evidence of caspase activation, apoptosis is believed to occur rather than necrosis as cell death mechanism in PD (Olanow and Tatton, 1999). Also excitotoxicity process has been suggested, even if there is no direct evidence of increased excitatory amino acid transmission prior to the onsent of symptomatic PD. Immunologic and inflammatory response to progressive dopaminergic degeneration was eventually reported (Huang et al., 2003).

1.4 α-Synuclein

As discussed previously, aS is a small (140 amino acids) cytoplasmic protein, highly expressed in the central nervous system and concentrated in presynaptic terminals, representing 0.5–1% of the total cytosolic protein in brain (Kruger, 2000). It is the major component of LBs, the cytoplasmic proteinaceous aggregates which characterize PD, and through which the pathology is diagnosed. These filamentous aggregates are composed by fibrillar structures with a cross-β-sheet core typical of amyloid. Mutations or overexpression of aS gene have been associated to early-onset autosomal dominant PD. The mechanisms by which the amyloid structure or other forms of aS aggregates exert their toxicity are not yet clarified. Moreover, it is not known if the fibrillar deposits of aS play a toxic, protective, or no role in the pathogenesis of the disease (Ellis et al., 2005).

1.4.1. Conformational properties

Depending on its biochemical and physical properties, the structure of aS can be divided into tree regions (Fig. 1.4). The N-terminal region, residues 1–60, includes the sites of three familial PD mutations and contains four out of seven imperfect repeats of 11 amino acids that are present in the protein. The repeats are characterized by a highly conserved hexameric motif (KTKEGV). This region is able to assume α-helix structure, interacting with lipids.

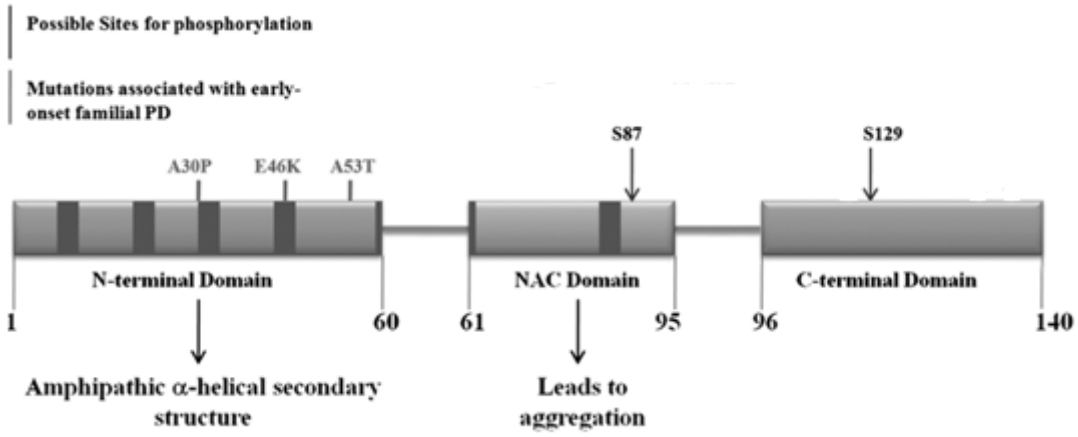


Fig. 1.4 Schematic representation of aS structure. The figure encloses the position of aS mutations and phosphorylation.

The central region, residues 61–95, comprises the highly aggregation-prone NAC sequence, and contains the other three imperfect repeats. This region was found in Aβ amyloid plaque of Alzheimer's disease patients (NAC is the acronym of non amyloid β–component) (Ueda et al., 1993; Weinreb et al. 1996, Goedert, 1997; Jo et al. 2000). The C-terminal region, residues 96–140, is highly dynamic in most conditions. It is highly enriched in acidic residues and prolines, and three highly conserved tyrosine residues are also located in this region.

Since the discovery of aS amyloid fibrils as main component of LBs inclusions, the protein has been deeply examined in a massive number of structural studies. aS has been cloned and produced in bacteria for the first time in 1994 (Jakes et al., 1994). The advantages of using recombinant protein for *in vitro* studies are several: human aS sequence is available to analyses, bypassing labored purification steps from human or animal brains or cell lines (which usually require the use of detergents), and producing high yield of protein with considerable purity grade. Moreover, minor equipment resources are needed to produce recombinant proteins in bacteria, therefore many more laboratories are able to study the protein of interest. Recombinant aS is able to form amyloid fibrils in vitro (Conway et al., 1998).

In vitro studies on recombinant aS show that the aS monomer lacks of ordered secondary structure under physiological conditions, detectable by far-UV circular dichroism (CD), Fourier transform IR (FT-IR) and NMR spectroscopy. For this reason, it is considered an intrinsically disorder, or natively unfolded, protein. In general, one of the main physical characteristic of this set of protein is the combination of low overall hydrophobicity and large net charge. However, aS does not fit this general trend. In its case, N- and C-terminal regions possess charge of opposite sign and are separated by an

extended hydrophobic region (Uversky et al., 2000). Nevertheless, the native state of recombinant aS should be considered as composed by an ensemble of structures that could be, on average, more compact than a random coil (Dedmond et al., 2004). Indeed, small angle X-ray scattering studies (SAXS) show that, at physiological conditions, aS is characterized by a radious of gyration, R_g, of about 40 Å, which is much larger than that predicted for a folded globular protein of 140 residues (15 Å), but significantly smaller than that for a fully unfolded random coil (52 Å) (Uversky et al., 2001). Paramagnetic relaxation enhancement (PRE) NMR spectroscopy gave a molecular characterization of this partial condensation proving that it is due to long-range tertiary interactions. Region 110-130 in the negatively charged C-terminal tail can contact residues 85-95 in the centre of the protein. Within the C-terminal region, residues 120-130 can contact residues 105-115, and the region about residues 120 also can interact with the N-terminus about residues 20 (Bertoncini et al., 2005).

In addition to this set of dynamic structures, natively unfolded, recombinant aS monomer owns an astonishing conformational plasticity and is able to adopt several different and structurally unrelated conformations. From a substantially unfolded native state, *in vitro* studies demonstrate that recombinant aS can assume various α-helical structures interacting with lipid vesicles (Davidson et al., 1998). It can also move from a fibrillation prone partially folded conformation to β-sheet species both in monomeric and oligomeric states. Its aggregates can be morphologically different, including spheres or ring oligomers, amorphous aggregates and amyloid fibrils (Uversky, 2003; Uversky, 2007) (see section 1.4.4). Based on this unusual conformational behaviour, Uversky proposed the concept of chameleon protein, which holds that the structure of aS monomer is modulated by its environment to a dramatic degree. Because of its plasticity, aS chameleon protein could be able to perform multiple functions. Indeed, Wright and Dyson proposed a reassessing of the structure-function paradigm due to the peculiarities of natively unfolded proteins (Wright and Dyson, 1999).

Data summarized so far refers to structural *in vitro* studies on recombinant aS. Some drawbacks of using recombinant are forcing non-physiological conditions, depriving the protein of the physiological modification it could be affected and of physiological binding partners and local ion environment. Outstanding data, published during the redaction of this Thesis, changed completely the way of considering the native state of aS. Bartel and colleagues proved that endogenous aS isolated from neuronal and non-neuronal cell lines, brain tissue and living human cells occurs mainly as folded tetramer. aS tetramers were purified and analyzed under non-denaturing conditions, including native gels and analytical ultracentrifugation detergent-free. Tetrameric state was hypothesized on the base of native gels results (about 48 kDa) and scanning transmission electron microscopy (STEM) measurements (55 kDa average). STEM imaging yielded a homogenous distribution of roughly spherical particles measuring about 3-3.5 nm in diameter (Fig. 1.5, A), and CD measurement demonstrate

that this particles have α-helical structure. Purified tetramers show higher affinity for negatively charged lipids than recombinant aS, and, most of all, less ability to aggregate. These findings are in disagreement with most of published results obtained studying recombinant aS, and redefine the concept of aS native state, with a number of functional implications on the physiological role. The divergence can be explained by the widespread use of denaturing agents, including sample heating, during the purification of recombinant aS from bacteria (Bartel et al., 2011).

A second study published independently by Wang and colleagues sustains the thesis of α-helical folded tetramer as native state of aS. In this case the study was performed on recombinant aS, produced as fusion GST-fusion protein (glutathione S-tranferase) which was removed enzymatically prior to analyses, but result in a 10-residues N-terminal extension. Again, heating treatment and the use of detergents were avoided during purification. Transmission electron microscopy (TEM) images of aS showed spherical particles consistent with STEM images reported by Bartel and colleagues. Reference free alignment and clustering were use to separate particles by size into three groups (fig 1.5 B and C). Small and medium particle averages resemble, respectively, three and four V-shaped repeating units arranged in a three- and tetra-fold symmetrical configuration (fig 1.5 D and E). Cross-linking experiments associated with mass spectrometry analysis confirm the presence of cross-linked trimer and tetramer. CD and NMR studies demonstrate the presence of α-helical structure (Wang et al., 2011).

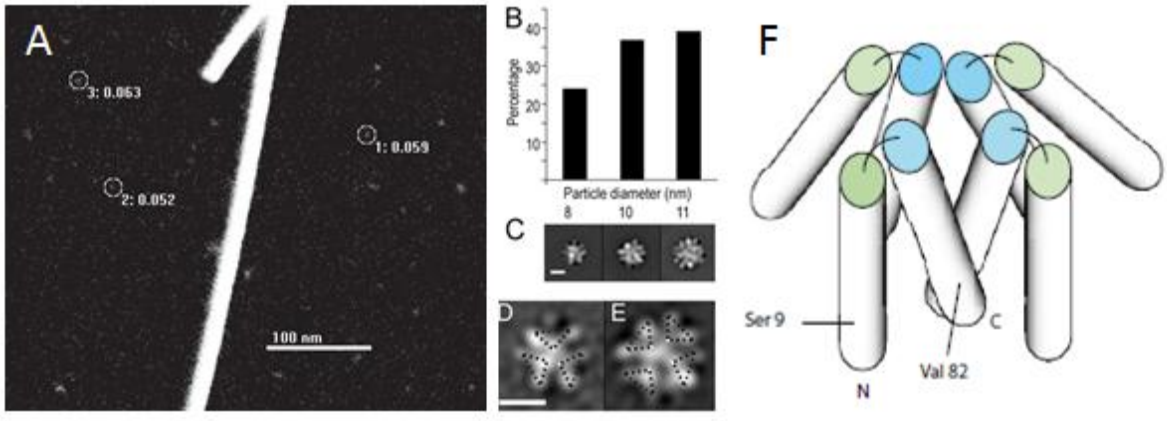


Fig. 1.5 aS tetramer. STEM image of aS purified from human red blood cells (scale bar, 100 nm) (A); distribution of aS particles sizes of TEM analyzed purified recombinant aS, and relative TEM images of overall class averages obtained from the three-sized particle groups (scale bar, 5nm) (B and C); representative class averages from the small- and medium-sized particle groups, with symmetry units shown as dashed triangles over the EM class averages (scale bar, 5nm) (D and E); model of compact aS tetramer based on EM reconstruction and PRE studies (F), helices are represented as cylinder, N- and C-terminals are indicated, the first helix end at residues 42, and the second starts at residues 50 and ends at residue 103, the remaining polypeptide, which is aspected to be disordered is not shown, the approximate position of Ser-9 (replaced by Cys for PRE experiments) and Val-82 is shown (A reprinted from Bartles et al., 2011; B, C, D, E and F reprinted from Wang et al., 2011).

On the base of the obtained results, Bartel and colleagues propose that the destabilization of the helically folded tetramer proceeds aS misfolding and amyloid aggregation.

1.4.2 Physiological role

Despite the evidence for a crucial role of aS in the pathogenesis of several neurodegenerative disorders, and its abundance in cytosolic brain fractions, the physiological function remains elusive (Iwai et al., 1995; Uversky, 2007). Several putative biological functions and possible interactions of aS have been reported. On the base of its abundance in the cytosol, its unfolded structure, and its prevention of protein aggregation induced by heat shock or chemical treatment , a putative role of aS is the chaperone function (Souza et al., 2000). Actually, aS can act as a molecular chaperone assisting in the folding and refolding of synaptic proteins called soluble NSF (N-ethylmaleimide sensitive factor) attachment receptors (Chandra et al., 2005). A recent proteomic study evidenced 587 proteins as aS-binding partners in a neuronal-hybrid cell line (Jin et al, 2007). Even considering the possibility of an overestimation, the chaperone activity could well fit with the huge number of protein-interactions.

aS is probably involved in synaptic plasticity, since it is expressed in regions of the brain that display highly synaptic connected regions, both in rat CNS, and in zebra finch, where it is transiently expressed in an epencephalic area associated with song acquisition during the critical period for song learning (Iwai et al., 1995; George et al., 1995). Moreover, aS knockout mice exhibit enhanced DA release at nigrostriatal terminals only in response to paired electrical stimuli, suggesting that aS is an activity-dependent, negative regulator of dopaminergic neurotransmission (Abeliovich et al., 2000). In particular, the potential function of aS in DA regulation and storage is correlated with its ability to interact with membranes and regulate vesicular trafficking (Lotharius and Brundin, 2002). In pre-synaptic termini, aS exists in an equilibrium between free and plasma membrane- or vesicle-bound states (McLean et al., 2000), with approximately 15% of aS being membrane-bound (Lee et al., 2002). Furthermore, A30P and A53T aS mutations, pathological in PD, abolish the protein's ability to bind to small phospholipidic vesicles or planar lipid membranes, respectively. This close connection with vesicular structures, together with the loss of function of mutations, led to the hypothesis that aS may control vesicular release and/or turnover and other synaptic functions in CNS (Ueda et al.,1993; Clayton & George, 1998, 1999; Davidson et al. 1998; Lavedan, 1998). Genome-wide screening in yeast showed that nearly one-third of genes that enhance the toxicity of aS are functionally related to lipid metabolism and vesicle trafficking (Willingham et al., 2003; Uversky, 2007). Moreover, membrane fluidity and in cellular fatty acid uptake and metabolism are strongly influenced by over-expression of aS in a neuronal cell line and in aS ko (Sharon et al., 2003; Castagnet et al., 2005; Golovko et al., 2005).

1.4.3 Post translational modification and alternative splicing

Alternative splicing of exon 3 and 5 of aS gives rise aS 126 and aS 112, respectively. The two isoforms are characterized by the in-frame deletions of amino acids residues 41-54 for aS 126 and 103-130 for aS 112 (Bayer et al., 2004; Bayer et al., 2006). Both of the isoforms are less abundant respect to the whole protein transcript, and differently expressed in diverse disorders (DLB, AD and PD). However, the detection of the isoforms of aS is difficult to carry out due to the small molecular weight differences (Beyer, 2006).

Several aS post-translational modifications have been reported, and some of them lead to the formation of stable oligomers. These include nitration, oxidation, phosphorylation, and interaction with iron (Estevans et al., 2011). Extensive phosphorylation of aS represents probably a significant pathogenic event. More that 90% of insoluble aggregates of aS in DLB brains contain phosphorylated aS, whereas phosphorylation involves only about 4% of soluble, monomeric aS (Fujiwara et al., 2002; Chen et al., 2005). aS is mainly phosphorylated at S87, while a second site is represented by S87 (Okochi et al., 2000). It has been described that aS could be glycosylated (Shimura et al., 2001). Ubiquitination is the targeting mechanism for proteosome degradation. Two E3 ubiquitin ligases have been identified to ubiquitinate aS: parkin and SIAH-2 (Shimura et al., 2001; Liani et al., 2004). aS ubiquitination in vivo occurs at K6, K10 and K12 (Liu et al., 2002). Sumoylation has also been described as aS post translational modification. As ubiquitination, this modification could be responsible for proteosomal targeted degradation (SUMO stands for small ubiquitin-like modifier), but it could also play a different role, since this modification is less known (Dorval et al., 2006). Eventually, as described in section 1.3.2, oxidation processes contributes to cell degeneration in PD (Jenner, 2003). aS is directly affected by oxidation and nitration, and these modification could lead to the formation of stable oligomers. Moreover, oxidative modification of aS via dopamine quinones may facilitate aggregation (Lotharius, 2002).

1.4.4 Fibril structure and aggregation process

Mature amyloid fibrils of aS, grown *in vitro*, are formed by two, or more, filaments and show a periodic twist (Fig. 1.6 A). They typically vary in length from about 500 nm to 3 μm and, based on AFM analyses, have an average height of 9.8 ±1.2 nm (Fink, 2006; Khurana et al., 2003). The core region of aS, defined by proteolysis studies, encompass the central, highly amyloidogenic NAC region (Miake et al., 2002; Quin et al. 2007). Electron paramagnetic resonance spectroscopy (EPR)

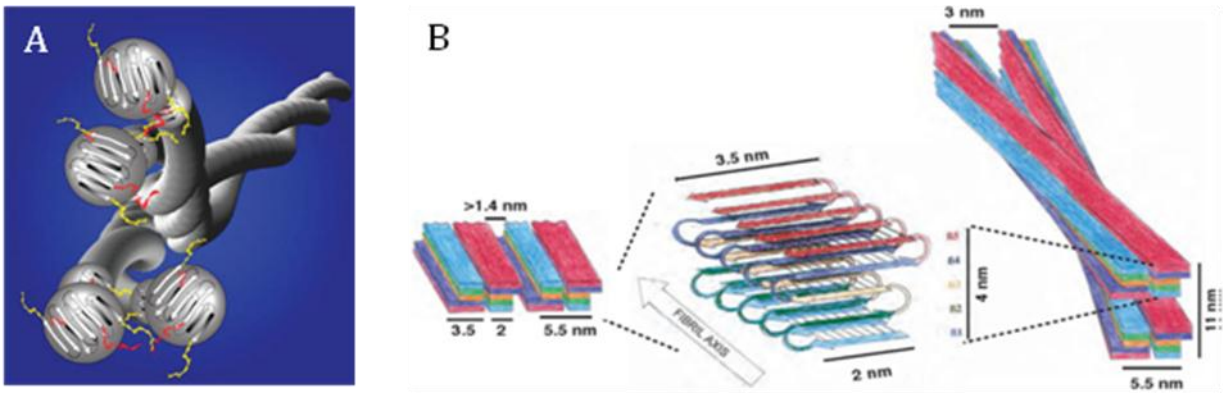


Fig. 1.6 Schematic representation of aS fibrils. A Model of the hierarchical structure of a aS in twisted fibrils (reprinted from Quin et al., 2007). **B** The proposed fold of a monomeric aS within a protofilament is shown in *Center*. The incorporation of a protofilament into the straight (*Left*) and twisted (*Right*) fibril type is indicated by a schematic drawing (reprinted from Vilar et al., 2008).

and site-directed spin labelling studies indicate that the fibril core is formed by parallel, in-register arrangement of multiple β-strands that run perpendicular to the fibril axis, where each layer contain a new molecules (Chen et al., 2003). On the base of D/H exchange experiments, monitored by NMR, AFM measurement and recent EPR studies, the β-strands which constitute the fibril core are actually supposed to be five (Fig. 1.6 B) (Vilar et al., 2008; Karyagina et al., 2011).

The *in vitro* kinetics of aS fibril formation show an initial lag-phase followed by an exponential growth phase and a final plateau (Fink, 2006). The process is nucleation-dependent (Wood et al., 1999). Early stages of fibril formation involve the partial folding of aS into a highly fibrillation-prone pre-molten globule-like conformation, which represents a key intermediate on the fibrillation pathway (Uversky et al., 2001). The process than progresses through a series of oligomeric intermediates, called oligomers or protofibrils, that normally disappear upon fibril formation (Fig. 1.7). Therefore, considering the possibility that aS could be a folded tetramer (Bartel et al., 2011; Wang et al., 2011) or that the unstructured monomer presents long-range interactions between the C-terminal and the central NAC regions (Bertoncini et al., 2005), a conformation change to a fully unfolded state should be expected prior to the aggregation. The aggregation-prone conformation may be attained early in the aggregation pathway, through the exposition of the hydrophobic NAC region and facilitating aggregation (Paleček et al., 2008). Thus, perturbations that increase the hydrophobicity of the system (i.e. low pH, which decrease the net charge of the C-terminal region) favour the aggregation (Wood et al., 1999; Hoyer et al., 2004). In addition to pH, also agitation and concentration strongly influence the aggregation kinetic.

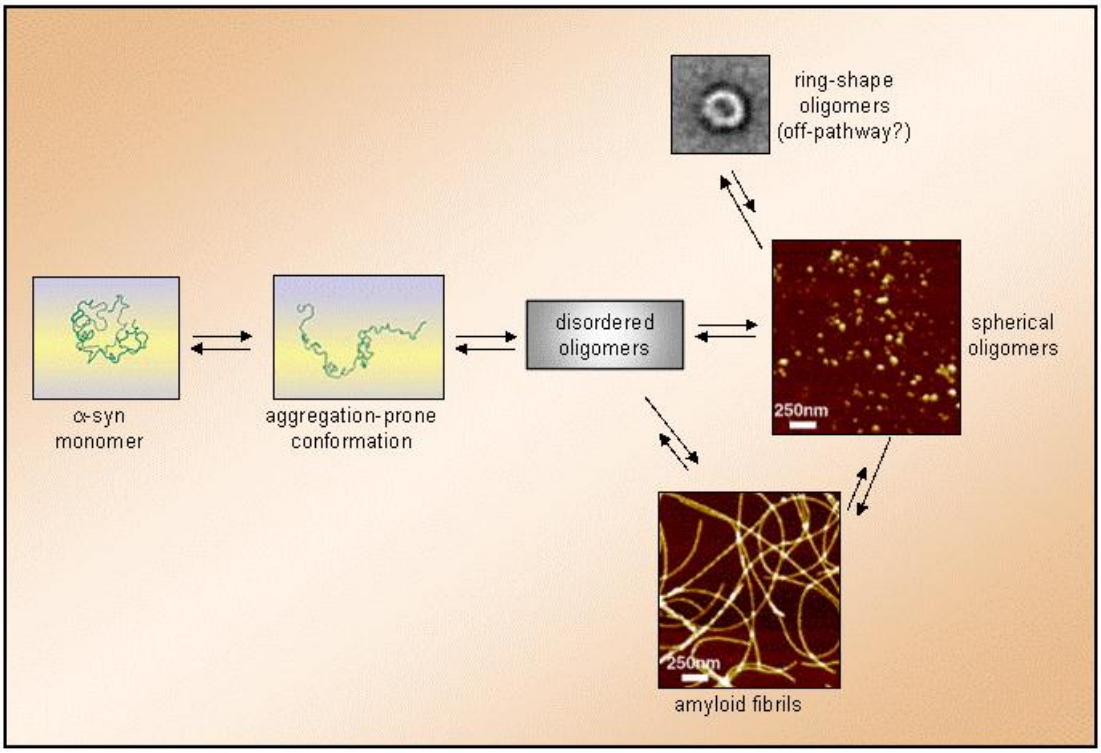


Fig.1.7 Schematic representation of the hypothetical aggregation pathways of aS. During aS fibrils formation several intermediate structures populate, including small oligomers, early-spherical protofibrils, membrane embedded pores or chain like aggregates. Therefore, aS fibrillation is not a simple two-state transition from monomer to fibrils, but rather a complex process that involves intermediates of various sizes and morphologies.

Microscopy techniques have revealed that different products may arise from the aggregation of aS depending on the experimental conditions: fibrils, on and off pathways stable oligomers, and insoluble amorphous aggregates. Although the central role of aS amyloid fibrils formation in the pathogenesis of PD is well-recognized, recent findings indicates that the fibril itself may not be the principal pathogenic species. Soluble oligomers and protofibrils seem to be the toxic species (Volles et al., 2002). As a result of the unstable and transient nature of these intermediates, the direct detection and characterization of prefibrillar oligomeric species has proven to be extremely difficult. The oligomeric species of aS are present during the lag-phase of the aggregation process, reaches a maximum concentration of 15-25% toward the end of the lag time and then decline as fibril growth accelerates (Goldberg et al., 2000; Conway at al. 2001; Fink 2006). Several oligomeric prefibrillar species with various morphologies have been described (Conway et al. 1998, 2000; Ding et al. 2002; Lashuel et al. 2002, Apetri et al. 2006). The earliest appreciable intermediates appeared to be predominantly spherical with heights varying between 2.5 and 4.2 nm (Conway et al. 2000; Ding et al. 2002). Lashuel and colleagues described annular oligomers having an average diameter of approximately 10-12 nm or less (rings of ~8 nm diameter were described) and tubular/rectangular oligomer with a mean diameter of 12 nm and varying length of 11-20 nm (Lashuel et al., 2002). The incubation of the spherical aS oligomers with brain

derived membranes was shown to also produce pore-like ring-type protofibrils (Ding et al. 2002). These types of oligomers are unable to bind Thioflavin T (ThT) and Congo Red (CR), that typically bind amyloid structures, but have β-structure, characterized by Raman-IR spectroscopy (Apetri et al., 2006). Other techniques used to characterize oligomers are photo-induced cross-linking (Li et al., 2006), electrochemical techniques (Paleček et al., 2007), intrinsic tryptophan fluorescence of Trp aS mutant (Dusa et al., 2006) and FRET measurements of Phe or Trp aS mutants (Kaylor et al., 2005), and fluorescence of pyrene-labeled aS (Thirunavukkuarasu et al., 2008). In addiction to transient oligomers, stable oligomers, results of several chemical modifications of aS, have been decribed. These modifications include oxidation of the four methionine residues to methionine sulfoxide (Hokenson et al., 2004), specific nitration of the tyrosine residues (Uversky et al., 2005), interaction with polyphenols such as bacalein (Zhu et al., 2004) and covalent modification by 4-hydroxynonenal (Qin et al., 2006). No significant fibrillation occurs from these modified forms of aS. Biophysical characterization of these oligomers suggests that they have significant secondary and tertiary structure and are substantially more compact than monomeric aS. The stability of these oligomers indicates that the underlying structure of their subunits is different than that in the transient oligomers (Fink, 2006).

1.5 Aim of the study

Parkinson's disease (PD) is a movement disorder resulting from the degeneration of dopaminergic neurons in the *substantia nigra*. The cause of the disease is unknown, but several evidences suggest that the aggregation of α-synuclein (aS) is a critical step in the etiology of PD. In PD patients, aggregated aS is the major component of LBs, the cytoplasmic proteinaceous aggregates which characterize PD, and through which the pathology is diagnosed. The structure of aS can be divided into three regions: residues 1-60, which contain four 11-amino acid imperfect repeats (KTKEGV); residues 61-95, which contain the amyloidogenic NAC region and the highly negative charged C-terminal region, residues 96-140. aS was for long time considered natively unfolded under physiological conditions, and hence highly dynamic. Recently, a tetrameric fold of aS in α-helix conformation was described in physiological condition. Indeed, aS acquires α-helix conformation upon interaction with lipids and membranes. On aggregation, aS undergoes a conformational change into amyloid fibrils, which contain a cross-β structure in which individual β-strands run perpendicular to the fiber axis.

Several factors have been shown to trigger the oligomerization of aS, including the formation of stable dimers of aS through oxidation or nitration mechanisms and the formation of di-tyrosine adducts (Takahashi et al., 2002; Krishnan et al., 2003). Dimerization of aS could represent a critical, rate-limiting step for the aggregation of the protein. Therefore, we decided to study the aggregation of several different dimers of aS, produced through molecular biology techniques. The dimers of this study have been

designed in order to link the protein at the terminals. A cysteine residue has been added at the N-terminal or at the C-terminal of aS, therefore producing NN dimer and CC dimer through the formation of a disulfide bond. A NC dimer, formed by two aS contiguous molecules, was obtained as a single polypeptide chain. During the project, another dimer, called DC dimer, was produced in order to further draw up the hydrophobic regions, and avoid the interferences of side chains within the molecule. DC is constituted by two consecutive central, highly amyloidogenic regions, containing aS residues from 1 to104 joined to residues from 29 to 140 (Fig. 1.8).

This Thesis work focuses on the characterization of the aggregation properties of dimers. All measurements were conducted in comparison with aS. The proposal is to collect information on the kinetic of fibrillation and the structure of the mature fibrils formed by the different molecules. The differences in the orientation and folding of dimers in respect to aS, would hypothetically differentiate the fibrillation process of the five protein structures. aS dimers would represent a suitable tool for the study of intra-molecular aS interaction pathway, to eventually obtain structural information on aS aggregation processes, which is still poorly understood.

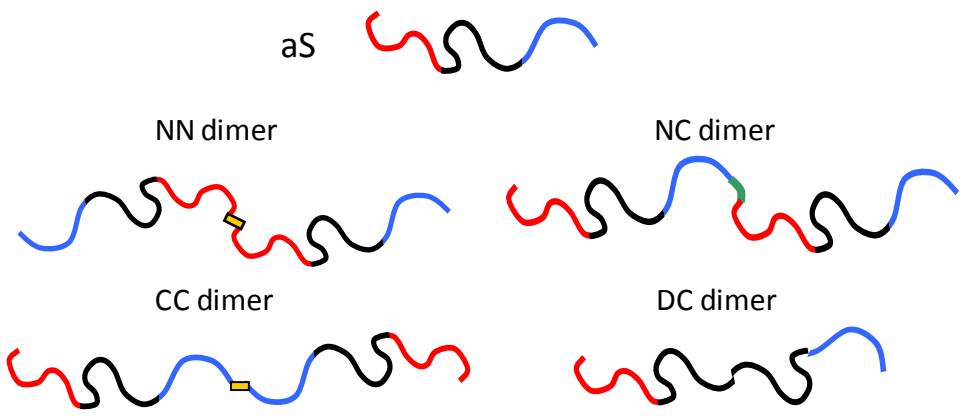


Fig. 1.8 Schematic representation of aS and aS dimers. N- and C-terminal regions are drawn in red and blue, respectively. Yellow rectangle represents disulphide bond, which links aS monomers in NN and CC dimers. NC and DC dimers were produced as single polypeptide chains.

3. Materials and methods

3.1 Materials

Proteinase K from *Tritirachium album* and porcine trypsin were purchased from the Sigma Chem. Co. (St. Louis, MO). All other chemicals were of analytical reagent grade and were obtained from Sigma or Fluka (Buchs, Switzerland).

3.2 Expression and purification of recombinant aS and aS dimers

Human aS cDNA was amplified by PCR with synthetic oligonucleotides (Sigma-Genosys) containing NcoI and XhoI restriction sites and designed to obtain the entire sequence of the protein (aS 1-140). After digestion with restriction enzymes, the two PCR products were subcloned into the NcoI-XhoI-linearized pET28b expression plasmid (Novagen). Further pcr were performed to mutagenize aS sequence and thus producing the sequence of NN, CC, NC and DC dimer sequences, which were introduced as well in pET28b expression plasmids (Novagen).

The expression and purification of aS and its dimers was conducted as indicated in the the following procedure. *E.coli* bacteria, BL21 (DE3) strain, are transformed with the pET-28b(+) plasmid containg kanamicine resitence and the correct gene sequence (aS, NN, CC, NC or DC dimers). A pre-colture is growth overnight in LB medium (1% Bacto tryptone, 0.5% yeast extract, 0.5% NaCl, supplemented with 25 μg/mlkanamicin) at 37 °C, agitating. Bacteria are then diluted in 1 l LB medium and growth to an OD_{600} of 0.6-0.8, followed by induction with 0.5 mM isopropyl β-thiogalactopyranoside (IPTG) for 4 hours. Bacteria are harvested by centrifugation at 6000 rpm for 10 minutes at room temperature and resuspended in 100 ml Osmotic Shock Buffer (30 mM Tris, 2 mM EDTA, 40% v/v sucrose, pH 7.2), according to Huang et al. (2005). The suspension is incubated at room temperature for 10 minutes, and then it is centrifuged at 12000 rpm for 10 minutes. Pellet is rapidly resuspended in 90 ml cold deionised water with the addition of 37.5 μl of $MgCl_2$ saturated solution. Protein release from periplasmatic space occurs in this passage. The suspension was kept on ice for 3 minutes, and then the bacteria were collected by centrifugation at 12000 rpm for 20 minutes. Bacterial pellet was wasted and the supernatant was boiled for 10 minutes. At this spec supernatant can be stored overnight at 4 °C. a centrifugation step (12000 rpm for 20 min at 4°C) proceeds the protein ammonia sulphate precipitation. 8 ml of 1 M Tris pH 8 is added to the supernatant, and the solution diluted to a final volume of 100 ml. Then, two steps of ammonium sulphate fractionation is performed: A centrifugation step is carried out at 35% saturation of ammonium sulphate solution, and then ammonium sulphate was added to the supernatant until a final concentration of 55% saturation. The precipitate is

recovered by centrifugation (12000 rpm for 20 min, at 4 °C), and resuspended in water. A dialysis agaist water is conducted over night to remove salts. Further purification is than performed by ion exchange chromatography with a Resource Q FPLC column (GE Healthcare). The column is preequilibrated with 3 column volume 20 mM Tris, pH 8. The elution was performed with a 30 min linear gradient from 0 to 100% 20 mM Tris, 500 mM NaCl, pH 8. Peaks corresponding to aS or aS dimers was collected and dyalized against deionized water overnight, then lyophilized. For NN and CC dimers, two elution peaks are evident; the first one corresponding to reduced Cystein and the second to the dimer (identification performed by MS). The purity of each purification batch is checked by RP-HPLC, performed with a Jupiter analytical C4 column (Phoenomenex, USA). Elution was performed with a linear gradient of acetonitrile (0.085 % TFA) versus water (0.1 % TFA), from 5% to 38% in 5 min and from 38% to 43% at a flow rate of 0.6 ml/min. Whether further purification was needed, this was performed with preparative RP-HPLC column (Phenomenex, USA), preequilibrated with 5% solvent A (0.1% trifluoroacetic acid in milliQ water) and eluted with a linear gradient from 44 to 50% solvent B (0.085% trifluoroacetic acid in acetonitrile) in 30 min, with 2 ml/min costant flow. The identity and integrity of the eluted material were assessed by mass spectrometry. Finally, after a dialysis phase to remove salts, the proteins were lyophilized and stored at -20 °C.

3.3 Aggregation studies

In order to analyze the aggregation process of aS, lyophilized protein obtained from purification was dissolved in PBS phosphate buffer (8 mM Na_2HPO_4, 137 mM NaCl, 2 mM KH_2PO_4, 2.7 mM KCl), pH 7.4, and filtering the protein solution with a 0,22 μm pore-size filter (Millipore, Bedford, MA, USA). Dissolved samples was incubated at 37 °C for up to 15 days at a protein concentration of 70 μM (1 mg/ml), under shaking at 500 rpm with a thermo-mixer (Compact, Eppendorf, Hamburg, DE). The same experiment has been conducted in parallel for NN, CC, NC and DC dimers. Aliquots of the samples during incubation were examined by native-PAGE, CD, Thioflavin T binding assay and TEM.

3.4 Gel filtration

Gel filtration chromatography (or size exclusion chromatography) was performed with a Superdex 200 10/300GL column (Amersham Biosciences, Uppsala, Sweden), using an ÄKTA FPLC system (Amersham Biosciences, Uppsala, Sweden). The matrix is composed of cross-linked agarose and dextrane with an average particle size of 13 μm and a 3-70 KDa separation range for globular proteins. The hydrodynamic volume of analytes was determined on the base of the the distribution coefficient, Kd, calculated as the following formula: $Kd= (V_e-V_o)/(V_t-V_o)$, where V_e, V_t and V_o are respectively the analyte elution volume, total and void volume. If the analyte is large and completely

excluded from the mobile phase within the gel, K_d = 0 whereas, if the analyte is sufficiently small to gain complete access to the inner mobile phase, K_d = 1. Due to variation in pore size between individual gel particles, there is some inner mobile phase that will be available to analytes of intermediate size; hence K_d values vary between 0 and 1. It is this complete variation of K_d between these two limits that makes it possible to separate analytes within a narrow molecular size range on a given gel.

In order to characterize aS and its dimer behaviour in solution, before the analysis of the proteins, a calibration controls with globular proteins of known MW was performed. 50 μg of each of the following standard calibration proteins were loaded: bovine α-lattalbumine (MW= 14 kDa), carbonic anhidrase (MW= 29 kDa), ovoalbumin (MW= 45 kDa), BSA(MW= 66 kDa), thyroglobulin (MW= 440 kDa) , β-amilase (MW= 200 kDa) and ferritin (MW= 669 kDa). 200 μg of bue dextran and 0.05% dimethyl sulfoxide (DMSO) were loaded to estimate the void and total volume of the column, respectively. The elution was monitored by recording on-line the absorbance at 214 nm. The calibration curve results by plotting Kd against the logarithm of the molecular weight of the calibration protein. 50 μg of samples (aS, NN, CC, NC and DC dimer), dissolved in the running buffer, were loaded into the column and eluted at 0.4 ml/min in 20 mM Tris-HCl, 15O mM NaCl, pH 7.4.

3.5 Circular Dichroism

Protein concentrations were determined by absorption measurements at 280 nm using a double-beam Lambda-20 spectrophotometer from Perkin Elmer (Norwalk, CT). The extinction coefficients at 280 nm were 5960 M^{-1} (aS), 12'045 M^{-1} (NN and CC dimer), 11'920 M^{-1} (NC dimer) and 7'450 (DC dimer) as evaluated from its amino acid composition by the method of Gill and von Hippel (1989).

Circular dichroism spectra were recorded on a Jasco J-710 (Tokyo, Japan) spectropolarimeter. Far-UV CD spectra were recorded using a 1 mm path-length quartz cell and a protein concentration of 3-10 μM. The mean residue ellipticity $[\theta]$ (deg·cm^2·dmol^{-1}) was calculated from the formula $[\theta] = (\theta_{obs}/10) \cdot (MRW/lc)$, where θ_{obs} is the observed ellipticity in deg, MRW is the mean residue molecular weight of the protein, *l* the optical pathlength in cm and *c* the protein concentration in g/mL. The spectra were recorded in PBS buffer, pH 7.4, both in the presence of different concentration of SDS (ranging from 0-5 mM) and during aggregation studies. For SDS titration, protein concentration was 20 μM for aS, NN, CC and NC, and 6.7 μM for DC dimer, while measurements of the aggregation aliquots were performed at 0.05-0.15 mg/ml.

3.6 FT-IR Fourier transformed infrared scpectroscopy

Deuterated aS and dimers were prepared by dissolving the proteins in D_2O, filtering the protein solution with a 20 nm pore-size filter (Whatman, Maidstone, UK) and incubating it at -80 °C for 40 min, followed by lyophilization.

The spectra of aS and dimers in solution were registered after dissolving the monomeric deuterated protein in 20 mM Tris·DCl, 150 mM NaCl, pH 7.2 at a 5 mg/ml concentration. The spectra of aS and dimers fibrils were registered on the same buffer at 3-5 mg/ml concentration. Fibrils obtained from deuterated protein aggregations after 30 days at 37°C, 500 rpm shaking in 20 mM Tris·DCl, 150 mM NaCl, pH 7.2 at 1 mg/ml concentration were collected by ultracentrifugation at 90'000 rpm for two hours at 4°C, (Optima MAX, Beckam Coulter). The concentration of the ultracentrifuged fibrils was evaluated by subtraction of the amount of protein presents in the supernatant (measured by UV absorbance).

FTIR spectra were recorded at 20-22 °C using a Perkin Elmer 1720X spectrometer (Norwalk, CT, USA), purged with a continuous flow of N_2 gas. Protein samples were placed between a pair of CaF_2 windows separated by a 50 µm Mylar spacer. For each protein sample, 50 interferograms were accumulated at a spectral resolution of 2 cm^{-1}. The spectra were analyzed using the Grams 32 program version 4.14 (Galactic Industries Corporation, Salem, NH). Buffer spectra were recorded under identical conditions to those used for protein samples and subtracted from the spectra of the latter. The second derivative of the amide I band was used to identify the different spectral components. Thereafter, curve fitting was performed with Gaussian and Lorentzian lineshapes, and with bandwidths varying between 15 and 25 cm^{-1} (Byler and Susi, 1986; Arrondo et al., 1993).

3.7 Nuclear Magnetic Resonance

Overexpression of proteins for NMR studies (15N-labeled proteins) was achieved by growing cells in M9 minimal medium. Per litre: 200 ml M9* salts solution, 2 ml 1 M MgSO4, 0.1 ml 1 M CaCl2, 200 ml 20% glucose; *M9 salts solution, per litre: 64 g Na2HPO4, 15 g KH2PO4, 2.5 g NaCl, 5.0 g NH4Cl supplemented with 1 g/L [15N]ammonium chloride.

Heteronuclear Single Quantum Correlation (HSQC) spectra were acquired on a Bruker Avance DMX spectrometer equipped with a gradient triple resonance probe. ^{15}N-labelled αsyn samples were dissolved in phosphate buffered saline (PBS) with 10% D_2O. The experiments (256 increments of 512 time points each) were acquired at 283 K with 16 transients. The spectral widths were 3 ppm (^1H) and 22 ppm (^{15}N) and the frequency offsets were 8 ppm (^1H) and 116 ppm (^{15}N). Prior to Fourier transformation, the data were multiplied by a 90° shifted *sin* function in both dimensions. Spectra were processed using MestReC software.

3.8 Thioflavin T binding assay (ThT)

The ThT binding assays were performed accordingly to LeVine (1993) using a freshly prepared 25 μM ThT solution in 25 mM sodium phosphate (NaH_2PO_4), pH 6, that had been passed through 0.22 μm filters. Aliquots (30 μl) of protein samples from the aggregation incubation were taken at specified times and diluted into the ThT buffer (final volume 500 μl). Fluorescence emission measurements were conducted at 25 °C using an excitation wavelength of 440 nm and recording the ThT fluorescence emission spectra between 460 and 560 nm.

3.9 Native-PAGE

Native (non-denaturing) polyacrylamide gel electrophoresis was performed at a constant 100 V using a Mini–PROTEIN II Bio-Rad electrophoresis system using a Tris–HCl 12% (w/v) polyacrylamide gel. Approximately 5 μg of protein were loaded into each well, and the protein bands were visualized by silver or Coomassie blue staining.

Silver staining was performed incubating the gel in 50% ethanol/10% acetic acid for 30 minutes followed by 15 minutes 5% ethanol/1% acetic acid. Three wash steps with water are performed prior to 2 minutes incubation with Thiosulphate solution (0,2 g/l $Na_2S_2O_3$), and followed by another wash step in water. Gel is then incubated in $AgNO_3$ (2 mg/ml) for 30 minutes and than developed with thiosulphate developing solution (for 100 ml: 2ml Thiosulphate solution, 6 g Na_2CO_3, 50 μl formaldehyde). Colorimetric reactuion is spopped with 5% acetic acid and the gel is then stored in 0.05 M HCl.

3.10 Transmission Electron Microscopy

In order to evaluate the morphology and the size of the species deriving from the self-assembly of DHA and from the aggregation process of aS, aliquots of the samples were examined by transmission electron microscopy (TEM). The samples relative to aggregation of the proteins were diluted 3 times with PBS. A drop of the samples solution was placed on a Butvar-coated copper grid (400-square mesh) (TAAB-Laboratories Equipment Ltd, Berks, UK), dried and negatively stained with a drop of uranyl acetate solution (1%, w/v). TEM pictures were taken on a Tecnai G^2 12 Twin instrument (FEI Company, Hillsboro, OR, USA), operating at an excitation voltage of 100 kV.

3.11 Proteolysis of the fibrils

Proteolysis experiments were carried out on ultra-centrifugated aggregation samples obtained after 15 days incubation (1mg/ml, PBS pH 7.4, 37°C, 500rpm).

Ultracentrifugation divides soluble protein monomers and aggregates (supernatant) from fibrils (pellet). Pellets were resuspended in PBS, pH 7.4, and divided for proteolytic analyses using two enzymes: proteinase K (Ebeling et al., 1974) at E/S ratio of 1:1000 (by weight) and trypsin at E/S ratio of 1:50 (by weight). The reactions were quenched at specified times by acidification with TFA in water (4%, v/v).

The proteolysis mixtures were again ultra-centrifuged (90'000 rpm for two hours at 4°C) and analyzed by RP-HPLC according to Schägger and von Jagow (1987). Prior to analysis pellets obtained after the reaction were incubated over night with guanidine-HCl 6 M. For all the proteins, the RP-HPLC analyses were conducted using a Vydac C_{18} column (4.6 mm x 250 mm; The Separations Group, Hesperia, CA), eluted with a gradient of acetonitrile/0.085% TFA *vs.* water/0.1% TFA from 5% to 25% in 5 min, from 25% to 28% in 13 min, from 28% to 39% in 3 min, from 39% to 45% in 21 min at a flow rate of 1 ml/min. The same column was used to analyze the proteolytic patter of ayn108-140 using a gradient of acetonitrile/0.085% TFA *vs.* water/0.1% TFA from 5% to 25% in 5 min, from 25% to 28.5% in 24 min, at a flow rate of 1 ml/min. For syn 1-99, syn1-52 and syn 57-102 the same gradient and the same columns utilized for the purification were used. Each column was provided by a HPLC security guard column C_4 (Phenomenex, USA).

The sites of cleavage along the polypeptide chains were identified by mass spectrometry analyses of the protein fragments purified by RP-HPLC. Mass determinations were obtained with an electrospray ionization (ESI) mass spectrometer with a Q-Tof analyzer (Micro) from Micromass (Manchester, UK). The measurements were conducted at a capillary voltage of 2.5-3 kV and a cone voltage of 30-35 V. The molecular masses of protein samples were estimated using the Mass-Lynx software 4.1 (Micromass).

3. Results

3.1 aS and dimers production

All the experiments reported in this Thesis were conducted on recombinant proteins expressed in *E. Coli* bacteria. aS dimers were produced by molecular biology techniques. Since aS does not contain Cys residues, the formation of NN and CC dimers, which link respectively N- and C- terminals of aS, were obtained through the addition of a Cys residue at aS terminals and the formation of disulphide bonds, which occurs spontaneously in solution. For NN, the Cys residue was introduced by V3C substitution, while for CC, GC residues were added at position 140 of aS. NC, which links N- and C-terminals of aS, was produced as single polypeptide molecule containing two aS sequences joined by RS residues. Similarly, DC dimer is a single polypeptide molecule which contains 1-104 and 29-140 residues of aS. The primary structures of aS, NN, CC, NC and DC are reported in Supplementary Material and their principal chemical characteristics are reported in Table 3.1.

3.2 Chemical and physicochemical characterization

Recombinant proteins were analyzed by RP-HPLC (Fig. 3.1) and the collected peaks were identified by mass spectrometry (MS). The experimental masses correspond to the theoretical MW of the proteins, identifying unequivocally the products of the expression (Table 3.1). RP-HPLC chromatograms evidence four different retention times (RTs), characteristic for each protein species under the used experimental conditions and gradient (linear gradient of water/acetonitrile containing 0.05% TFA, from 35 to 50%). The differences in RTs between the five proteins depend on their hydrophobic properties and on the different molecular masses. All the four dimers elute after aS because of the higher MW. DC dimer, which is the smallest dimer (~22 kDa), has the higher RT due to the large hydrophobic central region, that corresponds to two tandem NAC domains.

Table 3.1. Chemical characteristic of aS and its dimers.

Protein	*Number of amino acids*	*Theoretical pI*	*Estimated charge at pH 8.3*	*Average MW (Da)*	*Experimental MW (Da)*
aS	140	4.7	-10	14460.1	14461.1 ±0.7
NN dimer	280	4.7	-20.1	28927.2	28926.3 ±0.5
CC dimer	284	4.7	-20.1	29239.6	29240.2 ±0.3
NC dimer	282	4.74	-18.4	29145.5	29145.5 ±0.1
DC dimer	216	5.22	-7.2	21969.7	21969.5 ±0.1

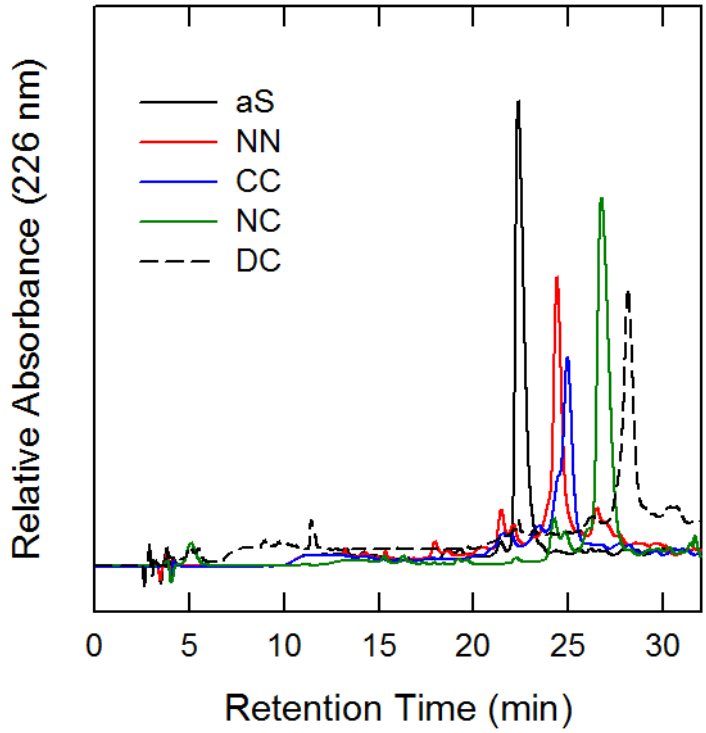


Fig. 3.1 RP-HPLC profiles of recombinant aS and its dimers after purification. HPLC profiles of aS and NN, CC, NC and DC are represented respectively with black, red, blue, green and black dotted lines. A C4 column (Phenomenex, California) was used at a flow of 0.6 mg/ml, using a linear gradient of water-acetonitrile containing 0.05% TFA from 35 to 50%.

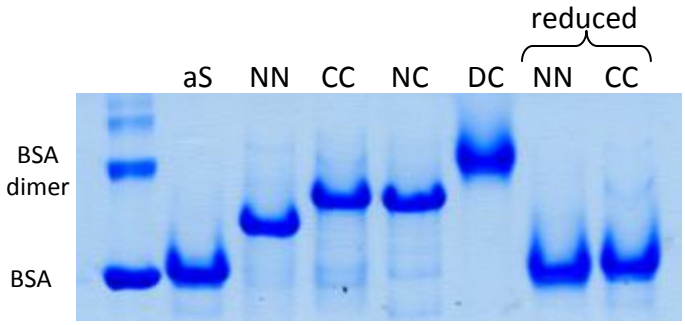


Fig. 3.2 Native gel electrophoresis. Native electrophoresis of aS and its dimers. BSA was used as marker (first lane), and NN and CC dimers were reduced with β-mercaptoethanol and loaded onto the gel.

Native gel electrophoresis was performed under basic conditions (pH 8.3) to negatively charge the proteins of interest (Fig. 3.2; estimated charge at pH 8.3 is reported in Table 3.1). The proteins migrate on the gel matrix as function of the ratio between hydrodynamic volume and the net charge (m/z ratio). aS migrates at the same level of BSA (MW=66 kDa), the dimers run slower on the gel, due to the higher molecular mass. DC dimer is the slowest protein, and migrates at the same level of BSA dimer, as a consequence of its lower net charge. Indeed, the negative charges of the carboxylic moieties are mainly due to the presence of glutammic acid (E) residues, which are mostly present at the C-terminal regions of aS. DC sequence contains only a C-terminal region while NN, CC and NC, which are formed by two-linked entire aS, contain these sequences twice. The production of NN and CC depends on the formation of a disulphide bond on Cys containing mutant of aS (see section 3.1). As expected, reduced NN and CC have the same electrophoretic properties of aS. The gel evidences, for all the protein species, the high homogeneity of the recombinant products and the absence of aggregates at high molecular levels.

The hydrodynamic volume of aS and its dimers was determined by using gel filtration chromatography (Fig. 3.3, A). The globular proteins α-lactalbumin (α-LA), carbonic anhydrase, ovalbumin, bovine serum albumin (BSA), thyroglobulin, β-amylase and ferritin were used as reference molecular weight markers (Table 3.2). A calibration curve was determined plotting different distribution coefficients (Kd) as a function of the molecular weight in a logarithmic scale. Kd was calculated as follow: $Kd= (V_e-V_o)/(V_t-V_o)$, where V_e, V_t and V_o are respectively the elution, total and void volume (Fig. 3.3 B). V_t and V_o were determined eluting dextran blue and DMSO, respectively. For each protein species, hydrodynamic volume was estimated on the base of the calculated Kd (Fig. 3.3 B, red spots). Table 3.2 reports the evaluated MW.

Proteins are resolved by gel filtration chromatography on the base on their hydrodynamic volume, which depends on both MW and the folding of the proteins. Assuming that aS is a globular protein, its elution profile would correspond to a protein of ~60 kDa. NN, CC and NC, which have about two times the MW of aS (Table 3.1), elute at 12.3 ml of elution volume, which correspond to, ~170 kDa, about 2.8 times more than aS. DC has a MW of ~21 kDa, between the MW of aS and the other dimers (Table 3.1). Its hydrodynamic volume is 98.9 kDa, again between aS and the other dimers, and 1.6 times aS. Dimers produced by the formation of disulphide bonds, NN and CC present a small portion of reduced monomer. The hydrodynamic volume of aS and the dimers was determined in order to study the behavior of the protein species in solution. aS is known to be an unfolded protein, therefore its estimated molecular mass, calculated on the base of its gel filtration chromatogram and compared with reference globular proteins, is bigger than its effective MW.

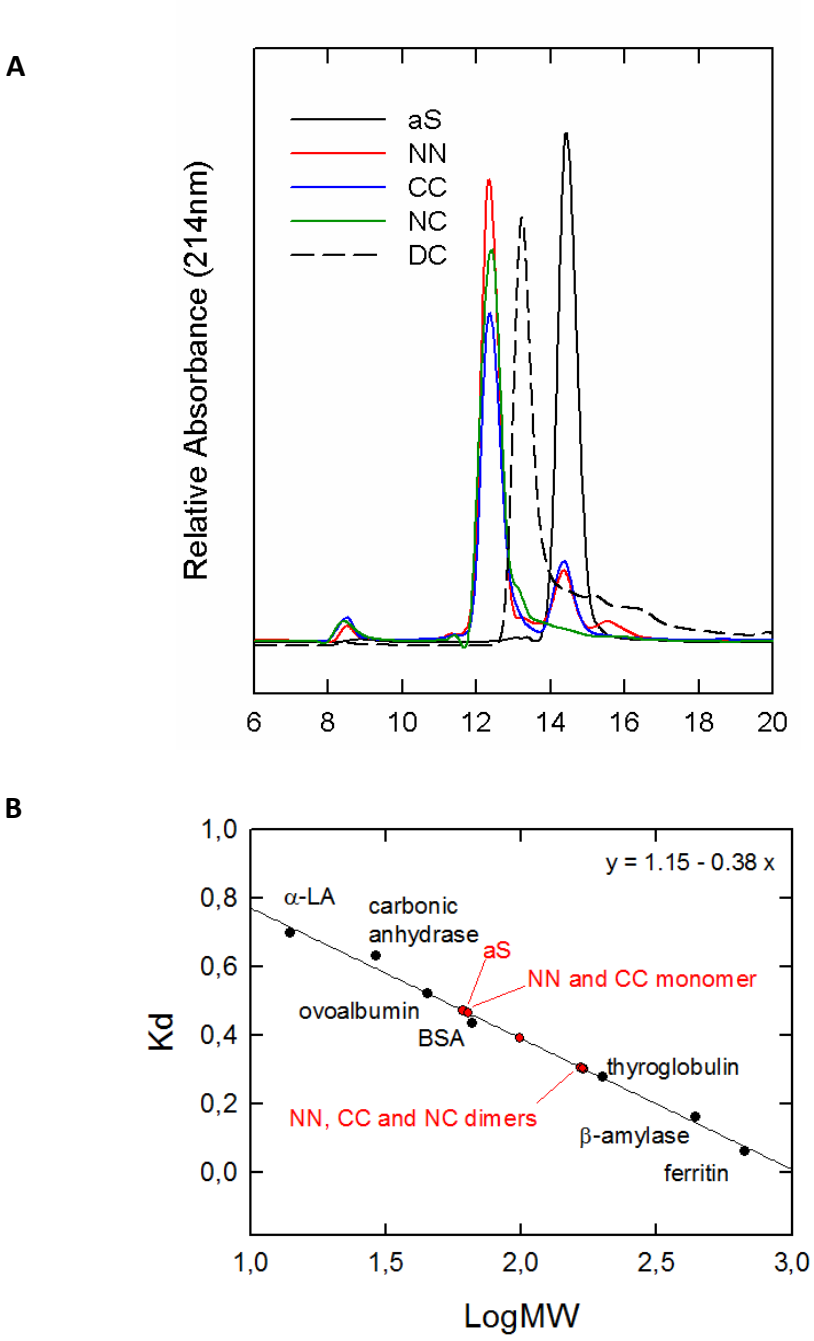


Fig. 3.3 Estimation of the hydrodynamic volume by gel filtration chromatography. A. Gel filtration chromatograms of aS and NN, CC, NC and DC (respectively black, red, blue, green and dotted line). Experiments were conducted with a Superdex200 column (Pharmacia, USA), eluted with Tris-HCl buffer (20 mM Tris, 150 mM NaCl, pH 7.4) and at flow rate of 0.4 ml/min. **B.** Calibration curve was estimated using α-LA, carbonic anhydrase, ovalbumin, BSA, thyroglobulin, β-amylase and ferritin as protein reference molecular markers. aS and dimers hydrodynamic volumes were than calculated on the base of their Kd (plotted in red).

Table 3.2 Estimation of the hydrodynamic volume of aS and its dimers

Calibration markers	*MW (kDa)*
α-Lactalbumin	14
Carbonic Anhydrase	29
Ovoalbumin	44
BSA	66
β-Amylase	200
Ferritin	440
Thyroglobulin	669

Protein	*Calculated MW (kDa)*
aS	61
NN and CC monomer	63.7
DC dimer	98.9
NN, CC and NC dimer	169
NN, CC, NC oligomers	off scale

Indeed, small angle X-ray scattering studies (SAXS) showed that, at physiological conditions, aS is characterized by a radius of gyration, R_g, of about 40 Å, which is much larger than that predicted for a folded globular protein of 140 residues (15 Å) (Uversky et al., 2001). The determined volume in solution corresponds to a globular protein of 61 KDa, a result in agreement with the electrophoretic properties of aS. Dimers elutes before aS, indicating that their hydrodynamic volume is bigger than the one of aS. This result suggest that they are, as well as aS, random in solution, since a fold state would result in a more compact volume.

The secondary structure of aS and NN, CC, NC and DC under physiological buffer conditions (phosphate buffer, pH 7.4) was evaluated by far-UV CD and FT-IR. All the analyzed species are unfolded in solution and do not have appreciable secondary structure. The CD spectra, recorded between 250 and 198 nm, are reported in Fig. 3.4. All the spectra show a marked minimum at ~200 nm corresponding to random coil structure, in analogy to aS behaviour (Weinreb et al, 1996).

FT-IR spectra of all the protein species were also recorded. Deuterated proteins were dissolved in a saline buffer (20 mM Tris·DCl, 150 mM NaCl, pH* 7.2, uncorrected for isotopic effects), and the IR absorbance was recorded between 1500 and 1750 cm^{-1} to evaluate the contribute of both amide I and II bands (Fig. 3.5, bold lines). The second derivatives of the curves were used to identify the different spectral components (Fig. 3.6), which were used for the curve fitting operation (Fig. 3.5, thin lines and Table 3.3). The sum of the fitted curves is shown as a thin dashed line, which closely overlaps the experimental trace, shown as a continuous bold line.

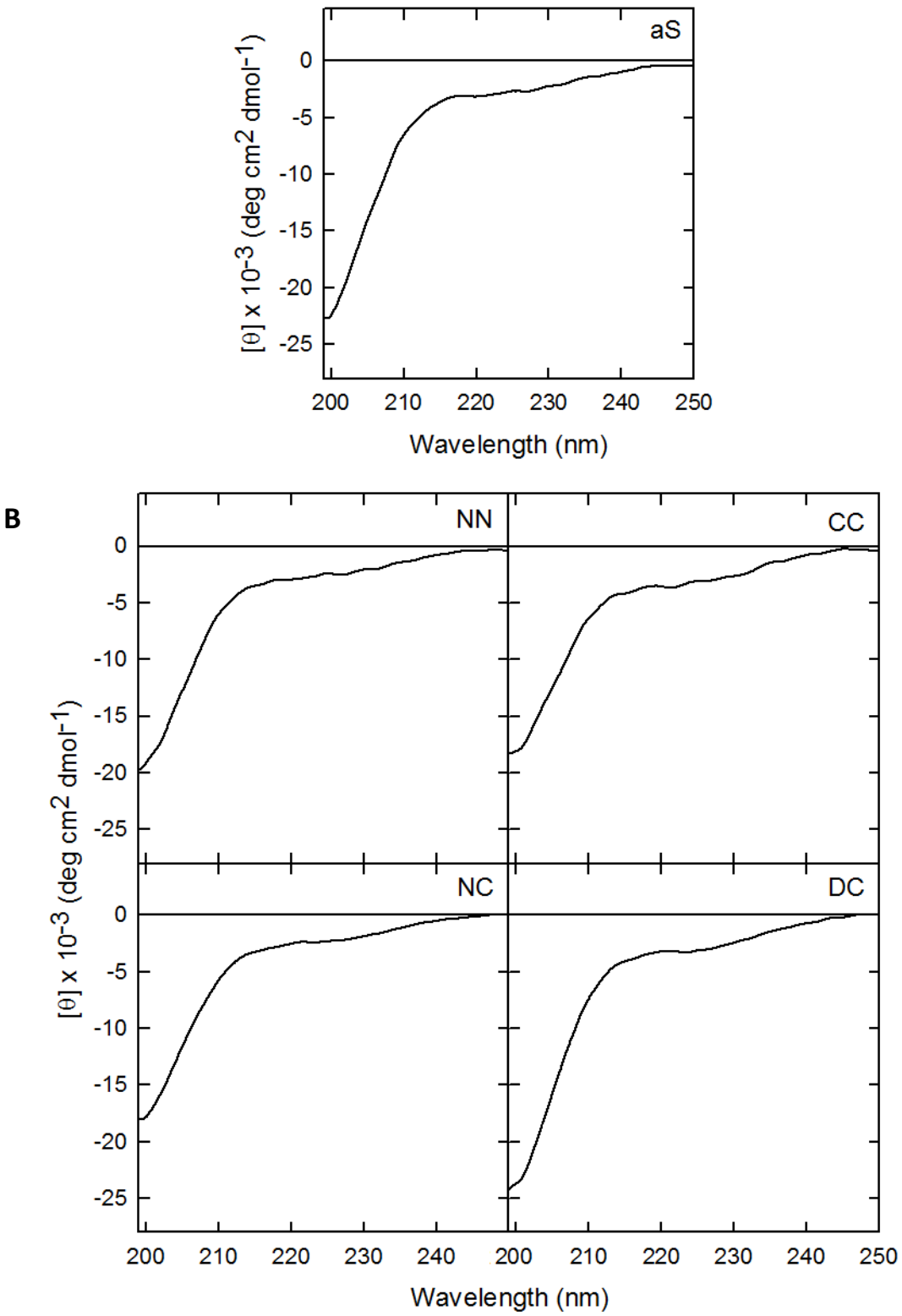
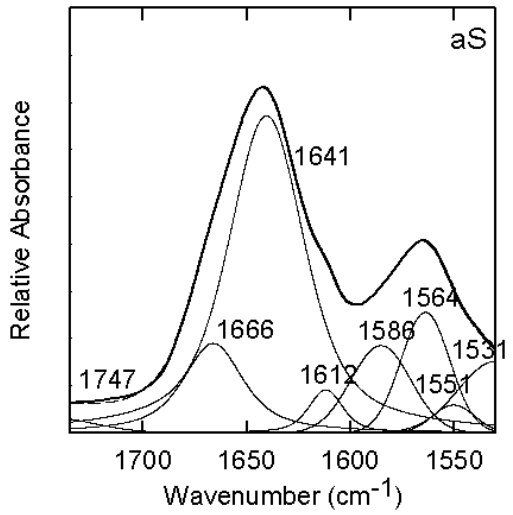


Fig. 3.4 Far UV CD of aS (A) and its dimers (B). The spectra were recorded in PBS buffer pH 7.4 at a protein concentration of 5 μM, using a quartz cuvette with 1 mm of pathlength.

A



B

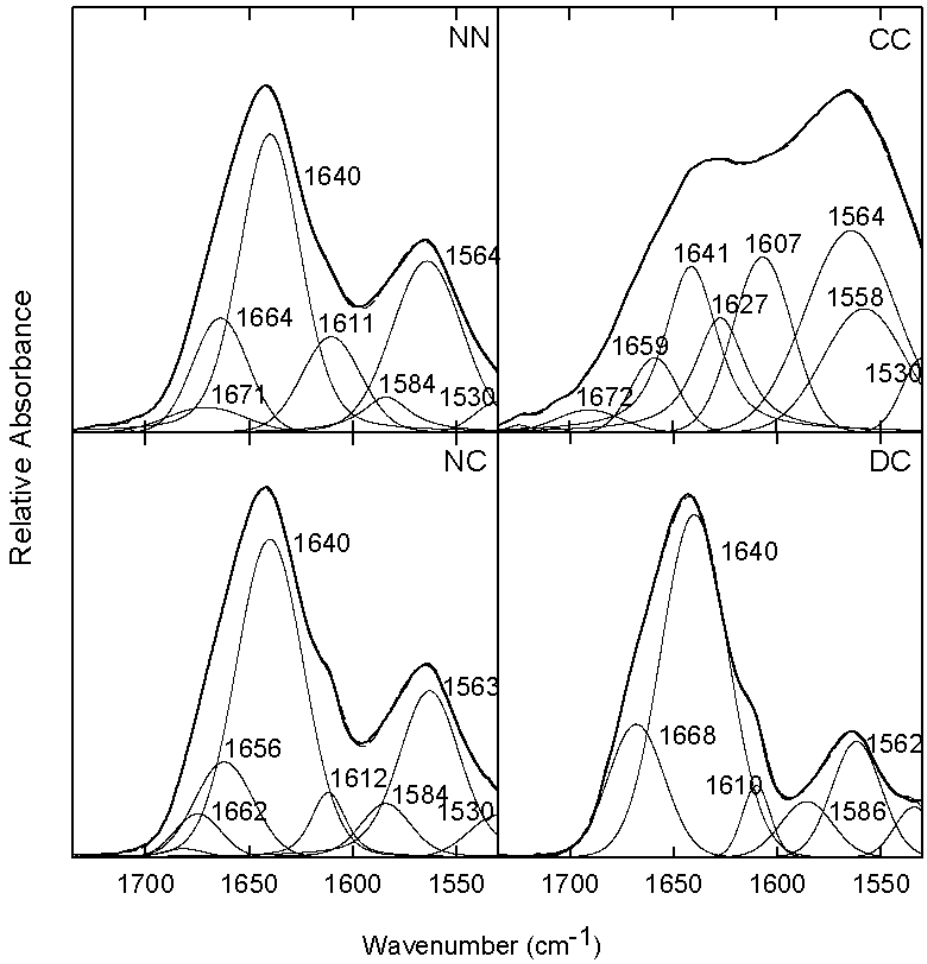


Fig. 3.5 FT-IR spectra of aS (A) and its dimers (B). Deuterated proteins were dissolved in 20 mM Tris·DCl, 150 mM NaCl pH* 7.2, uncorrected for isotopic effects. Curve fitting was performed with Gaussian and Lorentzian lineshapes. The heights, widths and positions of each band were optimized iteratively. The sum of the fitted curves is shown as a thin dashed line, closely overlapping the experimental trace, shown as a continuous bold line.

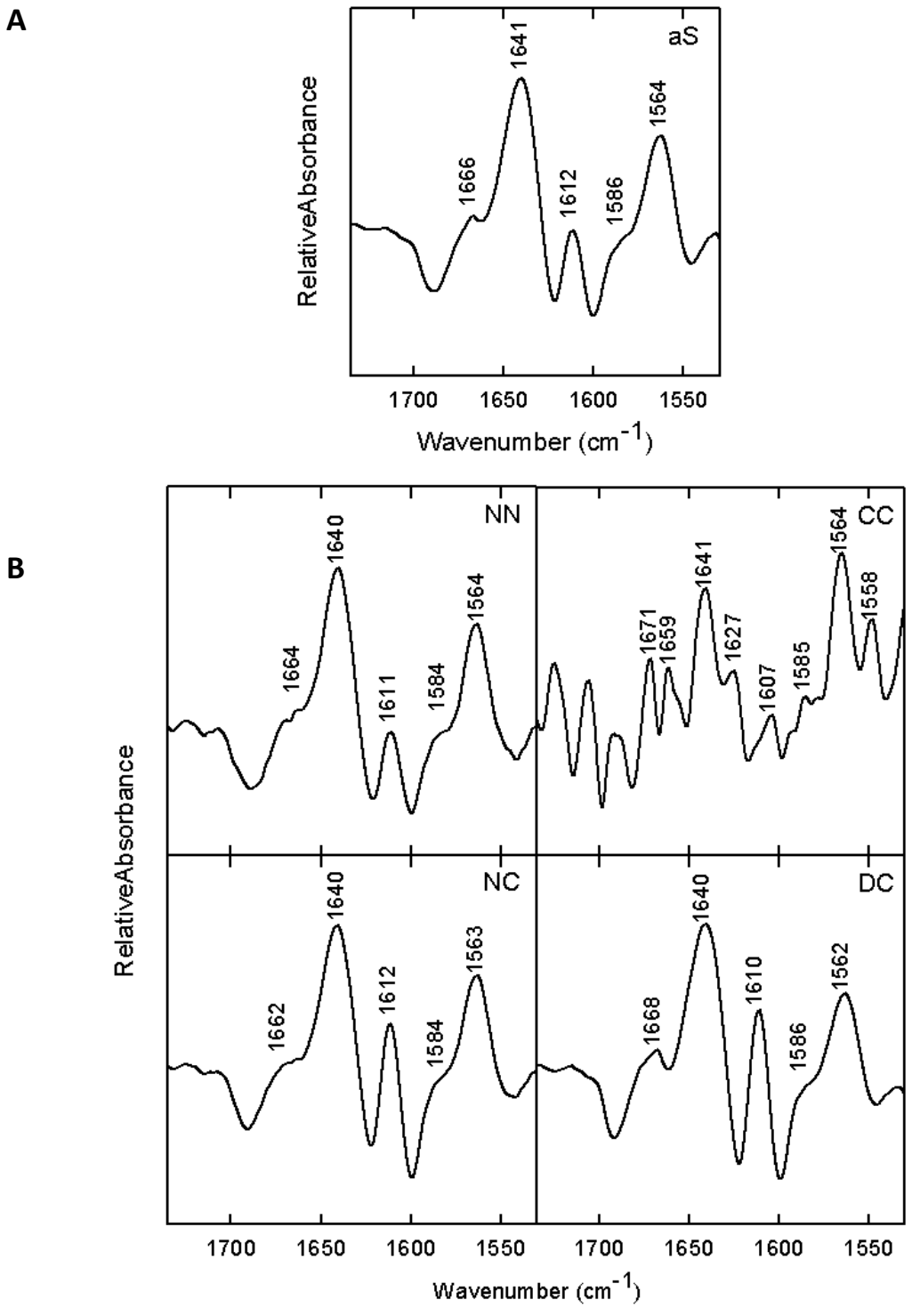


Fig. 3.6 Second derivative of FT-IR spectra of aS (A) and its dimers (B). The different spectra components band relative values indicated in figure were used as reference for the curve fitting operation.

The main component of aS, NN, NC and DC FTIR spectra is the band at 1640 cm^{-1}, which corresponds to the vibrational motions of the backbone amide moieties in random conformation. Another observable contribute of secondary structure is due to the presence of turns (band at 1656-1668 cm^{-1}). The Amide II region evidences the presence of deprotonated carboxylic moieties of Glu and Asp residues (1558-1568 cm^{-1} and 1580-1586 cm^{-1}, respectively), which significantly affects the FT-IR absorbance spectra. Of note, deuterated CC dimer showed reduced solubility in the buffer used for FTIR measurements, in respect to the other proteins. As a consequence, the preparation of the sample was repeated several times in order to obtain a clear solution and to reach an acceptable signal/noise ratio in the FTIR spectrum. At variance from the spectra of the other dimers, in the FTIR spectrum of CC dimer the Amide II band is prevalent on the Amide I one (Fig. 3.5), which indicates a larger contribute of the negatively charged Glu residues (bands at 1564 and 1558 cm^{-1}, Fig. 3.6), and of Ans/Arg residues (1607 cm^{-1}). Together with the random conformation (1640 cm^{-1}), the spectrum reveals the presence of intra-molecular β-sheet (band at 1627 cm^{-1}). The latter signals could also be linked to an aggregation process occurred during the preparation of the sample (lyophilization) or following the deposition of the protein solution in the 50 μm-thick cell.

Table 3.3. Secondary structure content of monomeric aS and its dimers as determined by FTIR spectroscopy.

Wavenumber[a] *(cm^{-1})*	*Structural Assignment*	*aS %*[b]	*NN %*[b]	*CC %*[b]	*NC %*[b]	*DC %*[b]
1530	β-sheet (Amide II)	9	3	4	5	3
1548	β-turns (Amide II)	2	-	-	-	-
1558-1568	Glu (COO^-)	11	25	40	22	14
1580-1586	Asp (COO^-)	9	5	4	7	7
1607-1610	Asn/Arg	-	12	15	12	5
1627-1636	Intra-molecular β-sheet	3	-	10	-	-
1640-1650	Random	52	38	15	42	54
1656-1668	Turns	14	17	10	12	17
1680-1689	Anti-parallel aggregated β-sheet	-	-	2	-	-

[a]Peak position of the amide I band components, as deduced by the second derivative spectra.

[b]Percentage area of the amide I band components, as obtained by integrating the area under each deconvoluted band.

To determine the presence of secondary structure with a high resolution spectroscopic technique, ^{1}H-^{15}N HSQC spectra of aS and its dimers were recorded. The spectrum of aS (Fig. 3.7) exhibits a dense cluster of cross-peaks over a narrow range, in agreement with the fact that the protein is largely unfolded at pH 7.4. As already described (Chandra et al., 2003), at 25 °C the number of visible peaks (~70) is lower than expected (135). This behaviour can be attributed to conformational exchange in the first 100 residues of aS (McNulty et al., 2006), or to fast chemical exchange between amide groups and the solvent (Croke et al., 2008).

Dimers present the same behaviour of aS in solution (Fig. 3.8 A, B, G and H). NN and CC were measured in both reduced and non-reduced state (Fig. 3.8 C and D). Reduced dimers perfectly overlap aS spectrum, and non-reduced dimers reproduce the same peak pattern, with the exceptions of the amino acids of the terminals involved in the dimer linkage (Fig. 3.8 E and F). As aspect, DC dimer has the most different spectrum respect to aS, due to the presence of two consecutive NAC regions, but noteworthy it doesn't show the presence of strong rearrangements or β-structures. Eventually, NMR measurements confirm the absence of secondary structure in all the four dimers, as for aS molecule, as previously described by CD and FT-IR.

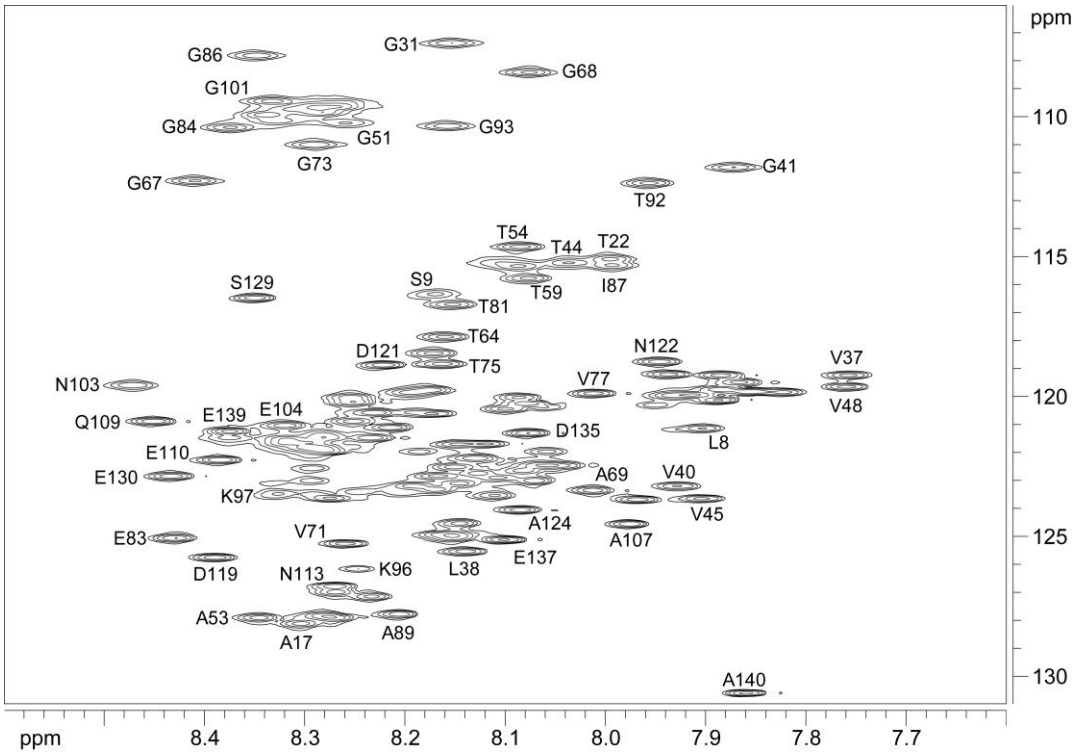


Fig. 3.7 HSQC spectrum of aS. Resonance assignments are indicated where space permits.

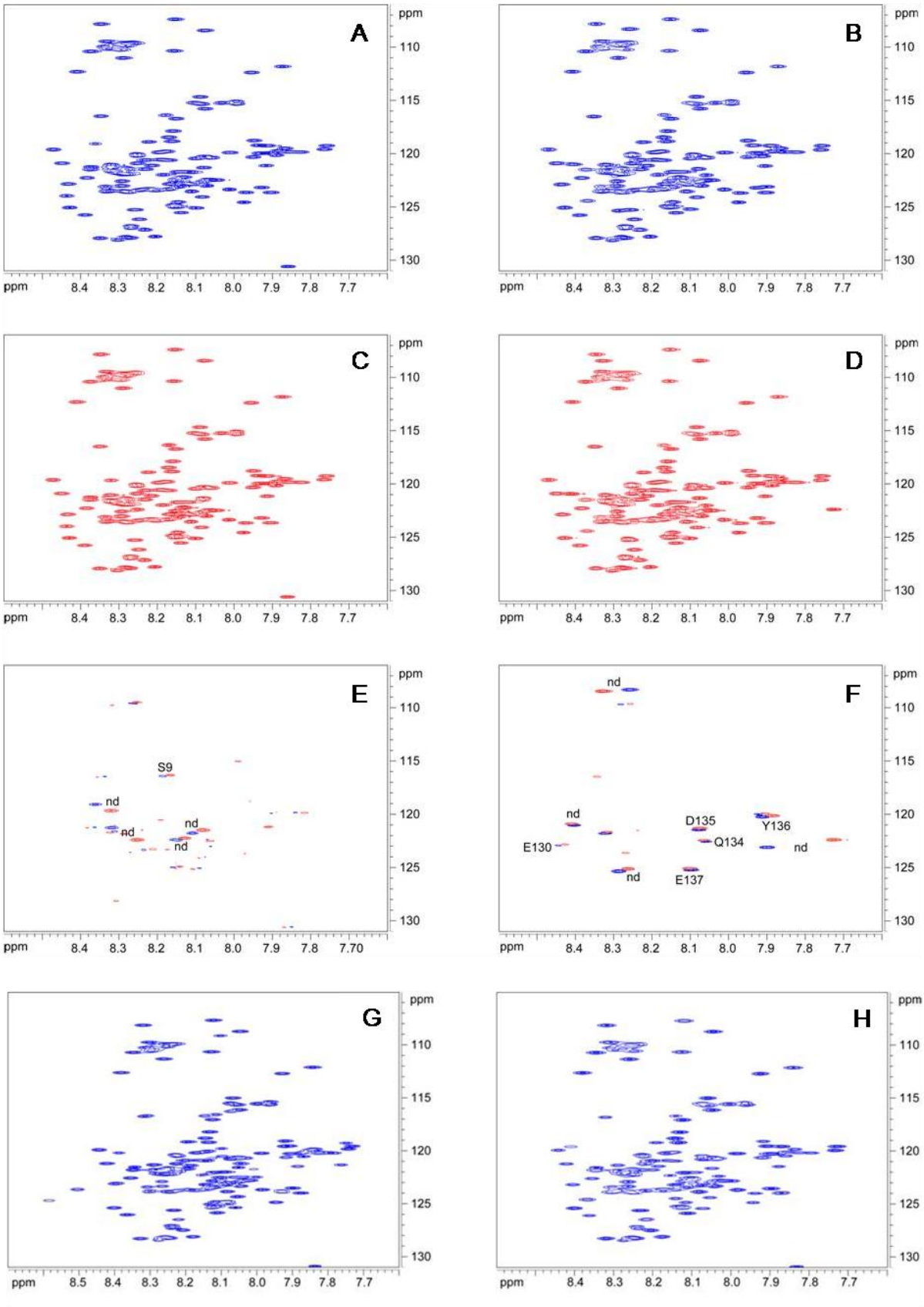


Fig.3.8 HSQC spectra of NN (A-C-E), CC (B-D-F), NC (G) and DC (F) dimers. For NN and CC, the experiments were performed in the absence (A, B) and in the presence (C, D) of 10 mM DTT. Difference maps (E, F) were obtained subtracting the data matrix obtained after 2 hours of incubation in the presence of DTT from the reference. NC and DC spectra (G, F) were obtained in absence of DTT. For each protein, spectra are reported using the same threshold.

3.3 α-Helical structure transition

The secondary structure of aS and its dimers was evaluated by far-UV CD in the presence of different concentrations of SDS. aS is unfolded in the absence of SDS (Fig. 3.4). aS CD spectra were recorded in the presence of increasing concentration of SDS (up to 2 mM) between 250 and 198 nm (Fig. 3.9 A). SDS induces aS to adopt α-helix conformation, as shown by the appearance of the two typical minima at 222 and 208 nm. The presence of an isodichroic point at 203 nm in the titration experiments (0-2 mM SDS) suggests a simple two-state conformational transition between random coil and α-helix. All the dimers present the same trend (Fig. 3.9 B-E).

To further analyze the structural transition of aS in the presence of SDS, in Fig. 3.6 F, the ellipticity of aS, NN, CC, NC and DC at 222 nm is plotted as a function of the SDS concentration. The decrement of ellipticity could be fitted by a sigmoid curve which indicates the complex dependence of aS structural properties on SDS. The sigmoid curve has similar features among the protein species: initial addition of SDS (0.2 mM) results in minimal changes in protein secondary structure. Further increases in SDS from 0.2 to ~1 mM result in the induction of α-helical structure. Up to 1 mM SDS, protein secondary structure remains practically unperturbed. aS reaches a plateau at an ellipticity of -24 $mdeg \cdot cm^2 \cdot dmol^{-1}$ and remains constant increasing SDS concentration. NN, CC, and NC dimer display a similar trend with plateau between -22 and -26 $mdeg \cdot cm^2 \cdot dmol^{-1}$. DC dimer ellipticity at 222 nm is approximately -17 $mdeg \cdot cm^2 \cdot dmol^{-1}$ at 1 mM SDS concentration and slowly decreases to -23 $mdeg \cdot cm^2 \cdot dmol^{-1}$ at 5 mM SDS, indicating a slow transition towards α-helical structure.

SDS is commonly known as a protein denaturing agent. It is well-established that the detergent can induce structure formation in aS, effectively mimicking phospholipids that have similar effects on the protein (Chandra et al., 2003; Ferreon and Deniz, 2007). All the effects of SDS on the different protein species at SDS concentration lower than 1 mM may be attributed mainly to the monomeric form of the detergent. This observation is consistent with the reported critical micelle concentration (CMC) of SDS in the specific buffer saline concentrations used, i.e. ~1 mM (the presence of low protein concentration in solution do not affect the SDS CMC) (Helenius et al., 1979; Ferreon and Deniz, 2007). Monomeric SDS molecules interact with aS and its dimers and induce structural formation in a highly cooperative fashion, probably due to a specific binding. The artificial constraints, which limit the dimers' movement, do not affect the ability of the dimers to interact with SDS. DC dimer, which stabilized the maximum of α-helical structure presence at higher SDS concentration, shows a slightly different behaviour respect to aS and the other dimers, probably due to the nearness of the consecutive NAC regions.

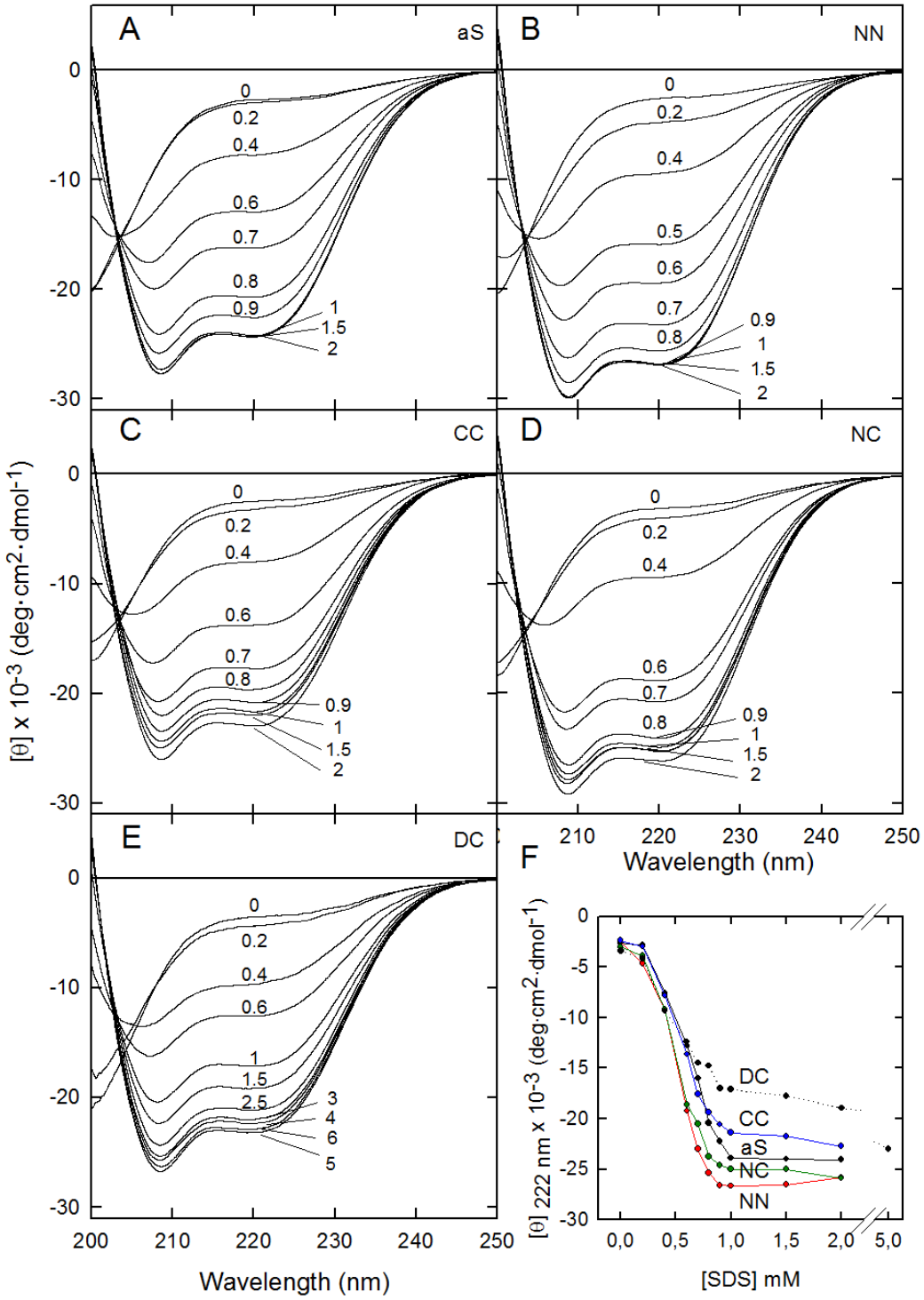


Fig. 3.9 Secondary structure changes induced by SDS. Far UV CD of aS and NN, CC, NC and DC (respectively **A**, **B**, **C**, **D** and **E**) in the presence of increasing concentration of SDS. The spectra were recorded in PBS buffer pH 7.4 at a protein concentration of 20 μM for aS, NN, CC and NC, and 6.7 μM for DC using a quartz cuvette with 1 mm of pathlength. The concentrations (mM) of SDS are indicated in the figure. **F.** Ellipticity at 222 nm ($[\theta]_{222\ nm}$) as a function of SDS concentration. aS and NN, CC, NC and DC are represented respectively by black, red, blue, green and black dotted lines.

3.4 Aggregation studies

3.4.1 Aggregation kinetic

Aggregation studies were carried on incubating the proteins of interest at 1 mg/ml in PBS buffer, pH 7.4, at 37°C, shaking at 500 rpm. Aliquots of aS aggregation samples were withdrawn at different time of incubation and analyzed by several biochemical and biophysical methods.

The formation of amyloid-like fibrils in the samples was monitored by Thioflavin-T (ThT) binding assay. All the dimers are able to form amyloid fibrils as aS, showing an increase of ThT fluorescence at 485 nm. The formation of amyloid fibrils was confirmed by TEM imaging (see section 3.4.2). Moreover, in order to monitor the kinetic of fibril formation, the aggregation of aS and NN, CC, NC was followed in parallel by ThT assay and by evaluating the amount of protein present in fibrils on time (Fig. 3.10). This calculation was indirectly performed measuring the absorbance of the supernatant obtained after centrifugation of each aliquot (Wood et al., 1999). By subtracting the amount of soluble protein from the total protein present in the aliquot, the percentage of protein in the fibrillar pellet was estimated. To minimize errors due to the loss of material during handling and to improve the measurement of each aliquot, this aggregation was carried on at 2.5 mg/ml concentration. Both the techniques evidenced the same trend. CC showed the same kinetic of aS in the formation of fibrils. Instead, NN and in particular NC present a slower kinetic than that of aS.

The kinetic of fibrillation directly depend on protein concentration (Uversky 2001). aS and dimers aggregation was further examined during incubation at a different protein concentration (1 mg/ml) and analyzed by ThT assay (Fig 3.11 A). In this case three independent aggregation mixture of aS and dimers were analyzed, and also DC aggregation was studied. Decreasing the concentration of the proteins, the speed of fibril formation decreases. In the case of NN and NC dimers, after 14 days the fluorescence signal is still low, indicating a prolonged lag phase respect to aS, CC and DC, in agreement with the experiments at higher concentration. Fig. 3.11 B refers to the normalized spectra of aS, CC and DC shown in Fig 3.11 A . aS reaches the fluorescence plateau after 2 days at a protein concentration of 2.5 mg/ml (Fig. 3.10), using a protein concentration of 1 mg/ml the protein reaches the plateau after 10 days, with 2 days of lag phase. CC and DC have the same lag time of 2 days but reach the fluorescence plateau after 7 and 5 days, respectively. The significant differences between these three kinetic lay on the slopes of the fitted sigmoid curves, which are 18.1 ± 7.2, 41.5 ± 5.8 and 51.2 ± 6.4 for aS, CC and DC (calculated as the curve tangent at the flex position). While the lag period corresponds to the time required for the nucleation of the protein species and the formation of oligomers, the slope depends on the kinetic of maturation of the fibrils, i.e. on the affinity of monomers to the growing fibrils (Fink, 2006).

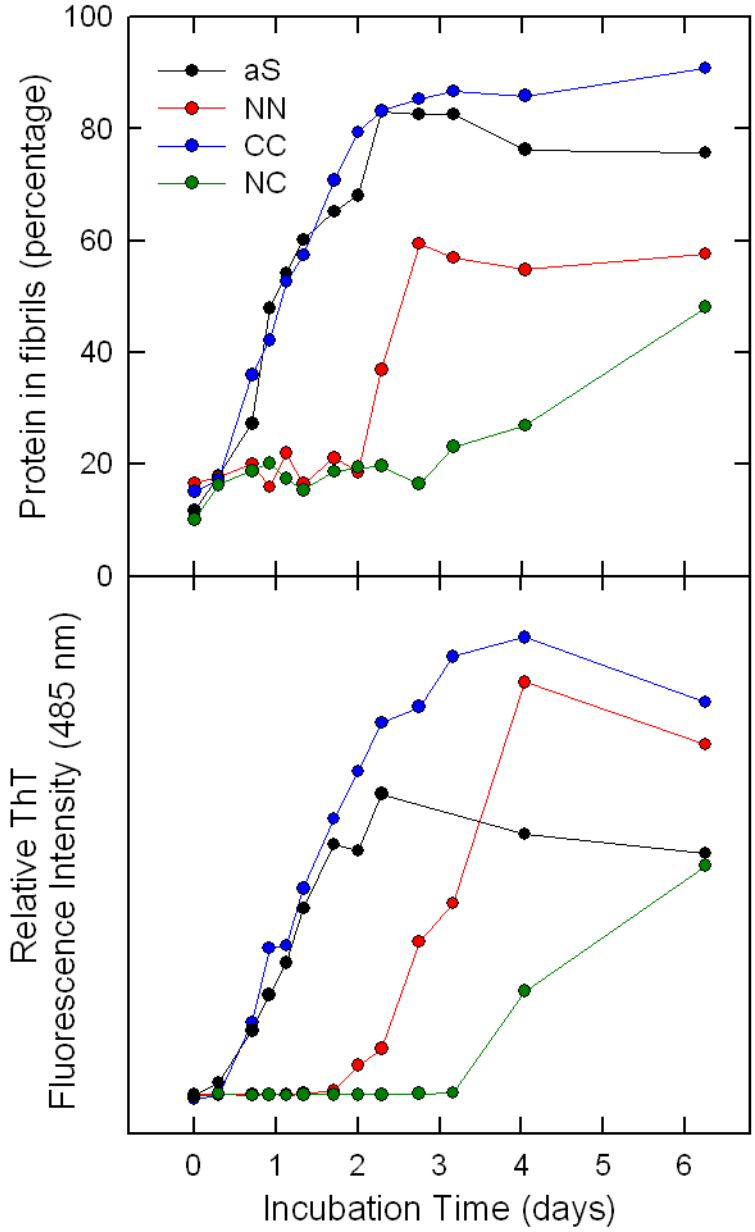


Fig. 3.10 Time-course analysis of the aggregation process of aS and NN, CC, NC followed by calculation of the percentage of protein in fibrils and by ThT fluorescence assay. aS and NN, CC and NC were incubated for 6 days at a concentration of 2.5 mg/ml (PBS buffer, 37°C, shaking at 500 rpm). Aliquots were withdrawn from the solution for ultracentrifugation and fluorescence assay. The percentage of protein into the fibrils was determined indirectly by absorbance measurement of the supernatant after centrifugation (upper panel). The aggregation was also monitored by ThT assay (lower panel). aS NN, CC and NC are represented respectively by black, red, blue and green and circles.

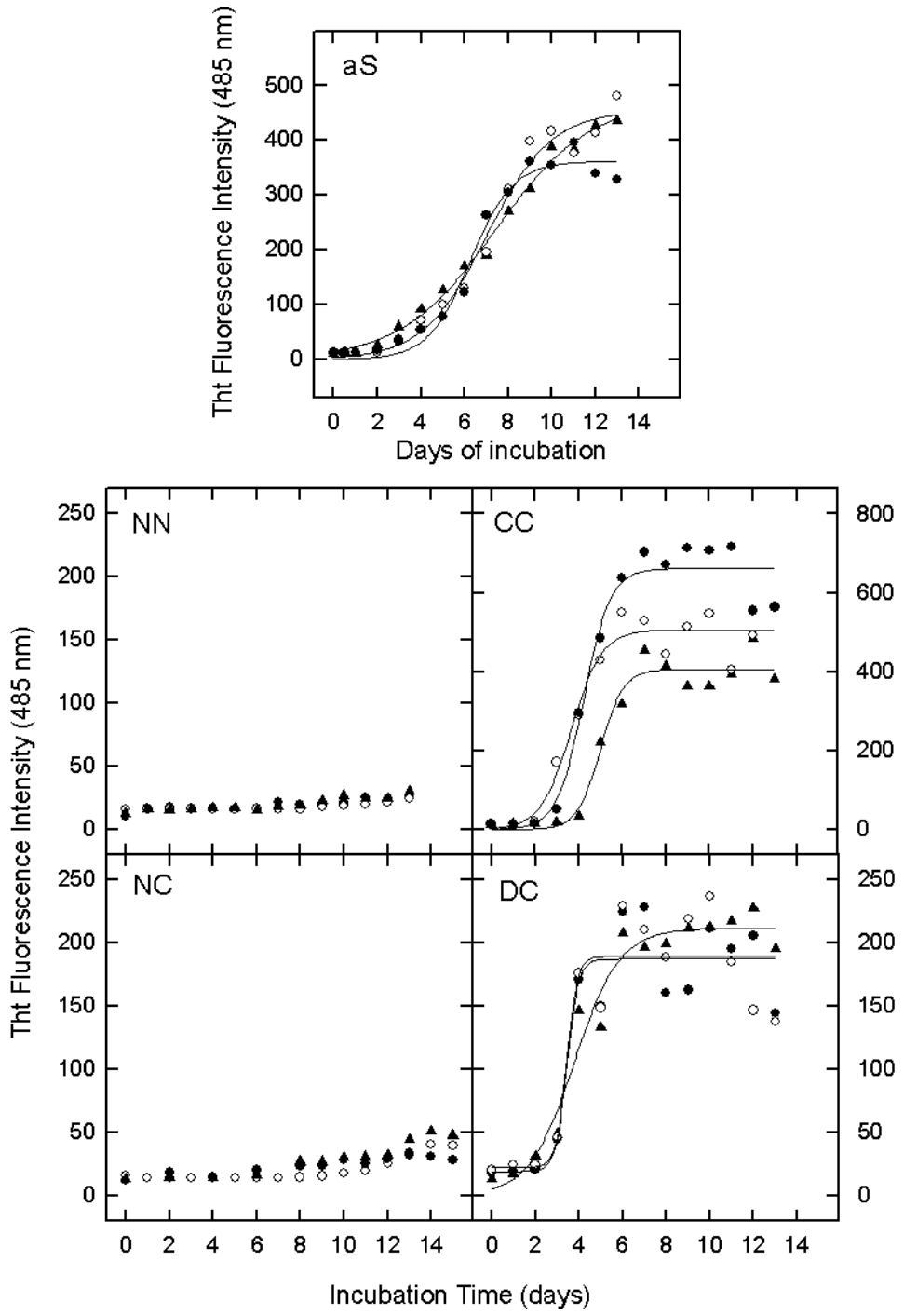


Fig. 3.11 A Time-course analysis of the aggregation process of aS, NN, CC, NC and DC followed by ThT fluorescence assay. The aggregation processes were conducted at a protein concentration of 1mg/ml. Three independent aggregation experiments were conducted for each ptotein.

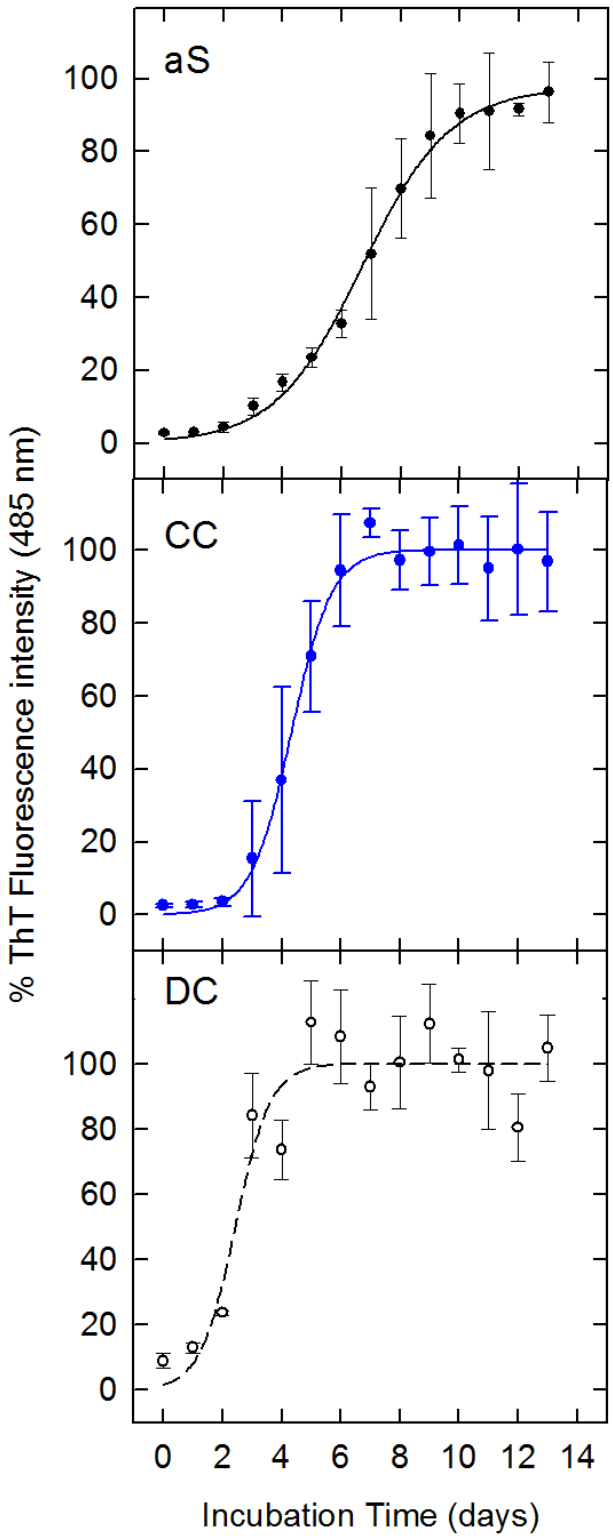


Fig. 3.11 B Time-course analysis of the aggregation process of aS, CC and DC followed by ThT fluorescence assay (normalized). The aggregation processes were conducted at a protein concentration of 1mg/ml. Fluorescence intensity of ThT is reported as percentage of the plateau of emission intensity corresponding to each curve. aS, CC and DC are represented respectively by black, blue and white circles. Error bars were calculated from three independent aggregation experiments.

Native gel electrophoresis was also used to monitor the aggregation and investigate the presence of stable aggregates during the process (Fig. 3.12). This method constitutes a good tool to reveal the presence of covalent and non-covalent oligomers and aggregates. The band relative to the aS monomer decrease its intensity during the aggregation in favour of bigger protein species present in the wells. The same trend is observed for the other dimers, and in particular, for DC dimer the decrease of intensity of the dimer is appreciable already from the first day after incubation. Both aS and NN and CC dimers show a band between the stacking and the running gel, which correspond to protein species of elevated MW. CD and fluorescence measurements of aliquots withdrawn at time zero and after a day of aggregation, do not show that the samples have acquired any secondary structure and they do not bind ThT. This suggests that these species with high MW, produced in the first stages of aggregation, could be amorphous aggregates (De Franceschi et al., 2011). Moreover, for CC a band corresponding to protein with reduced Cys-Cys bond is observable. Noteworthy, no stable oligomeric species (dimers, trimers, tetramers) were evidenced for neither aS nor dimers by the native electrophoresis during the incubation time that precedes amyloid formation.

A conformational analysis of the four protein species was carried out at different stages of aggregation by using CD measurements. At the beginning of the incubation, aS and all the dimers are unfolded (as shown in section 3.2). During incubation, aS undergoes a conformational rearrangement and acquires β-sheet structure (fig. 3.13 A). After 7 days of incubation, the spectrum is dominated by the band at 217 nm, which indicates the presence of a β-sheet structure. Similar trend is observed for NN, NC and DC, which reach the same minimum at different incubation times (fig. 3.13 B). In agreement with gel electrophoresis and ThT fluorescence measurement, DC shows the faster kinetic of β-sheet conversion. CC displays a decrease of random content without showing the minimum at 217 nm. To avoid the contribute of non-aggregated material, fibrils of CC were isolated by ultracentrifugation. The obtained spectrum doesn't evidence any minimum at 217 nm (data not shown). The absence of the minimum could be due to contributes and interferences of the negative charges which are present in a big cluster in the C-terminal regions of aS, and that are, in CC, consecutive. However, spectra underline the loss of random coil conformation, which is required for the amyloid structure formation (Uversky, 2001). β-Sheet transition of CC was confirmed by FT-IR analysis on the aggregate material after 15-20 days (see section 3.4.3).

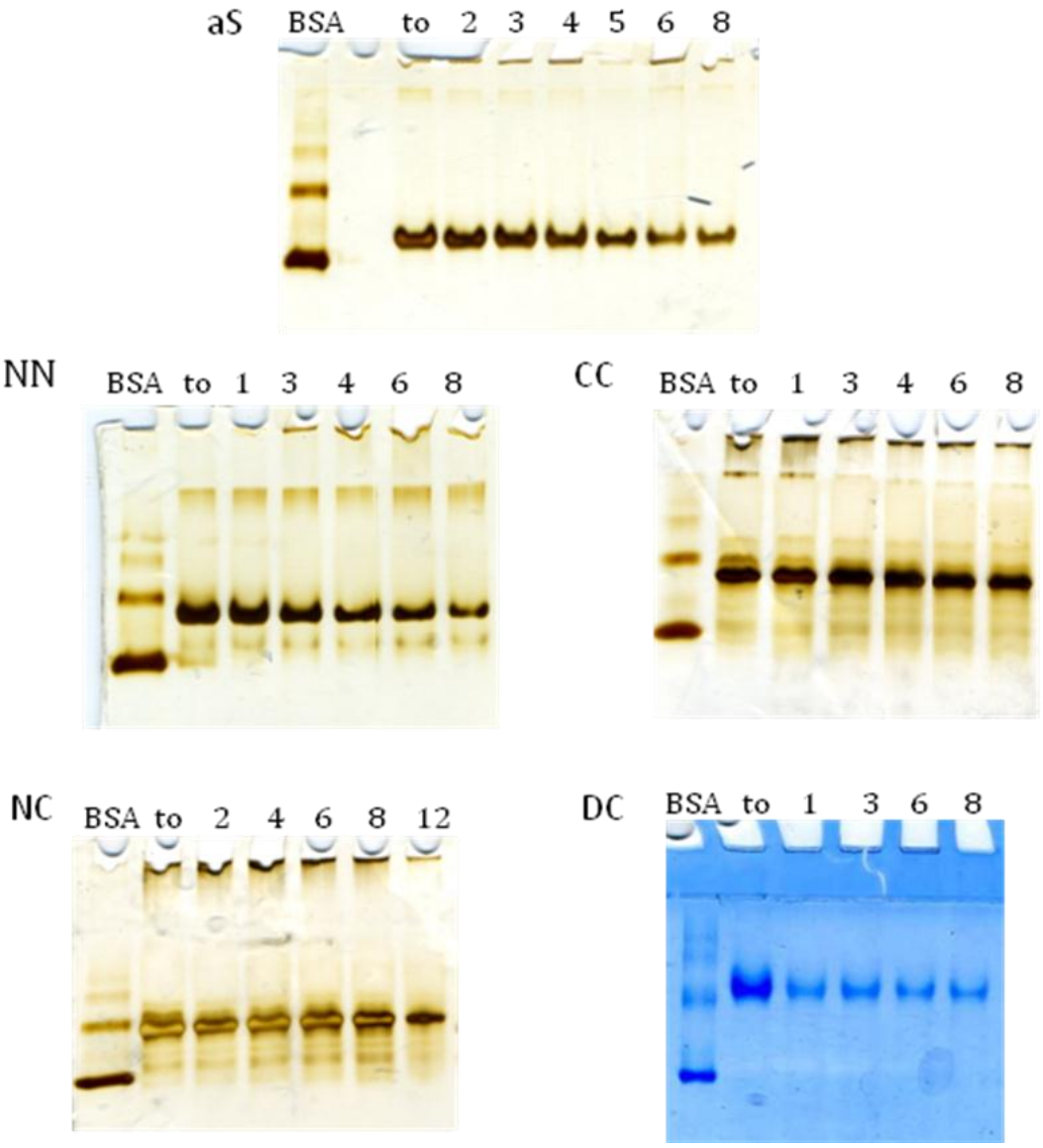
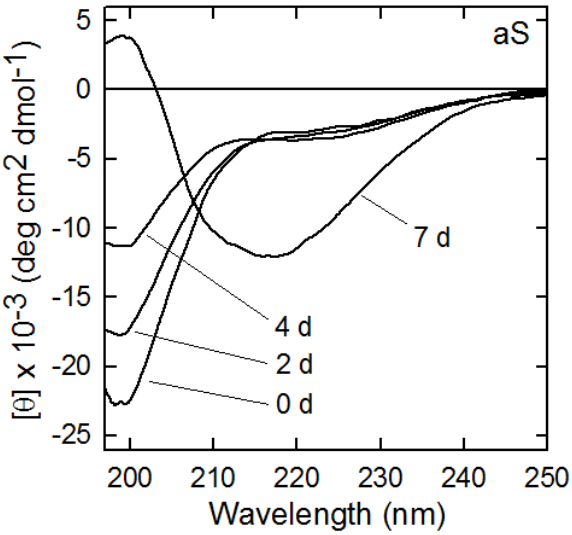


Fig. 3.12 aS and dimer aggregation monitored by native gels. Aliquots of aS and NN, CC, NC and DC samples were withdrawn during aggregation and analyzed by native electrophoresis. Numbers on each lanes refer to the days of incubation. The gels were stained by silver (aS, NN, CC and NC) or by Coomassie blue staining (E). BSA was used as marker.

A



B

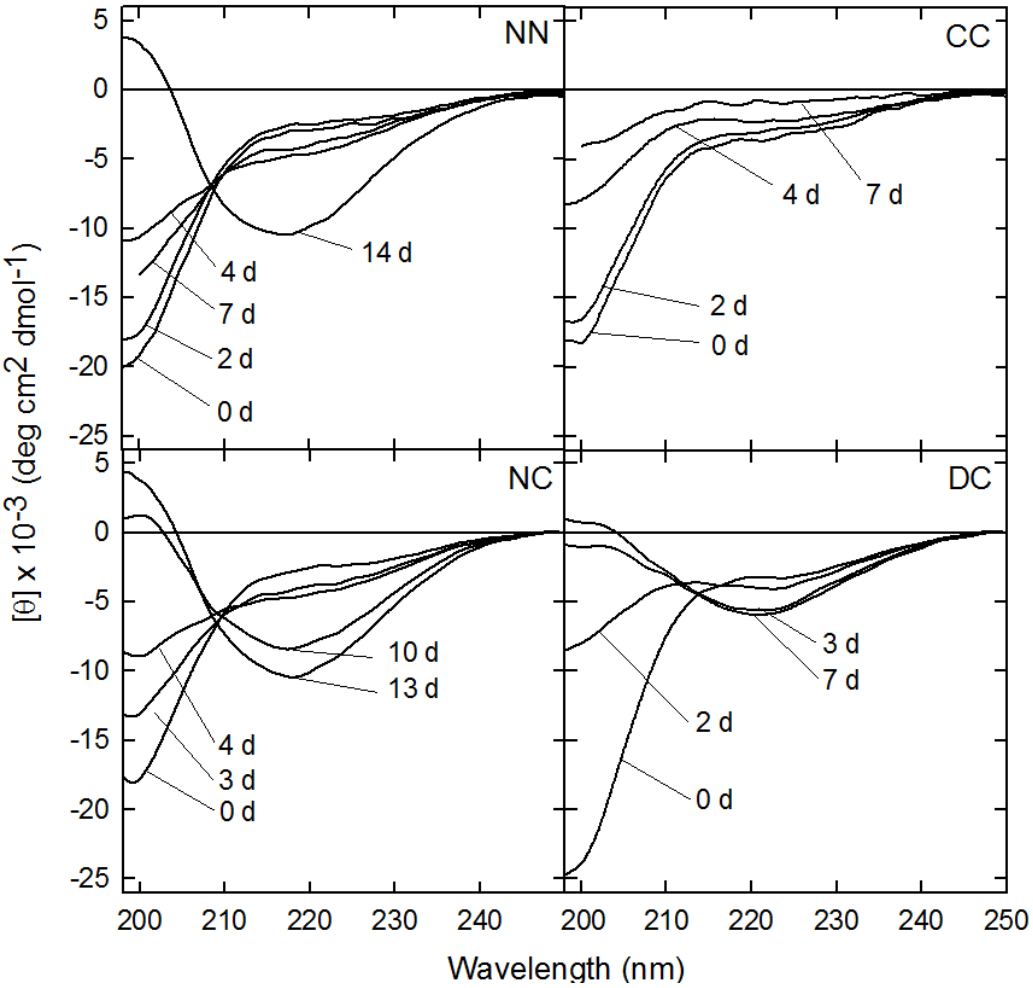


Fig. 3.13 Far UV CD of aS (A) and its dimers (B) during aggregation. The evolution of secondary structure of aS and its dimers during aggregation was monitored by far-UV CD. The spectra were recorded in PBS buffer pH 7.4 at a protein concentration of 0.05-0.15 mg/ml, using a quartz cuvette with 1 mm of pathlength.

3.4.2 Analysis of the aggregates morphology

Fibril morphology of dimers was investigated by electron microscopy (EM) and atomic force microscopy (AFM) imaging. As anticipated, all the dimers form amyloid fibrils that are apparently indistinguishable from the aS ones. Fibrils obtained from aS, NN, CC and NC are long and straight, and could have a length of ~200 nm to ≥2 μm. Fibrils obtained from DC are instead shorter (they reach a maximum length of 450 nm).

It was observed that, even after prolonged incubation (up to a month), fibrils derived from aS dimers display a lower organization level in the fibril overall structure. aS fibrils are made by two, or more, filaments and show a periodic twist. Fibrils derived from the dimers do not display this ultra-structure organization of filaments and are single filaments (Fig. 3.13). AFM studies confirmed that aS fibrils are formed by twisted filaments, and display a periodicity of 82 ±7 nm and an average height of 7.5 nm. Single aS filaments have an average height of 5.44 ±0.59 nm, while mature fibrils, composed by more filaments, have an average height of 7.47 nm ± 0.82. The diameters of NN, CC and NC filaments have a comparable measure of 5.37 ±1.09 nm. EM images show larger diameters, accordingly with literature data (Fink, 2006), and are about 9.4 ±1.3 nm for single filament. AFM measurement of DC fibrils was not performed for technical problems and the analysis is now under investigation. EM images show a diameter of 11.85 ±1.35 nm for DC, suggesting a larger dimension.

These results suggest a specific role of the non structured regions, i.e. regions not involved in fibril core, for the assembly of two filaments to mature, twisted fibrils. The terminal side chains, not involved in the core β-structure of the fibril, are responsable for the inter-molecular interactions among different filaments. NN, CC and NC have limited degree of freedom respect to aS. Indeed the intra-molecular re-arrangement, that allows the interaction of the two NAC regions, is limited by the fact that the N and C terminals of the proteins are linked for the formation of the dimer. This conformational limit (constraint) could affect the possibility to create intra-filaments interactions. Nevertheless, although DC has free terminal regions as aS, it does not form twisted fibrils. This could be due to higher dimension of the fibril core or to the instability of the fibrillar core (see section 3.4.4).

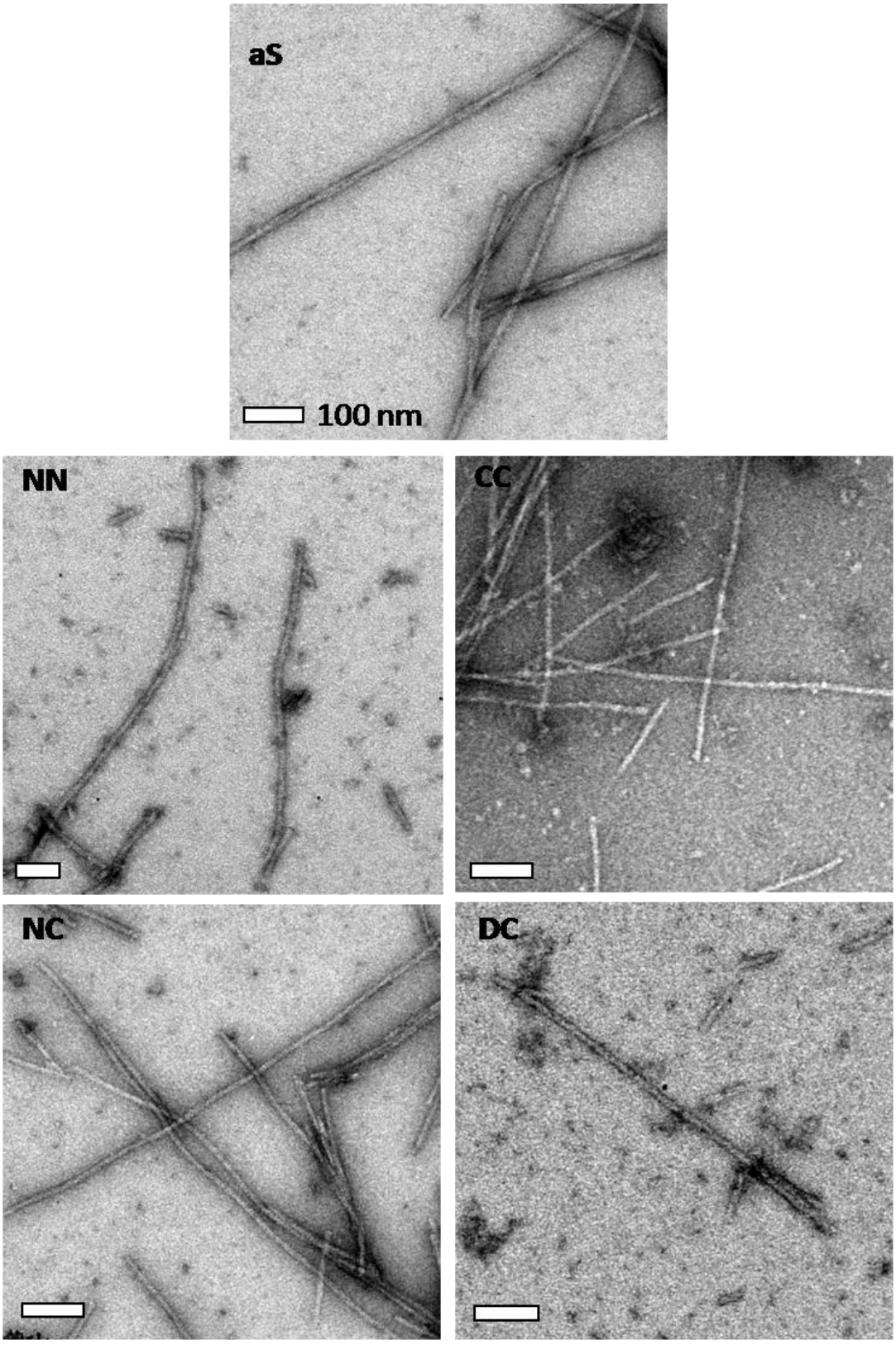


Fig. 3.14 TEM imaging of aS and dimer fibrils. TEM pictures relative to aS and dimers aggregation samples after 14 days of incubation are reported. Barrels indicate 100 nm.

3.4.3 Conformational analysis

FT-IR spectra of the proteins after prolonged incubation were recorded to evaluate the type and content of secondary structure in the fibrils. The protein species were previously deuterated and aggregation was induced by incubating them in 20 mM Tris·DCl, 150 mM $NaCl_2$ pH* 7.2 for a month. Therefore, fibrils were separated from non-aggregated species by ultracentrifugation, in order to evade their contribute and obtain suitable concentrations for the measurement. The IR absorbance was recorded between 1500 and 1750 cm^{-1} to evaluate the contributes of both Amide I and II bands (Fig. 3.15, bold lines). The second derivatives of the curves were used to identify the different spectral components (Fig. 3.16), which were used for the curve fitting operation (Fig. 3.15, thin lines). The sum of the fitted curves is shown as a thin dashed line, which closely overlaps the experimental trace, shown as a continuous bold line. Table 3.4 reports the contributes of fitted bands calculated on percentage of the total area. This extrapolation is useful for the formulation of consideration on the secondary structure content of fibrils, but it should be considered an approximation, since the spectra do not reach a minimum at 1530 cm^{-1}. As expected for amyloid fibrils, the spectra of all the protein species contain the band at 1612-1619 cm^{-1}, which corresponds to the vibrational motions of the backbone amide moieties in aggregate β-sheet structure. A further observable contribute of secondary structure among all the fibrils is due to the presence of turns (band at 1659-1668 cm^{-1}), and only in the case of aS fibrils, β-turns contribute is appreciable also in the Amide II regions. As observed for the spectra of monomeric proteins in solution (Fig. 3.5), the FT-IR absorbance is strongly affected by the contribute of deprotonated carboxylic moieties of Glu and Asp residues (1560-1568 cm^{-1} and 1580-1586 cm^{-1}, respectively) in the Amide II region. CC dimer fibrils show the strongest contribute of the band, followed by NN and aS fibrils. Apart from NC, all the other proteins show an increase of this band respect to non-aggregated proteins (see Table 3.3), indicating that the Glu residues, which are mostly present at the C-terminal of the protein, increase their interaction with the surrounding environment after the structural transition.

Even if CC fibril seems to have a lower contribute of aggregate β-sheet structure (13%), the strength of the bonds that held together the cross β-sheet structure of CC fibrils is higher, because such species vibrate at lower wavenumbers with respect to the other dimers fibrils (1612-1613 rather than 1616-1619 cm^{-1}). The same consideration could be done for aS fibrils. Moreover, another similarity between aS and CC fibrils secondary structure content is that they contain significantly less random structure than the other dimers and conversely show a significant contribute of intra-molecular β-sheet structure (band at 1627-1636 cm^{-1}). Overall, aS and CC fibrils seems to be more rigid and packed than NN, NC and DC fibrils.

Table 3.4. Secondary structure content of aS and dimers fibrils after 30 days incubation as determined by FTIR spectroscopy.

Wavenumber[a] *(cm^{-1})*	*Structural Assignment*	*aS %*[b]	*NN %*[b]	*CC %*[b]	*NC %*[b]	*DC %*[b]
1548	β-turns (Amide II)	26	-	-	-	-
1560-1568	Glu (COO^-)	17	33	48	3	25
1580-1586	Asp (COO^-)	4	-	-	3	-
1612-1619	Aggregated β-sheet	23	18	13	32	35
1627-1636	Intra-molecular β-sheet	13	-	18	-	-
1640-1650	Random	-	33	11	38	27
1659-1668	Turns	14	15	9	22	9
1680-1689	Anti-parallel aggregated β-sheet	3	1	1	2	4

[a]Peak position of the amide I band components, as deduced by the second derivative spectra.

[b]Percentage area of the amide I band components, as obtained by integrating the area under each deconvoluted band.

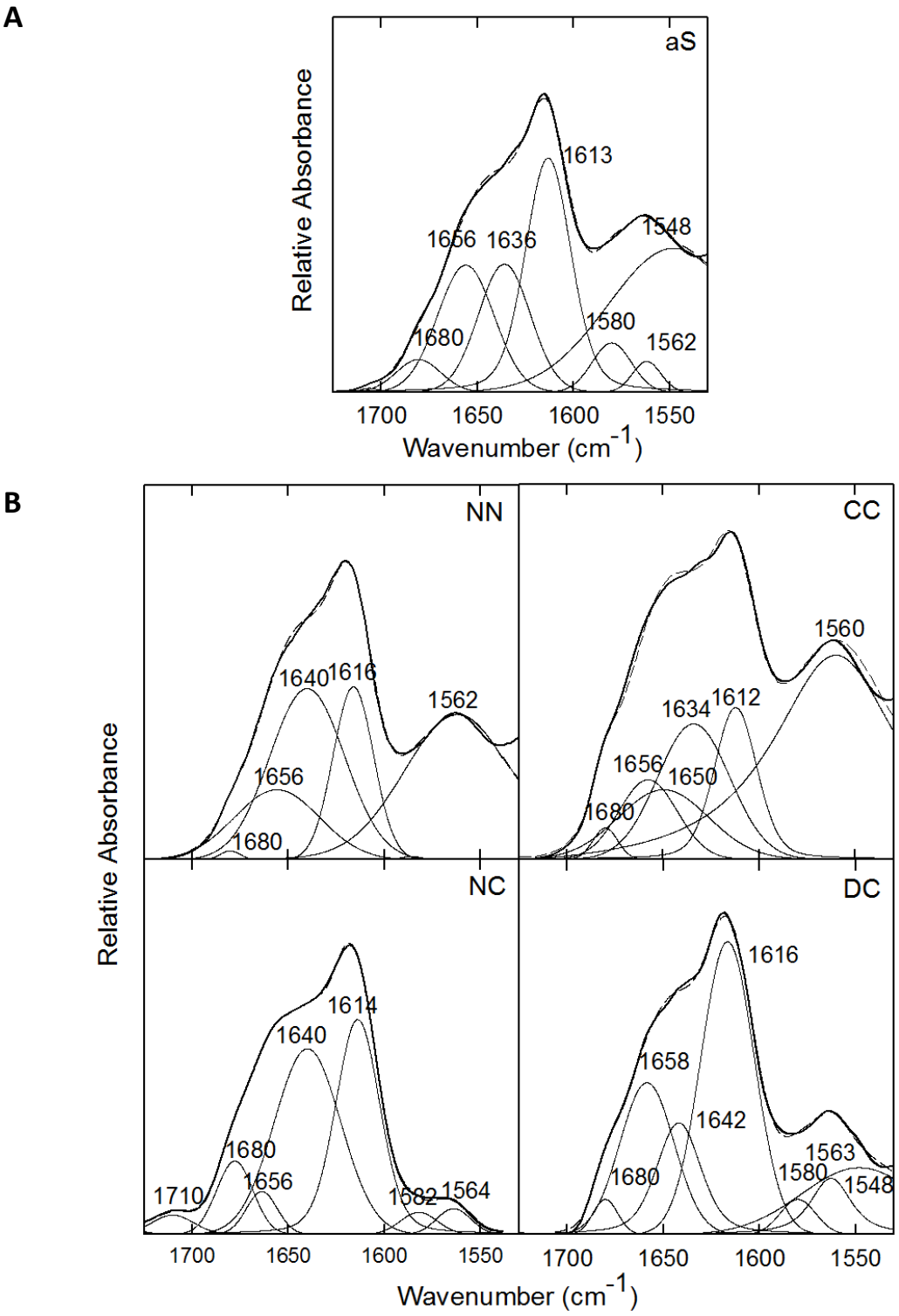


Fig. 3.15 FT-IR spectra of aS fibrils (A) and dimer fibrils (B). Protein fibrils obtained after one month incubation were dissolved in 20 mM Tris·DCl, 150 mM NaCl pH* 7.2. Curve fitting was performed with Gaussian and Lorentzian lineshapes. The heights, widths and positions of each band were optimized iteratively. The sum of the fitted curves is shown as a thin dashed line, closely overlapping the experimental trace, shown as a continuous bold line.

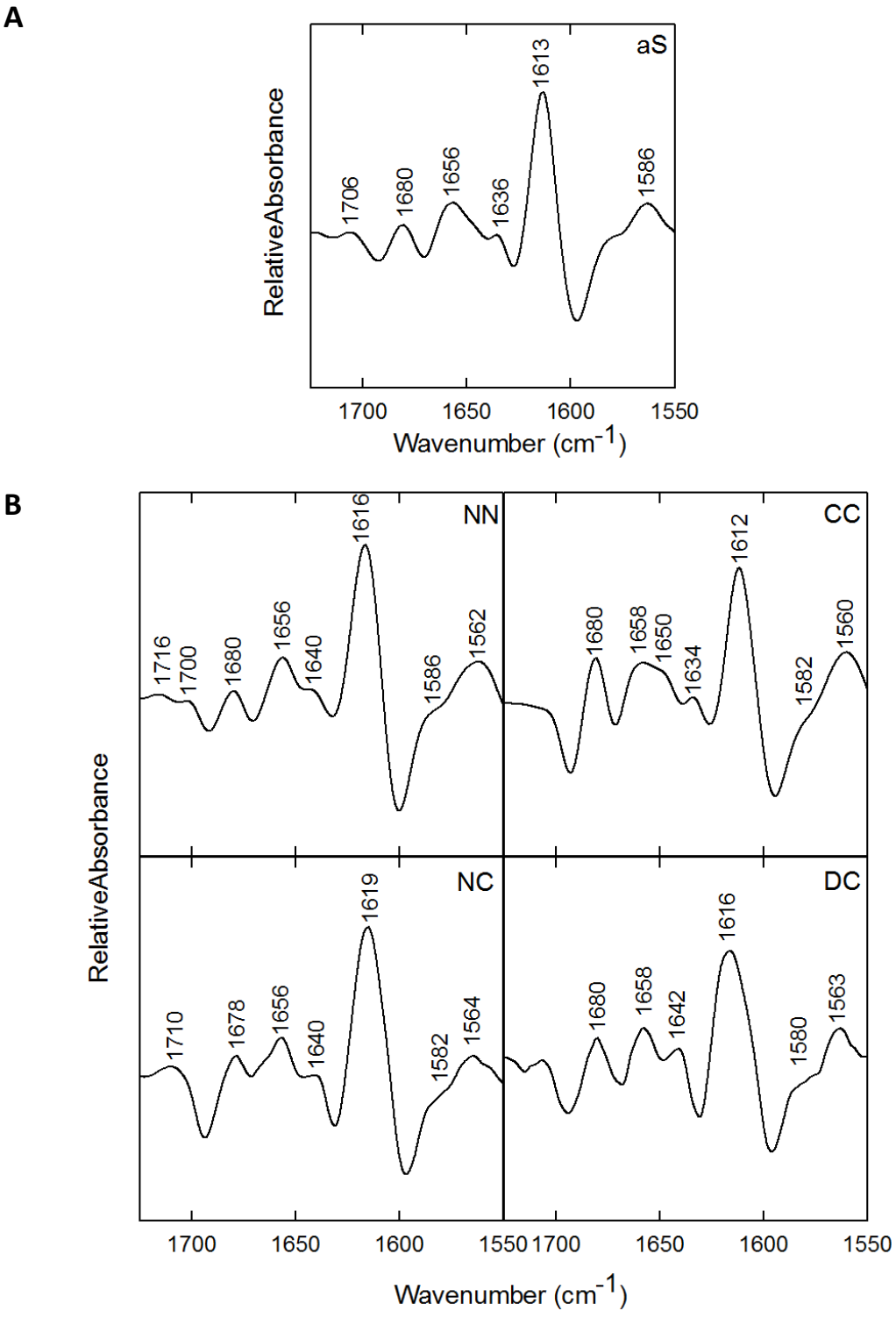


Fig. 3.16 Second derivatives of FT-IR spectra of aS fibrils (A) and dimer fibrils (B). The different spectra components and relative band values indicated in figure were used as reference for the curve fitting operation.

3.4.4 Proteolytic mapping

To directly define the region(s) of the proteins involved in the core of the fibrils, an approach based on proteolytic digestion and mass spectrometry was used. This strategy relies on the consideration that regions of the protein normally available to proteases exhibit limited accessibility when involved in β-sheet bounds (Fontana et al., 2004; Polverino de Laureto et al., 2006). Proteinase K (PK) (Ebeling et al., 1974), a particularly voracious protease, which displays broad substrate specificity, and trypsin (T), which specifically hydrolyzes peptide bonds containing basic residues at the C-terminus, were used. Proteolysis experiments were conducted on 15 days fibrils of each aS species and the proteolysis patterns were compared.

Fig. 3.17 reports the RP-HPLC profiles of aS fibrils digested with PK (upper panels) and trypsin (lower panels). Soluble and insoluble fragments obtained after proteolysis were separated by ultracentrifugation. The supernatants, which contain the soluble peptides, were loaded directly onto a C18 column (right panels). The pellets obtained after ultra-centrifugation contain the insoluble fractions (left panels), the peptides from those regions of the fibrils that were inaccessible to the proteases, i.e., the fibril cores. Fibrillar pellets were treated with guanidine-HCl prior to perform the RP-HPLC analysis. An aliquot from each sample of fibrils core fractions was withdrawn before the guanidine-HCl treatment and tested by TEM imaging to check the integrity of fibrils cores. The same procedure was performed on the dimer fibrils. The RP-HPLC profiles of fragment separation after proteinase K and trypsin digestion are shown in Fig. 3.18 and 3.19, respectively.

Subsequently, peptides isolated by RP-HPLC were identified by ESI-QTOF mass spectrometry. The obtained molecular masses corresponding to peptides are summarized in the Supplementary Material (Table I - X). The identification of proteolytic fragments was performed by using MassLynxs software. Fig. 3.17 reports some of the identified fragments isolated by RP-HPLC.

The proteolytic pattern of aS digested by PK evidences a protected region that encompasses amino acids 31 to 113, with a cleavage site between residues 56-57. The same region is evidenced by trypsin digestion (region 35-96), as confirmed by the peptides identified in the fibrillar pellet. These results are in agreement with some published data, which underline the presence of central NAC region in the amyloid fibril core (Miake et al., 2002; Quin et al. 2007). The hydrolysis at the N- and C- terminals indicates that both terminal regions of aS are localized at the surface of fibril core, therefore not involved in β-sheet structure, so that they are accessible to proteases. In the soluble fraction of trypsin digested fibrils, fragmentation of protected region is observable (hydrolysis of peptide bones between residues 45-46, 58-59, 34-35, 43-44, 58-59, 80-81). Since no evidence of these cleavages are reported in the fibrillar pellet fraction, this result should indicate that monomeric, non structured molecules of aS were present prior to the proteolysis, therefore they were digested completely by the

protease. Of particular interest is the PK hydrolysis site between residues E56-K57, which is situated inside the protected region individuated by the two proteases. Interestingly, trypsin proteolysis does not evidence cleavages after residues K57 and K60. This data suggest that few amino acids inside the protected, high structured region are accessible to the protease, probably corresponding to a turn in the β-strands arrangement. The exact number of β-strands, and the identification of amino acids for each β-strands in the fibrils of aS are actually not known. On the base of AFM measurements they were proposed to be five to seven (Quin et al. 2007), and on the base of D/H exchange experiments followed by NMR they were proposed to be five (Vilar et al., 2008). In particular, Vilar and colleagues identified a region of fast D/H exchange within 56-60 amino acids, in agreement with the proteolysis cleavage reported here between residues E56-K57.

The analyses of the proteolytic pattern of the dimer fibrils did not evidence remarkable difference among the species and respect to the aS proteolytic pattern. For NN, CC and NC dimer the protected region overlaps the one of aS, i.e. the 61 amino acids among 35-96 residues (Supplementary Material, Table III - VIII). This result confirms that the central region is highly prone to form amyloid structure and constitutes the same fibril core for aS and the three dimers. The presence of extended, unstructured, terminal regions between the two NAC regions that constitutes each dimers does not affect the possibility of internal arrangements and the formation of amyloid fibrils.

On the other hand, DC is constituted by two consecutive NAC regions. In particular, it is formed by linking 1-104 and 29-140 residues of aS in a unique molecule. On the bases of obtained results on aS, the fibrils core should involve region 35-96 and 111-172. No fragments which encompass the whole region (35-172) were identified in the fibrillar pellet fractions. On the contrary, some cleavage sites were found inside the region 96-111, such as 80-81, indicating that this region is available to protease attack. Some hydrolysis sites were also detected at extremes of the protected region, such as 43-44, 45-46, 136-137, 156-157 (Supplementary Material, Table IX – X, peptides which indicates cleavage sites inside aS core region are indicated in italic font). This result indicates that the presence of a region that is not prone to form β-strand between the two NAC (residues 96-139) somehow reduces the strength of the β-sheet that form the fibrillar core. Overall, DC dimer fibrils seems to be less rigid and packed than aS, and the other dimers, fibrils.

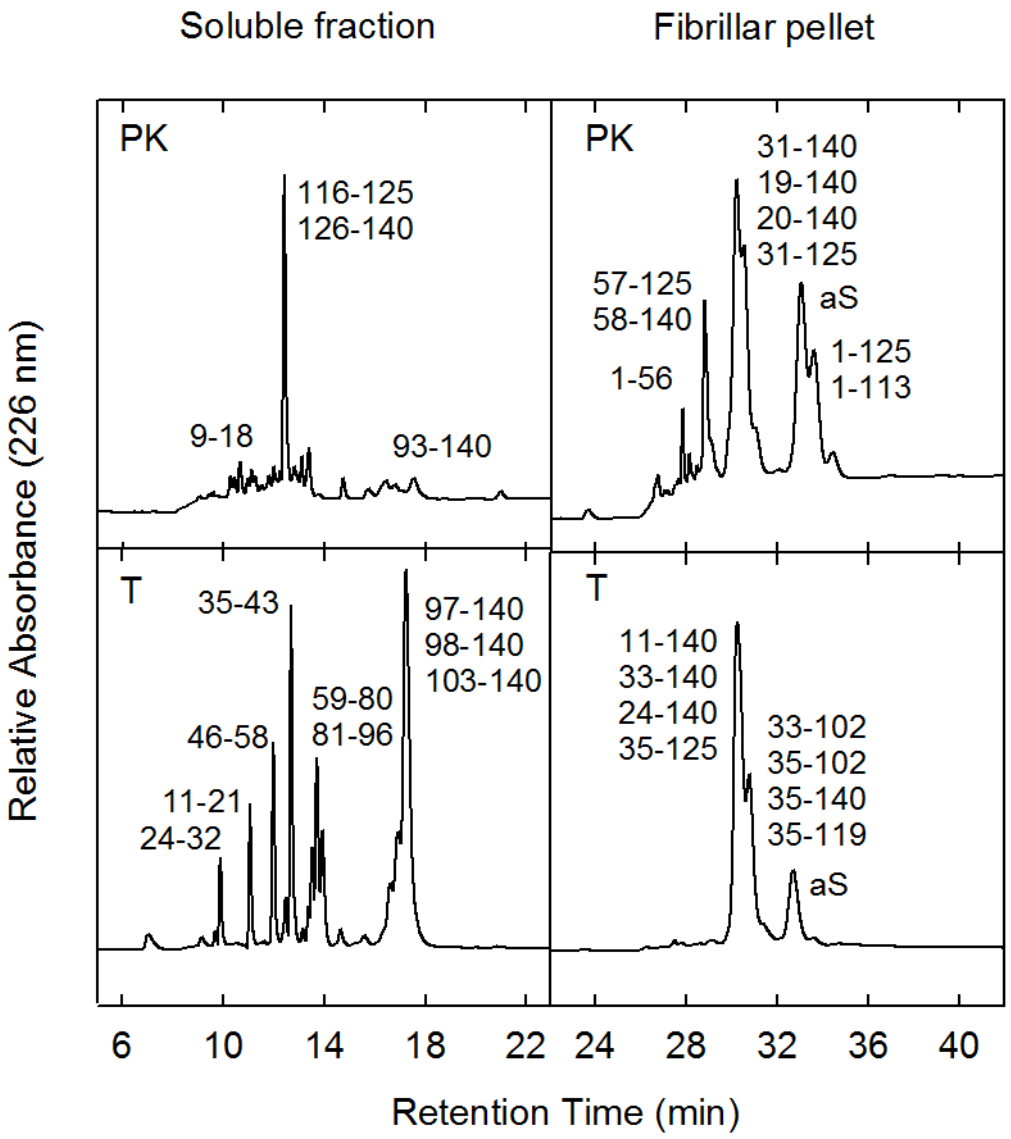


Fig. 3.17 Proteolysis of aS fibrils analyzed by RP-HPLC. Proteolysis of aS fibrils by proteinase K (PK) (upper panels) and trypsin (T) (lower panels). The reactions were conducted on ultra-centrifuged fibrils for 2 hours at a protein concentration of 70 μM in PBS buffer, using an E/S weight ratio of 1:1000 and 1:50, respectively. After the reaction, soluble peptides (left panels) were separated from the fibrillar core (right panels) by ultra-centrifugation and fibrillar pellets were treated overnight with guanidine-HCl prior to perform RP-HPLC.

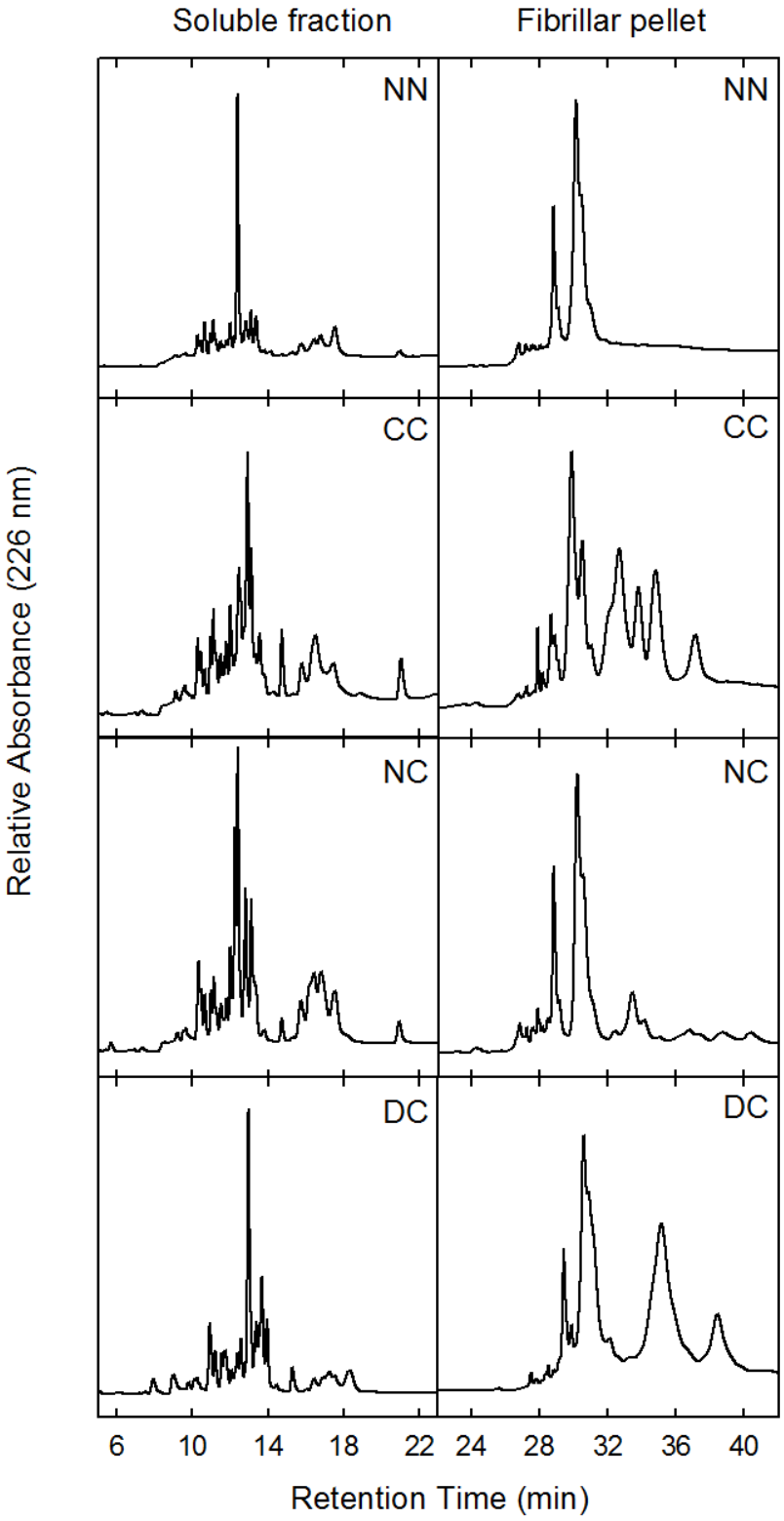


Fig. 3.18 Proteinase K proteolysis of NN, CC, NC and DC dimer fibrils analyzed by RP-HPLC. Proteolysis of dimers' fibrils by PK. The reactions were conducted on ultra-centrifuged fibrils for 2 hours at a protein concentration of 70 μM in PBS buffer, using an E/S ratio of 1:1000 (by weight). After the reaction, soluble peptides (left panels) were separated from the fibrillar core (right panels) by ultra-centrifugation and fibrillar pellets were treated overnight with guanidine-HCl prior to perform RP-HPLC.

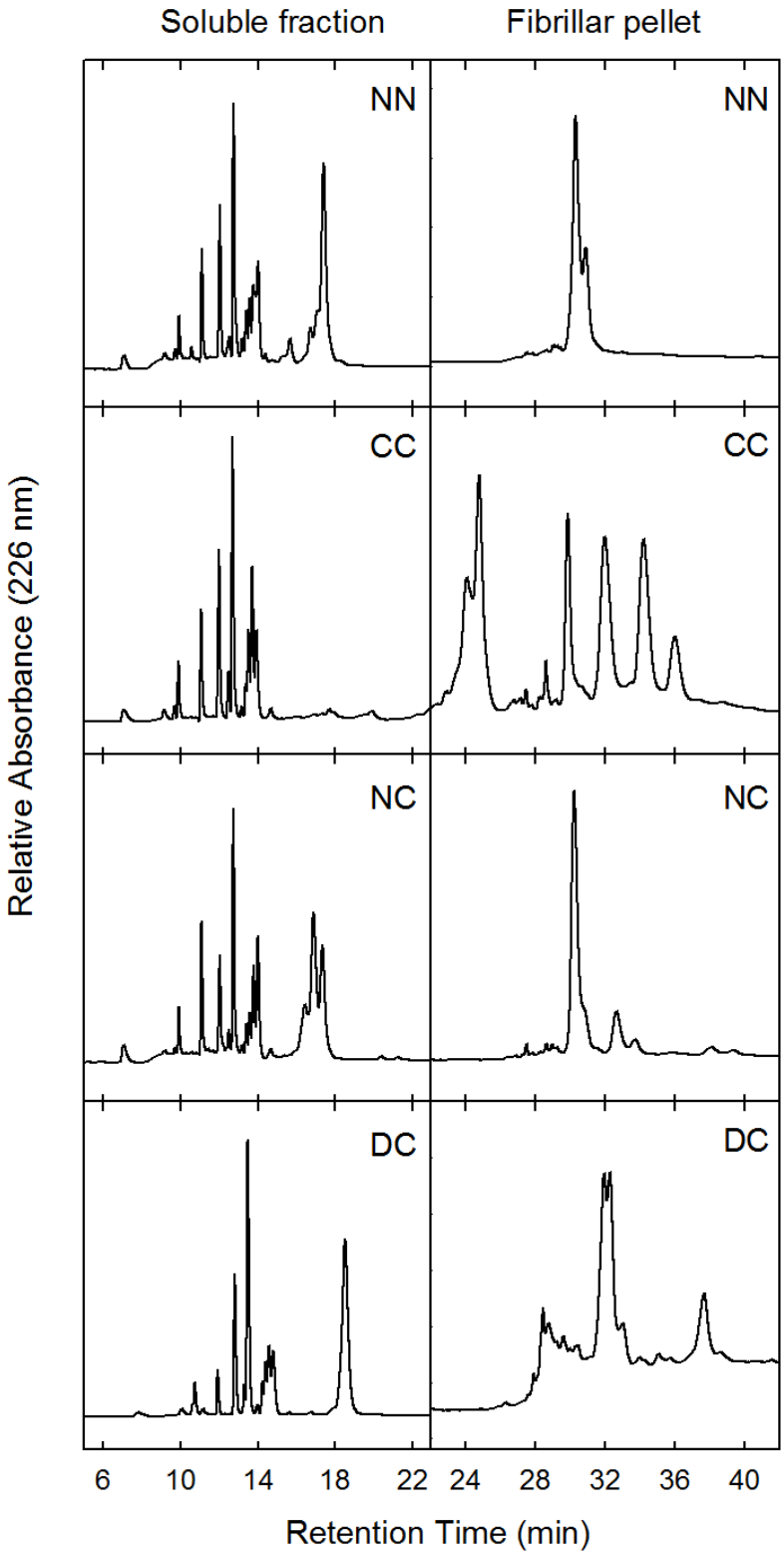


Fig. 3.19 Trypsin proteolysis of NN, CC, NC and DC dimer fibrils analyzed by RP-HPLC. Proteolysis of dimers' fibrils by trypsin. The reactions were conducted on ultra-centrifuged fibrils for 2 hours at a protein concentration of 70 μM in PBS buffer, using an E/S ratio of 1:50 (by weight). After the reaction, soluble peptides (left panels) were separated from the fibrillar core (right panels) by ultra-centrifugation and fibrillar pellets were treated overnight with guanidine-HCl prior to perform RP-HPLC.

3.5 Co-aggregation of aS and dimers

The aggregation process of aS was conducted in the presence of small percentage of dimers (at monomeric state). The aim was to verify if dimers were able to affect aS aggregation kinetic. Aggregation studies were carried on under the same conditions previously described (PBS buffer, pH 7.4, at 37°C, shaking at 500 rpm) incubating aS in the presence of 20% dimer molecules (weight percentage) at a total protein concentration of 1 mg/ml. Two independent aggregation experiments were conducted for each aS-dimer mixture. Aliquots of the aggregation samples were withdrawn at different time of incubation and analyzed by ThT binding assay.

All the aggregation mixtures have produced amyloid fibrils as confirmed by TEM imaging (Fig. 3.20) and the increase of ThT fluorescence at 485 nm (Fig. 3.21). All the protein mixtures, with the exception of aS-NN dimer, give rise to twisted fibrils composed by two or more filaments. Since the kinetic of aS-NN dimer is slower (Fig. 3.21), and aS is the main component of the mixture, it is plausible suppose that prolonged incubation times would result in twisted fibrils.

The aggregation processes of the protein mixtures were analyzed in comparison with that of aS alone. aS reaches the fluorescence plateau after 10 days, with 2 days of lag time. NN and NC dimers have slower kinetic than the others (in particular, NN dimer has the slowest one). aS-NN mix reaches the plateau in 13 days, while aS-NC mix shows the same kinetic of aS. The two experiments conducted for aS-DC mix show a different behavior, with a lag period of 6 and 8 days and the plateau after 8 and 12 days of incubation, respectively. A different kinetic respect to aS alone is obtained by the addition of CC to aS aggregation mixture. Indeed CC is able to strongly accelerate the process, as it is appreciable in Fig. 3.22. aS and aS with 20% CC dimer show the same lag period, but aS-CC mix reaches the fluorescence plateau in half time in respect to aS. As discussed in section 3.4, while the lag period correspond to the time required for the protein nucleation, the slope of the fitted sigmoid curve, which describe the aggregation kinetic, depends on the maturation (elongation) of the fibrils, i.e. on the affinity of monomers to the growing fibrils (Fink, 2006). In the cases reported in Fig. 3.22, the slopes are 18.1 ± 7.2 and 75.5 ± 10.5 for aS and aS in the presence of CC, respectively (calculated as the curve tangent at the flex position).

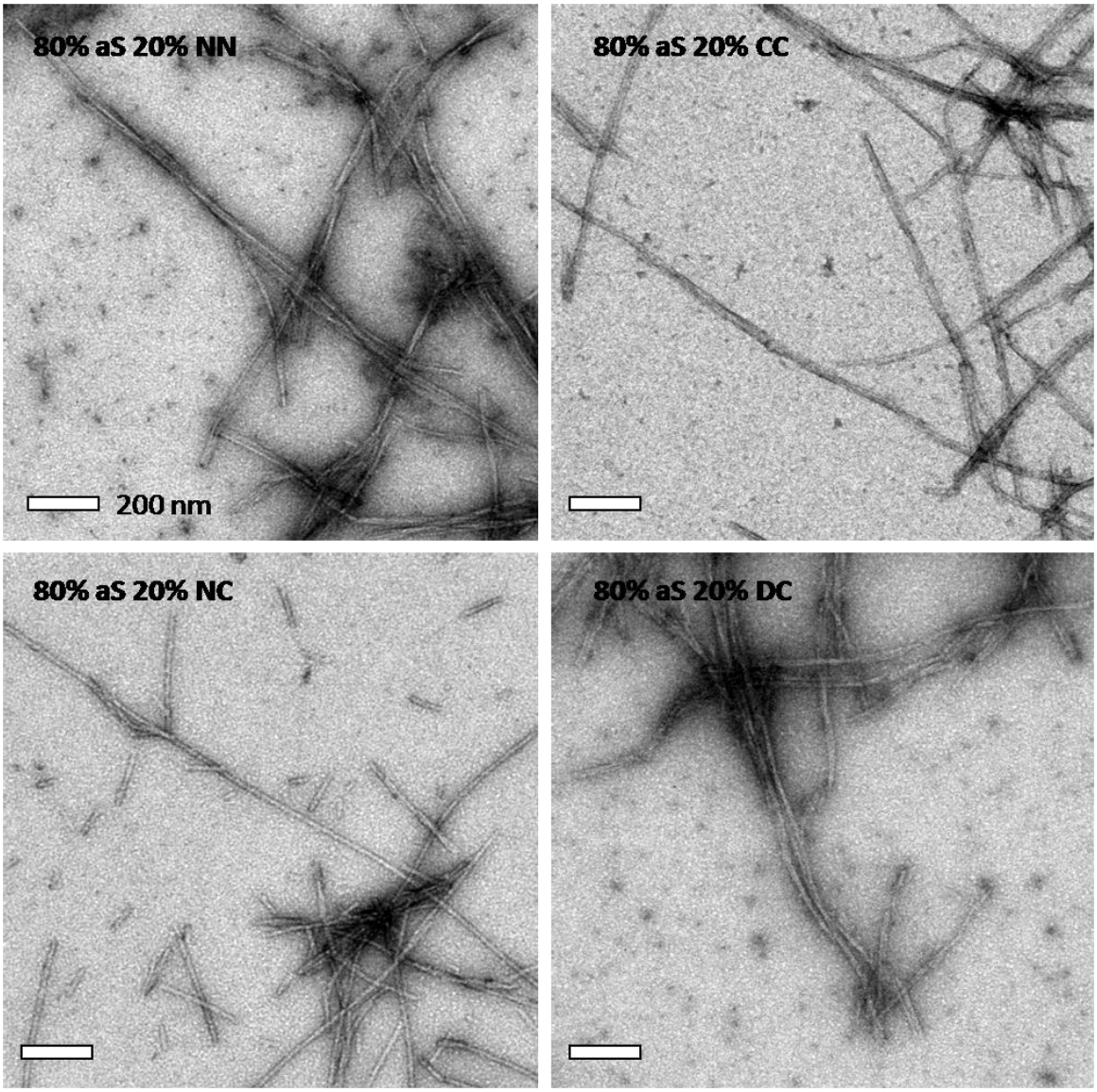


Fig. 3.20 TEM imaging of aS-dimers aggregation mixture. TEM pictures relative to aS incubated with 20% of NN, CC, NC and DC dimer after 13 days of incubation are reported. Barrels indicate 200 nm.

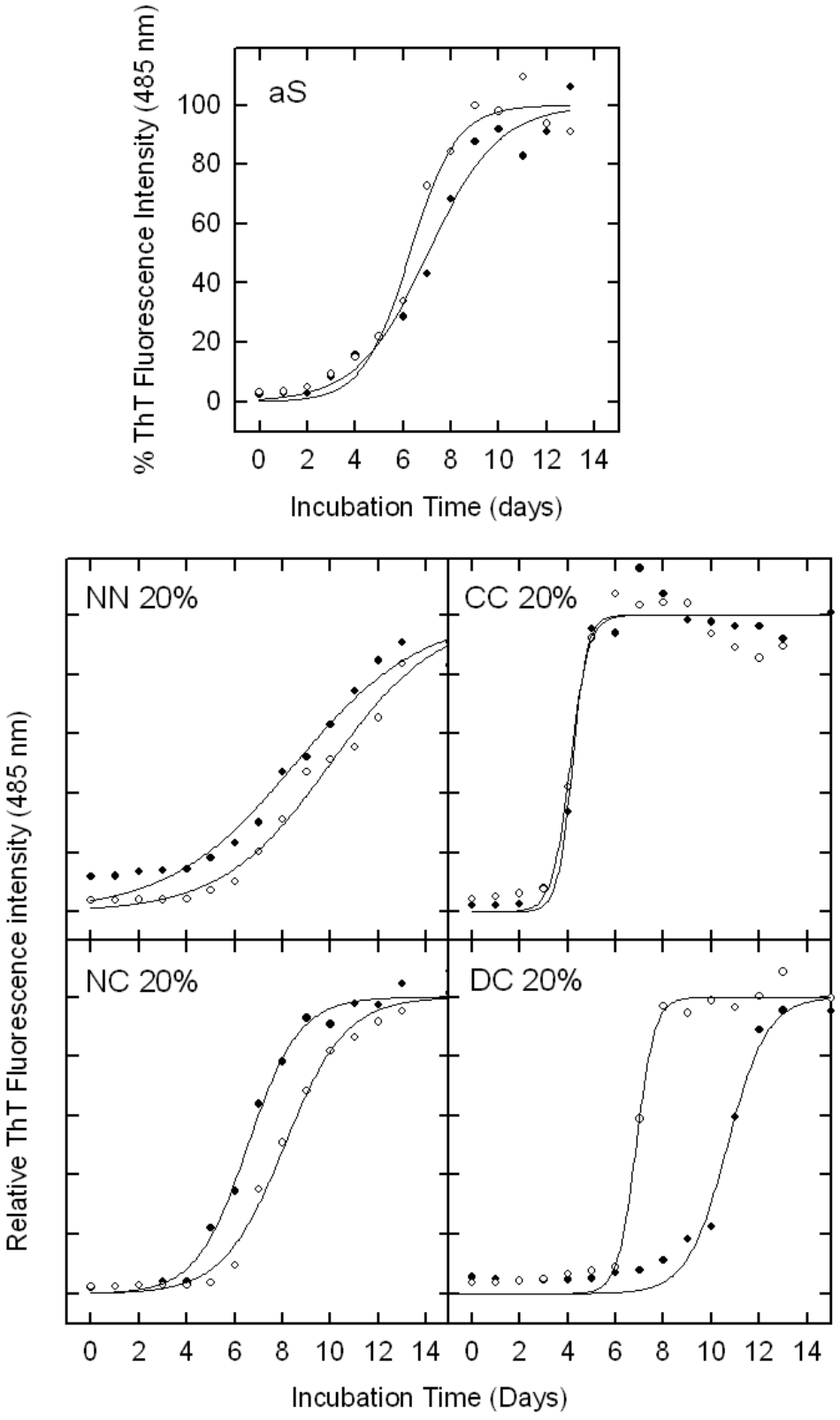


Fig. 3.21 Time-course analysis of the aggregation process of aS and of aS in the presence of 20% dimer followed by ThT fluorescence assay. aS and aS in the presence of 20% dimer were incubated for 14 days at a final concentration of 1 mg/ml (PBS buffer, 37°C, shaking at 500 rpm). Three independent experiments for each aggregation were followed. Aliquots were withdrawn from the solution for ThT fluorescence assay.

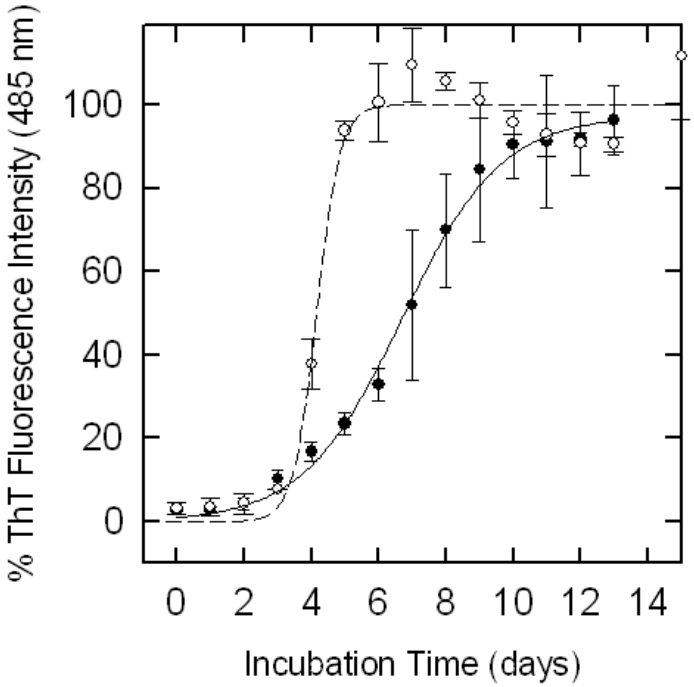


Fig. 3.22 Comparison of the aggregation process of aS and aS in the presence of 20 % CC. ThT fluorescence intensity for aS and aS in the presence of 20% CC are shown. The curves of three independent experiments were normalized for their plateau value (calculated by curve fitting operation)

In Fig. 3.23 the evolution of the secondary structure of aS during aggregation in the presence of 20% dimer was followed by far UV CD. At the beginning of the incubation process, aS and aS-CC mix show CD spectra indicative of the presence of unfolded protein structure. During the aggregation processes, both protein species show a decrease of the CD minimum at 197 nm. After 7 days of incubation, the spectrum of aS does not present the minimum at 217 nm, indicative of the presence of a β-sheet structure. This could be due to the presence of monomeric aS or non-β-structured oligomeric aS in the mixture also after 7 days incubation. Indeed, removing these non-fibrillar aggregates by ultra-centrifugation, the spectra of the pellet that is supposed that contains only fibrillar material displays the characteristic CD band at 217 nm (Fig. 3.23, Insert). Noteworthy, after 3 days of incubation this band is appreciable in the case of aS incubated with CC, also without the ultra-centrifugation.

Native gel electrophoresis was also used to analyze the final products of the aggregation of aS conducted in the presence of dimers. The aggregation was conducted for 15 days (Fig. 3.24). The soluble fraction was separated by the fibrillar material by ultra-centrifugation and quantified by UV absorbance. Then an aliquot was loaded into a gel, and the electrophoretic pattern was compared with an aliquot of the starting aggregation mixture. Interestingly, supernatant of the mixture of aS with CC or DC dimer do not show the band corresponding to the dimer species. This is an indication that all the dimer present at the beginning of the aggregation has been incorporated into the

growing fibrils. For NN and NC dimer-aS mixtures the bands correspondent to the dimers are still present, and in particular the band of NN dimer is less intense than the NC one.

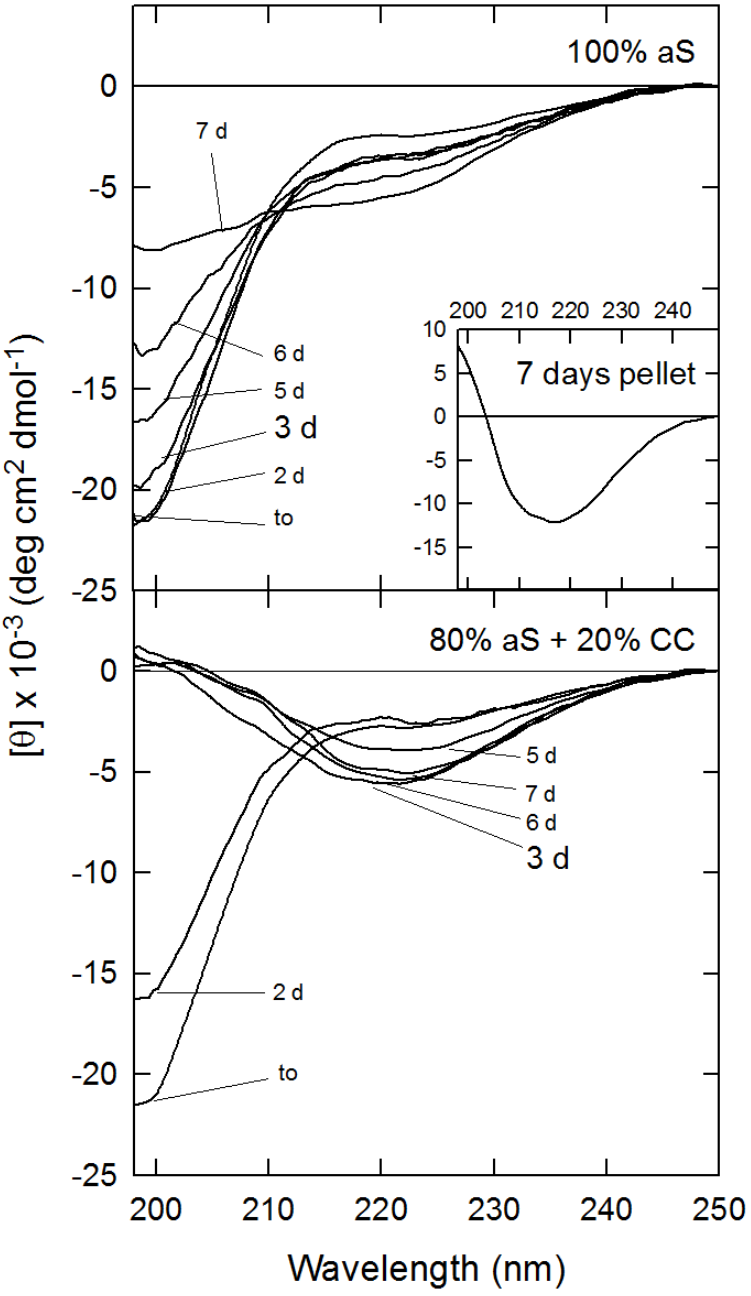


Fig. 3.23 Far UV CD of aS (upper panel) and aS-CC mixture (lower panel) during aggregation. Evolution of the secondary structure of aS in comparison with aS-CC mixture, monitored by far-UV CD during aggregation. The spectra were recorded in PBS buffer pH 7.4 at a protein concentration of 5 mg/ml, using a quartz cuvette with 1 mm of pathlength. **Insert**: CD spectra of fibrillar ultra-centrifuged pellet of aS aggregation at 7 days.

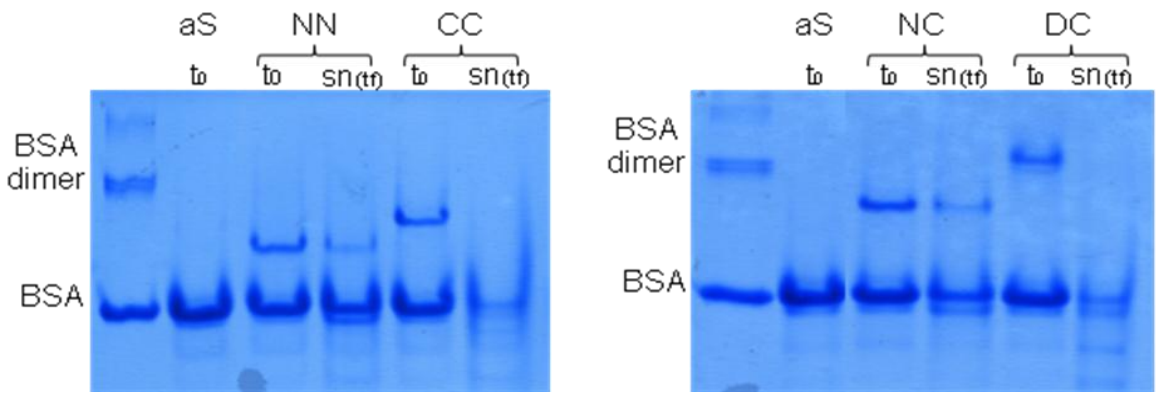


Fig. 3.24 Native gel electrophoresis. Native gel electrophoresis (12% acrylamide) of aS and aS in the presence of the different dimers (20%). For each experiment, an aliquot corresponding to the mixture at the initial time (t_0) and after 13 days of incubation and ultra-centrifugation (sn t_f) were loaded. BSA was used as reference marker.

There are two possible explanations for the presence of the band of NN and NC after 15 days of incubation. Fibril formation was not completed even after this incubation time. Another possibility is that the two dimers are less prone to be incorporated into the growing fibrils in respect with CC and DC dimer.

The analysis of the aggregation process of co-aggregation experiments evidenced that the presence of the dimer species do not affect the formation of the nucleation state, i.e. the dimers do not constitute an anchor for the formation of fibrils. Nevertheless, they, and in particular CC, affect fibril growth. In the case of CC, it seems that this dimer has a strong affinity to pre-fibrillar species and a high tendency to bind to the elongating filaments. This tendency is underlined by the fact that at the end of the process, there is no evidence of the presence of the dimer in the supernatant of the aggregation mixture. Among the different dimers, the affinity to the growing fibrils seems to be proportional to the kinetic properties of each dimer (for example, NN dimer show the slowest kinetic and the incubation of NN in aS mixtures decelerate aS aggregation). DC dimer presents a controversial behavior and requires further studies.

4. Discussion

The formation of aS amyloid fibrils represents a key event in the pathogenesis of PD (Cookson 2005 and 2009; Fink, 2006). The histological hallmark of PD is the presence of intracellular inclusions called Lewy bodies (LBs). LBs are composed of filamentous and aggregated proteins, mainly by aS. Despite its importance in neurodegeneration, little is known about aS function, native physiological state and mechanism of aggregation. aS was recently described as a folded tetramer (Bartel et al., 2011; Wang et al., 2011), but it is generally considered a natively unfolded protein (Uversky, 2001, 2003 and 2007). During PD pathological process, aS acquires β–sheet secondary structure and forms amyloid fibrils. aS fibrils are composed by *parallel* and *in register* β-strands (Fig. 1.6 B), i.e. single units of NAC forms five β-strands, and the β-structured NAC regions interact in parallel with each other (Chen et al., 2003; Vilar et al., 2008; Karyagina et al., 2011). The presence of aS soluble oligomers has also been described. These species are transient intermediates between the physiological form of aS and amyloid fibrils. Recently, oligomers were proposed to be the toxic species that lead to cell degeneration (El-Agnaf et al., 2001; Volles et al, 2001; Volles and Lansbury, 2003).

In order to study the mechanism of aggregation of aS, dimers of the protein have been obtained (Fig. 4.1). A cysteine residue has been added at the N- or C-terminal of aS, allowing the formation of covalent NN and CC dimers through a disulfide bond. A NC dimer was obtained as a single polypeptide chain. During the project, another dimer, called DC, was produced. It is constituted by the segment 1-104 linked to 29-140 one, in order to draw up the regions containing the high amyloidogenic NAC segment (Miake et al., 2002).

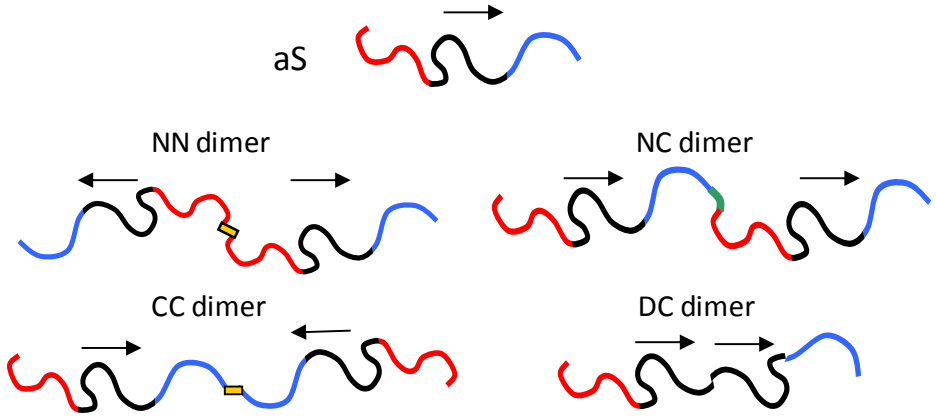


Fig. 4.1 Schematic representation of aS and aS dimers. N- and C-terminal regions are drawn in red and blue, respectively. Yellow rectangle represents disulphide bond, which links aS monomers in NN and CC dimers. NC and DC dimers were produced as single polypeptide chains. Arrows indicate the orientation of the central NAC region.

My research project has focused on the characterization of the aggregation properties of aS dimers in comparison with those of aS. aS dimers would constitute a suitable tool for the study of intra-molecular aS interaction pathway. The sequence of the protein is conserved, therefore all the amyloidogenic determinants are maintained. There are some differences that arise from the fact that two molecules have been joined, limiting the conformational mobility of the dimer molecules in respect to aS, hypothetically differentiating the fibrillation process of the five proteins.

All the dimers are able to form amyloid fibrils, as confirmed by a number of biophysical and biochemical techniques. The core structure of the fibrils of the different molecules seems to be constituted by the same amino acidic region, which encompasses the segment 35-96, in analogy with previous studies (Fink, 2006; Qin et al. 2007; Vilar et al., 2008). The main difference between aS and dimer fibrils was evidenced, by morphological analyses, in the organization of filaments within the fibril. The dimer fibrils are formed by single filaments, conversely from aS, whose mature fibrils are composed by two twisted filaments. Moreover, while NN and NC dimers show a slower kinetic of fibrillation than aS, the rate of fibril formation of CC and DC dimers is faster than aS. Another remarkable result concerns the ability of CC dimer to catalyze the fibrils elongation in the aggregation of aS. Indeed, aS, in the presence of a small amount of CC dimer, aggregates faster than aS alone, CC dimer molecules are all enclosed in the final fibrils. The lag time doesn't change in the two aggregation processes, but fibrils formed by the mixture of aS and CC grow faster, as evidenced by the curve trend relative to the process (Fig. 3.22).

We formulated three hypotheses about the molecular conformation of dimers within the fibril. As shown in the images of Fig. 4.2, both NAC regions of a dimer could form β-strands without interacting between each other (Fig. 4.2, A). Otherwise only one region folds into the cross β-sheet structure (Fig. 4.2, B) and the other one remains unfolded. Alternatively an intra-molecular rearrangement could take place between the two NAC regions of the same molecule allowing the formation of *parallel* and *in register* β-sheets, as it is described for aS fibrils (Fig. 4.2, C).

The first hypothesis was ruled out by AFM measurements. AFM techniques determines the height profile of structured regions, therefore extended core region (as hypothesized in Fig 4.2, A) would result in higher diameter. Nevertheless, our analysis has shown that mature fibrils of the three dimers have similar diameters both between each others (approximately 5.37 ±1.09 nm) and in respect to the single filament of aS (5.44 ±0.59 nm).

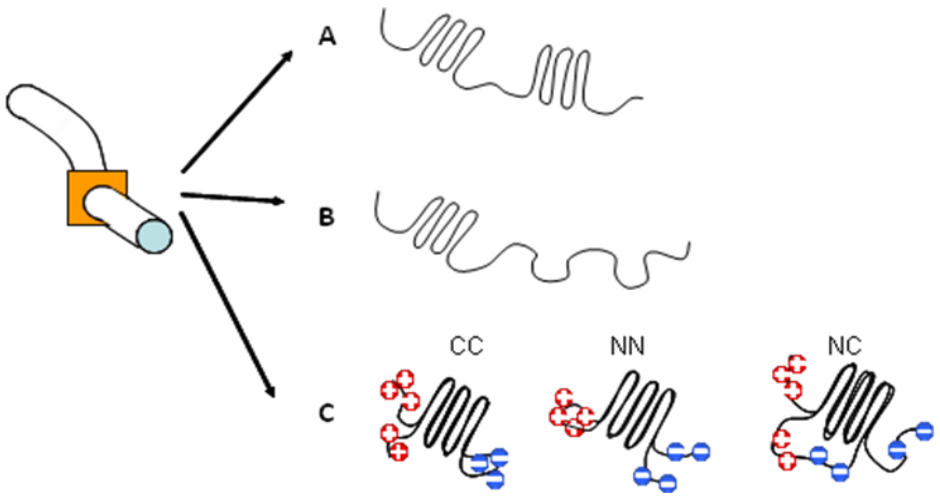


Fig. 4.2 Schematic representation of hypothesized dimer folding into the fibrils. A. The two NAC regions form consecutive β-strands. **B**. Only one NAC region is folded into β-sheet. **C**. Intra-molecular rearrangements allow the formation of parallel, in register β-strands.

Proteolysis experiments confirmed that both NAC regions of each dimer constitute the β-sheet structure core, excluding the possibility that only one chain of the dimer would be involved in the fibril structure. Indeed, dimer fibrils show the same proteolytic pattern of aS one. The protected region encompasses the segment 35-96. A proteinase K cut occurs inside the protected region, between residues 56-57, consistent with the hypothesis of a turn between the first two β-strand of the core (Vilar et al., 2008). In agreement with proteolysis data, FT-IR measurements showed that β-sheet structure content (Amide I region) is analogous among dimers and aS. Therefore, the most probable conformation assumed by the dimers into the fibril is that described in Fig. 4.1 C. To assume *parallel* and *in register* β-sheet structure, as described for aS fibrils (Fig. 1.6), intra-molecular rearrangements take place to allow the NAC segments to draw up. Then folded dimers interact with other molecules to build up the fibril.

To obtain the correct fold, the orientation of NAC region (represented with arrows in Fig. 4.1) could be a key aspect. NN and CC dimers are expected to be in a suitable orientation to form *parallel* and *in register* β-strands. Indeed we can assume that the molecule can bend by a simply pairing of the hydrophobic regions, as if it pivotes on the covalent link between the two units as a “nutcracker” (Fig. 4.3). On the other hand, NC dimer requires a more complex reorganization: to pair the NAC regions; one molecule of the dimer has to turn and roll over, justifying its slower kinetic in comparison with other molecules.

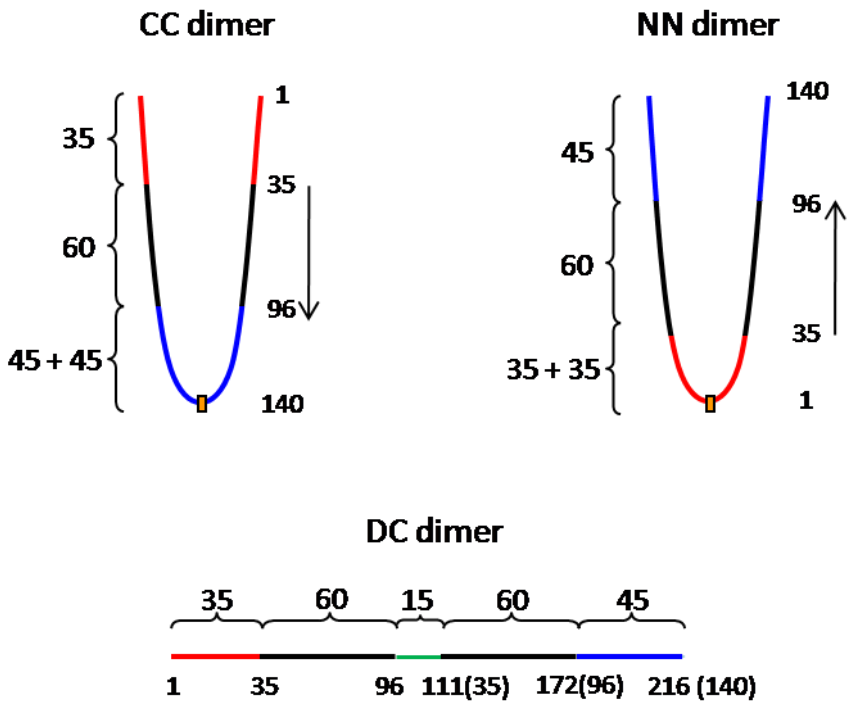


Fig. 4.2 Schematic representation of NN, CC and DC dimers folding. N- and C-terminal regions are drawn in red and blue, respectively. The yellow rectangle represents the disulphide bond. Numbers close to the brackets indicate the residues which constitute N- terminal, central and C-terminal regions (35, 60 and 45 residues, respectively).

However, on the base of these considerations, it is difficult to explain the different kinetic behaviour among NN and CC. We could take in account that drawing up the charged regions, a repulsive charge effect could occur and compromise interactions between NAC regions, and therefore fibril formation. In CC, there is an extended, negatively charged region (two consecutive C-terminal regions, 90 residues), but it does not seem to affect the kinetic of aggregation. In NN the region that turns involves 70 residues (Fig. 4.3). The reduced dimension of the segment among the NACs could increase repulsive charge effects. In the case of DC dimer, that shows the fastest kinetic, the hydrophobic region is such extended to make the molecule prone to aggregate. The same kinetic trends were observed in co-aggregation experiments, confirming the relative tendencies to aggregate among the different molecules. In particular, CC seems to have a considerable affinity to the growing fibrils, as determined by the major slope of the curves that describe the kinetic (Fig. 3.22).

Terminal regions were demonstrated to be highly dynamic not only in the native state, which are naturally disordered (Uversky, 2001), but also in the fibrillar state, as seen by NMR and computational studies (Der-Sarkissian et al., 2003). Although the two side regions of aS are not involved in the assembly of the core fibrils, they affect the organization of mature fibrils. aS mature fibrils are described as multiple, twisted filaments (Qin et al., 2007). In the dimers fibrils the covalent linkage between aS units

strongly conditions the ability of the side regions to interact with other filaments and form H bonds with them. TEM imaging proved the inability of our dimers to form fibrils with double twisted filaments. This observation confirms that the terminal side chains of aS, not involved in the core β-structure of the fibril, are indeed involved in the inter-molecular interactions among filaments. In the case of aS, terminal, unfolded chains allow the formation of double twisted filaments, whereas, in the dimers, the limited degrees of freedom affect the possibility to create these interactions.

Besides the faster aggregation kinetic of DC in comparison with other dimers and aS, further differences were evidenced. Unfortunately no definitive AFM measurements are available to evaluate the dimension of the DC fibrils, they are under investigation. An increased diameter would be consistent with the presence of consecutive β-strands belonging to the two contiguous NAC, as shown in Fig. 4.2, A for other dimers. Similar dimension would confirm an intra-molecular rearrangement. TEM imaging shows that DC fibrils have a larger diameter (11.85 ±1.35 nm) than other dimers, suggesting an extended core model. Proteolysis experiments have shown that the connecting loop between the two NAC segments (15 residues) is susceptible to proteases attack (Fig. 4.3). As a consequence it is not involved in persistent secondary structure, as in the cross-β-structure of fibrils core. Summing up, the two NAC could be contiguous within the fibril core, the connecting loop is too short to allow the bending of the two units, but it is unfolded and permits the protease cleavage.

The formation of aS dimers was proposed to represent a critical, rate-limiting step for the aggregation of the protein (Takahashi et al., 2002; Krishnan et al., 2003). In this PhD Thesis we investigated the behaviour of aS dimers that present different orientation and distance of the central amyloidogenic regions. Amyloidogenic determinants were maintained, and the resulting fibrils show similar fibrillar core. aS dimers do not strongly affect the formation of the nucleation event that starts fibril maturation (lag period), thus indicating that dimerization is not the only factor which affect the aggregation pathway of aS. Nevertheless, among the dimers that contain the whole sequence of aS, CC showed the major affinity to the growing fibrils, as confirmed by the co-aggregation experiments. This suggests that aS and CC could assume during aggregation a similar fold and thus CC fibrils are easily integrated in the aS fibril structure.

References

Abeliovich, A., Schmitz, Y., Fariñas, I., Choi-Lundberg, D., Ho, W. H., Castillo, P.E., Shinsky, N., Verdugo, J. M., Armanini, M., Ryan, A., Hynes, M., Phillips, H., Sulzer, D. & Rosenthal, A. (2000). Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron*, **25**, 239–252.

Abeliovic, A. (2010). Mitochondrial damage control in Parkinson's disease. *Nature*, **463**, 744-745.

Adibhatla, R. M., Hatcher, J. F. (2008). Phospholipase A2, reactive oxygen species, and lipid peroxidation in CNS pathologies. *BMB Rep.*, **41**:8, 560-567.

Andresen, J. K. (2009). Arvid Carlsson: an early pioneer in trasnational medicine. *Sci. Transl. Med.*, **1**, S1-S1.

Apetri, M. M., Maiti, N. C., Zagorski, M. G., Carey, P. R., Anderson, V. E. (2006). Secondary structure of α-synuclein oligomers: charcterizaton of Raman and atomic force microscopi. *J. Mol. Biol.*, **355**, 63-71.

Barth, A. (2007). Infrared spectroscopy of proteins. *Biochim. Biophys. Acta*, **1767**, 1073-1101.

Barth, A. (2007). The infrared absorbtion of amino acid side chain. *Prog. Biophys. Mol. Biol.*, **74**, 141-173.

Bekris, L. M., Mata, I. F., Zabetian, C. P. (2010). The genetics of Parkinson disease. *J. Geriatr. Psychiatry Neurol*., **23**:4, 228-242.

Bertoncini, C. W., Jung, Y. S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., Zweckstetter, M. (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. U S A.*, **102**, 1430-1435.

Beyer, K., Lao, J. I., Carrato, C., Mate, J. L., Lopez, D., Ferrer, I., Ariza, A. (2004). Upregulation of amyloid precursor protein isoforms containing Kunitz protease inhibitor in dementia with Lewy bodies. *Brain Res. Mol.*, **131**, 131–135.

Beyer, K., Humbert, J., Ferrer, A., Lao, J. I., Carrato, C., López, D., Ferrer, I., Ariza, A. (2006). Low a-synuclein 126 mRNA levels in dementia with Lewy bodies and Alzheimer disease. *NeuroReport*., **17**:12, 1327-1330.

Beyer, K. (2006). α-Synuclein structure, posttranslational modification and alternative splicing as aggregation enhancers. *Acta Neuropathol.*, **112**, 237-251.

Braak, H., Del Tredici, K., Rueb, U., de Vos, R. A., Jansen Steur, E. N., Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, **24**, 197-211.

Bonifati, V., Rizzu, P., van Baren, M. J., Schaap, O., Breedveld, G., J., Krieger, E., Dekker, M. C., Squitieri, F., Ibanez, P., Joosse, M., van Dongen, J. W., Vanacore, N., van Swieten, J. C., Brice, A, Meco, G., van Duijn, C. M., Oostra, B. A., Heutink, P. (2003). Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science*, **299**, 256-259.

Camandola, S., Poli, G., Mattson, M. P. (2000). The lipid peroxidation product 4-hydroxy-2,3-nonenal increases AP-1-binding activity through caspase activation in neurons. *J. Neurochem.*, **74**, 159-168.

Castagnet, P. I., Golovko, M. Y., Barceló-Coblijn, G. C., Nussbaum, R. L. & Murphy, E. J. (2005). Fatty acid incorporation is decreased in astrocytes cultured from alpha-synuclein gene-ablated mice. *J. Neurochem.* **94**, 839–849.

Chandra, S., Chen, X., Rizo, J., Jahn, R., Sudhof, T. C. (2003). A broken α-helix in folded α-synuclein. *J. Biol. Chem.*, **278**, 15313-15318.

Chartier-Harlin, M. C., Kachergus, B.S., Roumier, C., Mouroux, V., Lincoln, S., Levecque, C., Larvor, L., Andrieux, J., Hulihan, M., Waucquier, N., Defebvre, L., Amouyel, P., Farrer, M., Destee, A. (2004). α-Synuclein locus duplication as a cause of familiar Parkinson's disease. *The Lancet*, **364**:9440, 1167-1169.

Chen, L., Feany, M. B. (2005). Alpha-synuclein phosphorylation controls neurotoxicity and inclusion formation in a Drosophila model of Parkinson disease. *Nature Neuroscience*, 8, 657–663.

Chen, M., Margittai, M., Chen, J., Langen, R. (2007). Investigation of a-synuclein fibril structure by site-directed spin labeling. *J. Biol. Chem.*, **282**:34, 24970-24979.

Clark, I. E., Dodson, M. W., Jiang, C., H. C., Joseph, Huh, J. R., Hong Seol, J., Ji Yoo, S., Hay, B., A. & Guo, M. (2006). *Drosophila pink1* is required for mitochondrial function and interacts genetically with *parkin*. *Nature*, **441**, 1162-1166.

Clayton, D. F., George, J. M. (1998). *Trends. Neurosci.*, **21**:6, 249-254.

Clements, C.M., McNally, R.S., Conti, B.J., Mak, T.W., Ting, J.P. (2006) DJ-1, a cancer- and Parkinson's disease-associated protein, stabilizes the antioxidant transcriptional master regulator Nrf2. *Proc. Natl Acad. Sci.*, **103**, 15091-15096.

Conway, K. A., Harper, J. D., Lansbury, P. T. (1998). Accelerated *in vitro* fibril formation by a mutant a-synuclein linked to early-onset Parkinsonian disease. *Nature Madicine*, **4**, 1318-1320.

Conway, K. A., Rochet, J. C., Biegansky, R. M., Lansbury, P. T. (2001). Kinetic stabilization of the α-Synuclein protofibril by a Dopamine-α-Synuclein adduct. *Science*, **294**, 1346-1349.

Cookson, M. R. (2005). The biochemistry of Parkinson's disease. *Annu. Rev. Biochem.*, **74**, 29-52.

Cookson, M. R. (2009). α-Synuclein and neuronal cell death. *Molecular Neurodegeneration*, **4**:9, 1-14.

Cookson, M. R., Bandmann, O. (2010). Parkinson's disease: insights from pathways. *Human Molecular Genetics*, **19**:1, R21-R27.

Craft, S. and Watson, G. D. (2004). Insulin and neurodegenerative disease: shared and specific mechanisms. *The Lancet Neurology*, **3**, 169-178.

Croke, R. L., Sallum, C. O., Watson, E., Watt, E. D. & Alexandrescu, A. T. (2008). Hydrogen exchange of monomeric alpha-synuclein shows unfolded structure persists at physiological temperature and is independent of molecular crowding in Escherichia coli. *Protein Sci*., **17**, 1434–45.

Dauer, W., Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron*., **39**, 889-909.

Davidson, W. S., Jonas, A., Clayton, D. F. & George, J. M. (1998). Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. *J. Biol. Chem.* **273**, 9443–9449.

Dawson, T. M., Dawson, V. L. (2010). The role of parkin in familiar and sporadic Parkinson's disease. *Movement Disorders*. **25**:1, S32-S39.

Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., Dobson, C. M. (2005). Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc*., **127**, 476-477.

Der-Sarkissian, A., Jao, C. C., Chen, J., Langen, R. (2003). Structural organization of α-synuclein fibrils studied by site-directed spin labelling. *J. Biol. Chem.*, **39**:26, 37530-37535.

Ding, T. T., Lee, S. J., Rochet, J. C., Lansbury, P. T. Jr. (2002). Annular alpha-synuclein protofibrils are produced when spherical protofibrils are incubated in solution or bound to brain-derived membranes. *Biochemistry*, **41**, 10209-10217.

Dorval, V., Fraser, P. E. (2006). Small ubiquitin-like modiWer (SUMO) modiWcation of natively unfolded proteins tau and alpha-synuclein. *J. Biol. Chem.*, **281**, 9919–9924

Dusa, A., Kaylor, J., Edridge, S., Bodner, N., Hong, D. P., Fink, A. L. (2006) Characterization of oligomers during alpha-synuclein aggregation using intrinsic tryptophan fluorescence. *Biochemistry*, **45**, 2752□2760.

Ebeling, W., Hennrich, N., Klockow, M., Meta, H., Orth, D., and Lang, H. (1974). Proteinase K from *Tritirachium album Limber*. *Eur. J. Biochem.* **47**, 91–97.

El-Agnaf, O. M., Jakes, R., Curran, M. D., Middleton, D., Ingenito, R., Bianchi, E., Pessi, A., Neill, D., Wallace, A. (1998). Aggregates from mutant and wild-type a-synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of beta-sheet and amyloid-like filaments. *FEBS letters*, **1**:2, 71-75.

El-Agnaf, O. M., Nagala, S., Patel, B. P., Austen, B.M. (2001). Non-fibrillar oligomeric species of the amyloid ABri peptide, implicated in familial British dementia, are more potent at inducing apoptotic cell death than protofibrils or mature fibrils. *J. Mol. Biol*., **310**, 157-68.

Ellis, C. E., Murphy, E. J., Mitchell, D. C., Golovko, M. Y., Scaglia, F., Barceló-Coblijn, G. C. & Nussbaum, R. L. (2005). Mitochondrial lipid abnormality and electron transport chain impairment in mice lacking alpha-synuclein. *Mol. Cell. Biol*. **25**, 10190–10201.

Esteves, A. R., Arduino, D. M., Silva, D. F. F., Oliveira, C. R., Cardoso, S. M. (2011). Mitochondrial dysfunction: the road to alpha-synuclein oligomerization in PD. *Parkinson's Disease*, ID 693761.

Ferreon, A. C., Deniz, A. A. (2007). a-Synuclein multistate folding thermodynamics: implications for the protein misfolding and aggregation. *Biochemistry*, **46**:15, 4409-4509.

Fink, A. L. (2006). The aggregation and fibrillation of α-synuclein. *Acc. Chem. Res.*, **39**, 628-634.

Fleming, S. M., Fernagut, P. O., Chesselet, M. F. (2005). Genetic mouse models of parkinsonism: strengths and limitations. *Neuro. Rx*., **2**, 495-503.

Farrer, M., Kachergus, J., Forno, L., Lincoln, S., Wang, D. S., Hulihan, M., Maraganore, D., Gwinn-Hardy, K., Wszolek, Z., Dickson, D., Langston, J. W. (2004). Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. *Ann. Neurol*., **55**, 174-179.

Fontana, A., Polverino de Laureto, P., Spolaore ,B., Frare, E., Ricotti, P. & Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim Pol.* **51**, 299–321.

Fujiwara, H., Hasegawa, M., Dohmae, N., Kawashima, A., Masliah, E., Goldberg, M. S., Shen, J., Takio, K., Iwatsubo, T. (2002). alpha-Synuclein is phosphorylated in synucleinopathy lesions. *Nat. Cell. Biol.*, **4**, 160–164.

Gai, W. P., Yuan, H. X., Li, X. Q., Power, J. T. H., Blumberhs, P. C., Jensen, P. H. (2000). *Experimental Neurology*, **166**, 324-333.

Gandhi, S., Muqit, M. M., Stanyer, L., Healy, D. G., Abou-Sleiman, P. M., Hargreaves, I., Heales, S., Ganguly, M., Parsons, L., Lees, A. J., Latchman, D. S., Holton, J. L., Wood, N. W., Revesz, T. (2006). PINK1 protein in normal human brain and Parkinson's disease. *Brain*, 1**29**, 1720-1731.

Gautier, C. A., Kitada, T., Shen, J. (2008). Loss of PINK1 causes mitochondrial functional defects and incresed sensitivity to oxidative stress. *PNAS*, **105**:32, 11364-11369.

George, J. M., Jin, H., Woods, W. S., Clayton, D. F. (1995). Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron*, **15**, 361–372.

Giovannone, B., Tsiaras, W. G., de la Monte, S., Klysik, J., Lautier, C., Karashchuk, G., Goldwurm, S., Smith, R. J. (2009). GIGYF2 gene disruption in mice results in neurodegeneration and altered insulin-like growth factor signalling. *Hum. Mol. Genet.*, **18**: 23, 4629-4939.

Green, J. C., Whitworth, A. J., Kuo, I., Andrews, L. A., Feany, M. B, Pallanck, L. (2002). Mitochondrial pathology and apoptotic muscle degeneration in *Drosophila parkin* mutants. *PNAS*, **100**:7, 4078-4083.

Greggio, E., Jain, S., Kingsbury, A., Bandopadhyay, R., Lewis, P., Kaganovich, A., van der Brug, M. P., Beilina, A., Blackinton, J., Thomas, K. J., Ahmad, R., Miller, D. W., Kesavapany, S., Singleton, A., Lees, A., Harvey, R. J., Harvey, K., Cookson, M. R. (2006). Kinase activity is required for the toxic effects of mutant LRRK2/dardarin. *Neurobiol. Dis.*, 23, 329-341.

Greggio, E., Bisaglia, M., Civiero, L., Bubacco, L. (2011). Leuchine-rich repeat kinase 2 and alpha synuclein: intersecting pathways in the pathogenesis of Parkinson's disease? *Mol. Neurodeg*., **6**:6.

Gregory, A., Westaway, S. K., Holm, I. E., Kotzbauer, P. T., Hogart, P., Sonek, S., Coryyell, J. C., Nguyen, T. M., Nardocci, N., Zorzi, G., Rodriguez, D., Desguerre, I., Bertini, E., Simonati, A., Levinson, B., Dias, C., Barbot, C., Carrilho, I., Santos, M., Malik, I., Gitschier, J., Hayflick, S. J. (2008). Neurodegeneration associated with genetic defects in phospholipase A2. *Neurology*, **71**:18, 1402-1409.

Goedert, M. (1997). Familial Parkinson's disease: the awakening of □-synuclein. *Nature*, **388**, 232-233.

Goldberg, M. S., Lansbury, P. T. Jr. (2000). Is there a cause-and-effect relationship between alpha-synuclein fibrillization and Parkinson's disease? *Nat. Cell. Biol*., **2**, E115-119.

Golovko, M. Y., Rosenberger, T. A., Feddersen, S., Faergeman, N. J., Murphy, E.J. (2007). Alpha-synuclein gene ablation increases docosahexaenoic acid incorporation and turnover in brain phospholipids. *J. Neurochem.* **101**, 201–211.

Hamilton, B. A. (2004). α-Synuclein A53T substitution associated with Parkinson disease also marks the divergence of old word and new world primates. *Genomics*, **83**, 739-742.

Hampshire, D. J., Roberts, E., Crow, Y., Bond, J., Mubaidin, A., Wriekat, A. L., Al-Din, A., Woode, C. G. (2001). Kufor-Rakeb syndrome, pallido-pyramidal degeneration with supranuclear upgaze paresis and dementia, maps to 1p36. *J. Med. Genet.*, **38**, 680-682.

Hardy, J., Lewis, P., Revesz, T., Lees, A., Paisan-Ruiz, C. (2009). The genetics of Parkinson's syndromes: a critical review. *Current Opinion in Genetics & Development*, **19**, 254-265.

Hardy, J. (2010). Genetic analysis of pathways to Parkinson disease. *Neuron*, **68**:2, 201-206.

Hatcher, J. M., Pennell, K. D., Miller, G. W. (2008). Parkinson's disease and pesticides: a toxicological perspective. *Trends Pharmacol. Sci*., **29**, 322-329.

Healy, D. G., Abou-Sleiman, P. M., Cases, J. P., Ahmadi, K. R., Lynch, T., Gandhi, S., Muqit, M. M., Foltynie, T., Barker, R., Bhatia, K. P., Quinn, N. P., Lees, A. J., Gibson, J. M.,

Holton, J. L., Revesz, T., Goldstein, D. B., Wood, N. W. (2006). UCHL-1 is not a Parkinson's disease susceptibility gene. *Ann. Neurol.*, **59**,627-633.

Helenius, A., McCaslin, D. R., Fries, E:, Tandford, C. (1979). Properties of detergents. *Methods Enzymol.*, **56**, 734-749.

Henchcliffe, C., and Beal, M. F. (2008). Mitochondrial biology and oxidative stress in Parkinson disease pathogenesis. *Nature Clinical Practice Neurology*, **4**:11, 600-609.

Hokenson, M. J., Uversky, V. N., Goers, J., Yamin, G., Munishkina, L. A., Fink, A. L. (2004). Role of individual methionines in the fibrillation of methionine-oxidized alpha-synuclein. *Biochemistry*, **43**, 4621–4633.

Hoyer, W., Cherny, D., Subramaniam, V., Jovin, T.M. (2004). Impact of the acidic C-terminal region comprising amino acids 109-140 on alpha-synuclein aggregation in vitro. *Biochemistry*, **43**, 16233-16242.

Hyun, D. H., Lee, M., Hattori, N., Kubo, S., Mizuno, Y., Halliwell, B., Jenner, P. (2002). Effect of wild-type or mutant parkin on oxidative damage, nitric oxide, antioxidant defenses, and the proteasome. *J. Biol. Chem*., **277**, 28572-28577.

Huang, Z., de la Fuente-Fernandez, R., Stoessl, A. J. (2003). Etiology of Parkinson's disease. *Can. J. Neurol. Sci.*, 30, S10-S18.

Irvine, G. B., El-Agnaf, O. M., Shankar, G. M., Walsh, D. M. (2008). Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Mol. Med*., **14**, 451-464.

Issidorides, M.R., Mytilineou, C., Panayotacopoulou, M.T., Yahr, M. D. (1991). Lewy bodies in parkinsonism share components with intraneuronal protein bodies of normal brains. *J. Comp. Neurol.*, **309**, 150-160.

Iwai, A., Masliah, E., Yoshimoto, M., Ge, N., Flanagan, L., de Silva, H. A., Kittel, A., Saitoh, T. (1995). The precursor protein of non-A beta component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. *Neuron.,* **14**, 467-475.

Jakes, R., Spillantini, M. G., Goedert, M. (1994). Identification of two distinct synucleins from human brain. FEBS Letter, **345**:1, 27-32.

Jenner, P. (2003). Oxidative stress in Parkinson's disease. *Ann. Neurol.*, **53**, S26-S38.

Jin J., Li G. J., Davis J., Zhu D., Wang Y., Pan C. and Zhang J. (2007). Identification of novel proteins interacting with both a-synuclein and DJ-1. *Mol. Cell Proteomics*, **6**, 845–859.

Jo, E., McLaurin, J., Yip, C. M., George-Hyslop, P., Fraser, P. E. (2000). α-Synuclein membrane interactions and lipid specificity. *J. Biol. Chem.*, **275**:44, 34328-34334.

Johansen, K. K., White, L. R., Sando, S. B., Aesly, J. O. (2010). Biomarkers: Parkinson disease with dementia and dementia with Lewy bodies. *Parkinsonism. Relat. Disordes.*, **16**, 307-315.

Karyagina, I., Becker, S., Giller, K., Riedel, D, Jovin, T. M., Griesinger, C., Bennati, M. (2011). Electron paramagnetic resonance spectroscopy measures the distance between the external β-strands of folded α-synuclein in amyloid fibrils. *Biophysical Journal*, L01-L03.

Khurana, R., Ionescu-Zanetti, C., Pope, M., Li, J., Nielson, L., Ramírez-Alvarado, M., Regan, L., Fink, A. L., Carter, S. A. (2003). A general model for amyloid fibril assembly based on morphological studies using atomic force microscopy. *Biophys. J*., **85**, 1135-1144.

Kim, Y., Park, J., Kim, S., Song, S., Kwon, s. H., Kitada, T., Kim, J. M., Chung, J. (2008). PINK1 controls mitochondrial localization of Parkin through direct phosphorylation. *Biochemical and Biophysisical Research Communication*, **377**:3, 975-980.

Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., Shimizu, N. (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*, **329**, 605-608.

Kempster, P. A., Hurwitz, B., Lees, A.J. (2007). A new look at James Parkinson's *Essay on the Shaking Plasy*. *Neurology*, **69**, 482-485.

Krueger, R., Kuhn, W., Mueller, T., Woitalla, D., Graeber, M., Koesel, S., Przuntek, H., Epplen, J. T., Schoels, L., Riess, O. (1998). Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. **18**:2, 106-108.

Krueger, R., Mueller, T., Riess, O. (2000). Involvement of α-synuclein in Parkinson's disease and other neurodegenerative disorders. *Journal of Neuronal transmission*, **107**: 31-40.

Krueger, R., Sharma, M., Riess, O., Gasser, T., Van Broeckhoven, C., Theuns, J., Aasly, J., Annesi, G., Bentivoglio, A. R., Brice, A., Ana Djarmati, Elbaz, A., Farrer, M., Ferrarese, C., Gibson, J. M., Hadjigeorgiou, J. M., Hattori, N., Ioannidis, J. P. A., Jasinska-Myga, B., Klein, C., Lambert, J. C., Lesage, S., Lin, J. J., Lynch, T., Mellick, G. D., de Nigris, F., Opala, G., Prigione, A., Quattrone, A., Ross, O. A., Satake, W., Silburn, P. A., King Tan, E., Toda, T., Tomiyama, H., Wirdefeldt, K., Wszolek, Z., Xiromerisiou, G., Maraganore′ D. M., the Genetic Epidemiology of Parkinson's disease consortium (2011). A large-scale genetic association study to evaluate the contribution of *Omi/HtrA2* (PARK13) to Parkinson's disease. *Neurobiology of Aging*, **32**:3, 548.e9-348.e18.

Langston, J. W. (1985). MPTP neurotoxicity: an overview and characterization of phases of toxicity. *Life Sci*., **36**, 201-206.

Larsen, K., Bendixen, C. (2011). Characterization of the porcine FBX07 gene: the first step towards generation of a pig model for Parkinsonian pyramidal syndrome. *Mol. Biol. Rep.*, **39**, 1517-1526.

Lashuel, H. A., Petre, B. M., Wall, J., Simon, M., Nowak, R. J., Walz, T., Lansbury, P. T., (2001). α-Synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. *J. Mol. Biol.*, **322**, 1089-1102.

Lautier, C., Goldwurm, S., Duerr, A., Giovannone, B., Tsiaras, W. G., Pezzoli, G., Brice, A., Smith, R. J. (2008). Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease. *Am. J. Hu. Genet.*, **82**:4, 822-33.

Lavedan, C., Leroy, E., Dehejia, A., Buchholtz, S., Dutra, A., Nussbaum, R. L., Polymeropoulos, M. H. (1998). Identification, localization and characterization of the human gamma-synuclein gene. *Hum. Genet*., **103**, 106-112.

Lee, M., Hyun, D., Halliwell, B., Jenner, P. (2001). Effect of the overexpression of wild-type or mutant alpha-synuclein on cell susceptibility to insult. *J. Neurochem.*, **76**, 998-1009.

Lee, H. J, Choi, C., Lee, S. J. (2002). Membrane-bound alpha-synuclein has a high aggregation propensity and the ability to seed the aggregation of the cytosolic form. *J. Biol. Chem*., **277**, 671-678.

Lees, A. J. (2009). The Parkinson chimera. *Neurology*, **72**, S2-S11.

Lees, A. J., Hardy, J., Revesz, T. (2009). Parkinson's disease. *Lancet*, **373**, 2055-2066.

Leroy, E., Boyer, R., Auburger, G., Leube, B., Ulm, G., Mezey, E., Harta, G., Brownstein, M. J., Sobhanadditya, J., Chernova, T., Dehejia, A., Lavedan, C., Gasser, T., Steinbach, P. J., Wilkinson, K. D., Polymeropoulos, M. H. (1998). The ubiquitin pathway in Parkinson's disease. *Nature*, **395**, 451-452.

Lewy, F. H. (1923). Die Lehre vom Tonus und der Bewegung: zugleich systematische Untersuchungen zur Klinik, Physiologie, Pathologie und Pathogenese der Paralysis agitans. *Julius Springer Verlag*.

Li, H. T., Lin, X. J., Xie, Y. Y., Hu, H. Y. (2006). The early events of alpha-synuclein oligomerization revealed by photo-induced cross-linking. *Protein Pept. Lett*., **13**, 385-390.

Li, L., Funayama, M., Tomiyama, H., Li, Y., Yoshino, H., Sasaki, R., Kokubo, Y., Kuzuhara, S., Mizuno, Y., Hattori, N. (2010). No evidence for pathogenic role of GIGYF2 mutation in Parkinson disease in Japanese patients. *Neurosci. Lett.*, **479**:3, 245-248.

Liani, E., Eyal, A., Avraham, E., Shemer, R., Szargel, R., Berg, D., Bornemann, A., Riess, O., Ross, C. A., Rott, R., Engelender, S. (2004). Ubiquitylation of synphilin-1 and alpha-synuclein by SIAH and its presence in cellular inclusions and Lewy bodies imply a role in Parkinson's disease. *Proc. Natl. Acad. Sci.* USA, **101**, 5500-5505.

Lincoln, S., Vaughan, J., Wood, N., Baker, M., Adamson, J., Gwinn-Hardy, K., Lynch, T., Hardy, J., Farrer, M. (1999). Low frequency of pathogenic mutations in the ubiquitin carboxyterminal hydrolase gene in familial Parkinson's disease. *Molecular Neuroscience*, **10**:2, 427-429.

Liu, Y., Fallon, L., Lashuel, H. A., Liu, Z., Lansbury, P. T. (2002). The UCH-L1 gene encodes two opposing enzymatic activities that aVect alpha-synuclein degradation and Parkinson's disease susceptibility. *Cell*, **111**, 209-218.

Maraganore, D. M., Lesnick, T. G., Elbaz, A., Chartier-Halin, M. C., Gasser, T., Krueger, R., Hattori, N., Mellick, G. D., Quattrone, A., Satoh, J., Toda, T., Wang, J., Ioannidis, J. P., de Andrade, M., Rocca, W. A. (2004). UCHL1 is a Parkinson's disease susceptibility gene. *Ann. Neurol.*, **55**, 512-521.

Martins, L. M., Morrison, A., Klumpsch, K., Fedele, V., Moisoi, N., Teismann, P., Abuin, A., Grau, E., Geppert, M., Livi, G. P., Creasy, C. L., Martin, A., Hargreaves, I., Heales, S. J., Okada, H., Brandner, S., Schulz, J. B., Mak, T., Downward, J. (2004). Neuroprotective role of the reaper-related serine protease HtrA2/Omi revealed by targeted deletion in mice. *Molecular and Cellular Biology*, **24**:22, 9848-9862.

Mata, I. F., Wedemeyer, W. J., Farree, M., J., Taylor, J. P., Gallo, K., A. (2006). LRRK2 in Parkinsoni's disease: protein domains and functional insights. *TRENDS in Neuroscience*, **29**:5, 286-293.

McNulty, B. C., Tripathy, A., Young, G. B., Charlton, L. M., Orans, J. & Pielak, G. J. (2006). Temperature-induced reversible conformational change in the first 100 residues of alpha-synuclein. *Protein Sci.* **15**, 602–608.

McLean, P. J., Kawamata, H., Ribich, S., Hyman, B. T. (2000). Membrane association and protein conformation of alpha-synuclein in intact neurons. Effect of Parkinson's disease-linked mutations. *J. Biol. Chem*. **275**, 8812–8816.

Miake, H., Mizusawa, H., Iwatsubo, T., Hasegawa, M. (2002). Biochemical characterization of the core structure of alpha-synuclein filaments. *J. Biol. Chem*., **277**, 19213-19219.

Miller, D. W., Hague, S. M., Clarimon, J., Baptista, M., Gwinn-Hardy, K., Cookson, M. R., Singleton, A. B. (2004). Alpha-synuclein in blood and brain from familial Parkinson disease with SNCA locus triplication. *Neurology*, **62**, 1835-1838.

Mori, F., Tanji, K., Yoshimoto, M., Takahashi, H., Wakabayashi, K. (2002). Demonstration of alpha-synuclein immunoreactivity in neuronal and glial cytoplasm in normal human brain tissue using proteinase K and formic acid pretreatment. *Experimental Neurology*, **176**, 98-104.

Mounsey, R. B., and Teismann, P. (2011). Mitochondrial sysfunction in Parkinson's disease: pathogenesis and neuroprotection. *Parkinson's disease*. ID 617472.

Narendra, D., Tanaka, A., Suen, D. F., Youle, R. J. (2008). Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *Journal of Cell Biology*, **183**:5, 795-803.

Narendra, D., Min Jin, S., Tanaka, A., Suen, D. F., Gautier, C. A., Shen, J., Cookson, M. R., Youle, R. J. (2010). PINK1 is selectively stabilized on impared mitochondria to activate Parkin. *PLOS Biology*, **8**:1, e1000298.

Okochi, M., Walter, J., Koyama, A., Nakajo, S., Baba, M., Iwatsubo, T., Meijer, L., Kahle, P. J., Haass, C. (2000). Constitutive phosphorylation of the Parkinson's disease associated alpha-synuclein. *J. Biol. Chem*., **275**, 390–397.

Olzmann, J. A., Brown, K., Wilkinson, K. D., Rees, H. D., Huai, Q., Ke, H., Levely, A. I., Li, L., Chin, L. S. (2004). *Journal of Biological Chemistry*, **279**:27, 8506-8515.

Outerio, T. F., Putcha, P., Tetzlaff, J. E., Spoelgen, R., Koker, M., Carvalho, F, Hyman, B. T., McLean, P. J. (2008). Formation of toxic oligomeric α-synuclein species in living cell. *Plos*, **3**:4, e1867.

Paisan-Ruiz, C., Jain, S., Evans, E. W., Gilks, W.P., Simon, J., van der Brug, M., Lopez de Munain, A., Aparicio, S. , Gil, A. M., Khan, N., Johnson, J., Martinez, J. R., Nicholl, D., Carrera, I. M., Pena, A. S., de Silva, R., Lees, A., Marti-Masso, J. F., Perez-Tur, J., Wood, N. W., Singleton, A. B. (2004). Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron.*, **44**:4, 595-600.

Paisan-Ruiz, C., Bathia, K. P., Li, A., Hernandez, D., Davis, M., Wood, N. W., Hardy, J., Houlden, H., Singleton, A., Schneider, S. A. (2008). Characterization of PLA2G6 as a locus for dystonia-parkinsonism. *Annals of Neurology*, **65**:1, 19-23.

Paisan-Ruiz, C., Guevara, R., Federoff, M., Hanagasi, H., Sina, F., Elehi, E., Schneider, S. A., Schwingenschuh, P., Bajaj, N., Emre, M., Singleton, A. B., Hardy, J., Bhatia, K. P., Brandner, S., Lees, A. J., Houlden, H. (2010). Early-onset L-dopa-responsive parkinsonism with pyramidal signs due to *ATP13A2*, *PLA2G6*, *FBXO7* and *Spactasin* mutation. *Movement Disorders*, **25**:12, 1791-1800.

Paleček, E., Ostatná, V., Masarík, M., Bertoncini, C. W., Jovin, T. M. (2008). Changes in interfacial properties of alpha-synuclein preceding its aggregation. *Analyst*, **133**, 76-84.

Papapetropoulos, S., Paschalis, C., Athanassiadou, A., Papadimitriou, A., Ellul, J., Polymeropoulos, M., Papapetropoulos, T. (2001). Clinical phenotype in patients with α-synuclein Parkinson's disease living in Greece in comparison with patients with sporadic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry*, **70**:5, 662-665.

Pankrats, N., Nichols, W. C., Uniacke, S. K., Halter, C., Rudolph, A., Shults, C., Conneally, P. M., Foroud, T. (2002). Genome screen to identify susceptibility genes for Parkinson disease in a sample without parkin mutations. *Am. J. Hum. Genet.*, **71**, 124-135.

Park, J., Lee, S. B., Lee, S., Kim, Yongsung, K., Song, S., Kim, S., Bae, E., Kim, J., Shong, M., Kim, J, M. & Chung, J. (2006). Mitochondrial dysfunction in *Drosophila PINK1* mutants is complemented by *parkin*. *Nature*, **441**:29, 1157-1161.

Plun-Favreau, H., Klupsh, K, Moisoi, N., Gandhi, S., Kjaer, S., Frith, D., Harvey, K., Deas, E., Harvey, R. J., McDonald, N., Wood, N. W., Martins, L. M., Downward, J. (2007). The mitochondrial protease HtrA2 is regulated by Parkinson's disease-associated kinase PINK1. Nature Cell Biology, **9**:11, 1243-1252.

Polverino de Laureto, P., Tosatto, L., Frare, E., Marin, O., Uversky, V. N. & Fontana, A. (2006). Conformational properties of the SDS-bound state of alpha-synuclein probed by limited proteolysis: unexpected rigidity of the acidic C-terminal tail. *Biochemistry,* **45**, 11523–11531.

Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E. S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W. G, Lazzarini, A. M., Duvoisin, R. C., Di

Iorio, G., Golbe, L. I. & Nussbaum, R. L. (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*, **276**, 2045–2047.

Qin, Z., Hu, D., Han, S., Reaney, S. H., Di Monte, D. A., Fink, A. L. (2007). Effect of 4-hydroxy-2-nonenal modification on alpha-synuclein aggregation. *J. Biol. Chem.*, **282**, 5862-5870.

Quin, Z., Hu, D., Han, S., Hong, D. P., Fink, A. L. (2007). Role of different regions of α-synuclein in assembly of fibrils. *Biochemistry,* **46**, 13322-13330.

Ramirez, A., Heimbach, A., Gruendemann, J., Stiller, B., Hampshire, D., Cid, P., Goebel, I., Mubaidin, A. F., Wriekat, A. L., Roeper, J., Al-Din, A., Hillmer, A. M., Karsak, M., Liss, B., Woods, C. G. Beherens, M. I., Kubish, C. (2006). Hereditary parkinsonism with dementia is caused by mutation in *ATP13A2* gene, encoding a lysosomal type 5 P-type ATPase. *Nature Genetics*, **38**, 1184-1191.

Rodrigues e Silva, A. M., Geldsetzer, F., Holdorff, B., Kielhorn, F. W., Balzer-Geldsetzer, M., Oertel, W. H., Hurtig, H., Dodel, R. (2010). Who was the man who discovered the „Lewy Bodies“? *Movement Disorders*, **25**, 1765-1773.

Ross, O. A., Soto, A. I., Vilarino-Guell, C., Heckman, M. G., Diehl, N. N., Hulihan, M. M., Aasly, J. O., Sando, S., Gibson, J. M., Lynch, T., Krygowska-Wajs, A., Opala, G., Barcikowska, M., Czyzewsky, K., Uitti, R. J., Wszolek, Z. K., Farrer, M. J. (2008). Genetic vaiation of *Omi/HtrA2* and Parkinson’s disease. *Parkinsonins and Related Disorders*, **14**:7, 539-543.

Saha, S., Guillily, M.D., Ferree, A., Lanceta, J., Chan, D., Ghosh, J., Hsu, C.H., Segal, L., Raghavan, K., Matsumoto, K., Hisamoto, N., Kuwahara, T., Iwatsubo, T., Moore, L., Goldstein, L., Cookson, M., Wolozin, B. (2009). LRRK2 modulates vulnerability to mitochondrial dysfunction in Caenorhabditis elegans. *J. Neurosci*., **29**, 9210-9218.

Sharon, R., Goldberg, M. S., Bar-Josef, I., Betensky, R. A., Shen, J. & Selkoe, D. J. (2001). Alpha-Synuclein occurs in lipid-rich high molecular weight complexes, binds fatty acids, and shows homology to the fatty acid-binding proteins. *Proc. Natl. Acad. Sci. U S A*, **98**, 9110–9115.

Shendelman, S., Jonason, A., Martinat, C., Leete, T., Abeliovic, A. (2004). DJ-1 is a redox-dependent molecular chaperone that inhibits α-synuclein aggregate formation. *PLoS Biol.*, **2**:11, e362.

Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K., Suzuki, T. (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat. Genet*., **25**, 302-305.

Shimura, H., Schlossmacher, M. G., Hattori, N., Frosch, M. P., Trockenbacher, A., Schneider, R., Mizuno, Y., Kosik, K. S., Selkoe, D. J. (2001). Ubiquitination of a new form of alpha-synuclein by parkin from human brain: implications for Parkinson’s disease. *Science*, **293**, 263–269.

Shults, C. W. (2006). Lewy bodies. *Proc. Natl. Acad. Sci. U S A*, **103**, 1661-1668.

Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., Lincoln, S., Crawley, A., Hanson, M.,

Maraganore, D., Adler, C., Cookson, M. R., Muenter, M., Baptista, M., Miller, D., Blancato, J., Hardy, J. & Gwinn-Hardy, K. (2003). alpha-Synuclein locus triplication causes Parkinson's disease. *Science*, **302**, 841.

Spillantini, M.G., Schmidt, M.L., Lee, V. M., Trojanowski, J.Q., Jakes R., Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature*, **388**, 839-840.

Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. (1998). alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci. USA*, **95**, 6469–6473.

Sriram, S. R., Li, X., Ko, H. S., Chung, K. K., Wong, E., Lim, K. L., Dawson, V. L., Dawson, T. M. (2005). Familial-associated mutations differentially disrupt the solubility, localization, binding and ubiquitination properties of parkin. *Hum. Mol. Genet*., **14**, 2571-2586.

Surguchov, A., Surgucheva, I., Solessio, E., Baehr, W. (1999). Synoretin : a new protein belonging to the synuclein family. *Molecular and Cellular Neuroscience*. **13**:2, 95-103.

Souza, J. M., Giasson, B. I., Lee, V. M. & Ischiropoulos, H. (2000). Chaperone-like activity of synucleins. *FEBS Lett.* **474**, 116–119.

Tan, E. K., Ho, P., Tan, L., Prakash, K. M., Zhao, Y. (2009). *PLA2G6* mutations and Parkinson's disease. *Annals of Neurology*, **67**:1, 148.

Thirunavukkuarasu, S., Jares-Erijman, E. A., Jovin, T. M. (2008). Multiparametric fluorescence detection of early stages in the amyloid protein aggregation of pyrene-labeled alpha-synuclein. *J. Mol. Biol*., **378**, 1064-1073.

Thomas, B., and Beal, M. F. (2007). Parkinson's disease. *Hum. Mol. Genet*., **16**, 183-193.

Ueda, K., Fukushima, H., Masliah, E., Xia, Y., Iwai, A., Yoshimoto, M., Otero, D. A., Kondo, J., Ihara, Y., Saitoh, T. (1993) Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl. Acad. Sci. U S A*, **90**, 11282-11286.

Uversky, V. N., Gillespie, J. R., Fink, A. L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins*, **41**, 415-427.

Uversky, V. N., Li, J., Fink, A. L. (2001). Evidence for a partially folded intermediate in α-synuclein fibril formation. *Journal of Biological Chemistry*, **276**:14, 10737-10744.

Uversky, V. N. (2003). A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J. Biomol. Struct. Dyn*., **21**, 211-234.

Uversky, V. N., Yamin, G., Munishkina, L. A., Karymov, M. A., Millett, I. S., Doniach, S., Lyubchenko, Y. L., Fink, A. L. (2005). Effects of nitration on the structure and aggregation of alpha-synuclein. *Brain Res. Mol. Brain Res*., **134**, 84-102.

Uversky, V. N. (2007) Neuropathology, biochemistry, and biophysics of α-synuclein aggregation. *J. Neurochem*. **103,** 17–37.

Valente, E. M., Abou-Sleiman, P. M., Caputo, V., Muqit, M. M., Harvey, K., Gispert, S., Ali, Z., Del Turco, D., Bentivoglio, A. R., Healy, D. G., Albanese, A., Nussbaum, R., González-Maldonado, R., Deller, T., Salvi, S., Cortelli, P., Gilks, W. P., Latchman, D. S., Harvey, R. J., Dallapiccola, B., Auburger, G., Wood, N. W. (2004). Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, **304**, 1158-1160.

Vilar, M., Chou, H. T., Luehrs, T., Maji, S. K., Riek-Loher, D., Vere, R., Manning, G., Stahlberg, H., Riek, R. (2008). The fold of α-synuclein fibrils. *Proc. Natl. Acad. Sci. USA*, **105**:25, 8637–8642.

Volles, M. J., Lee, S. J., Rochet, J. C., Shtilerman, M. D., Ding, T. T., Kessler, J. C., Lansbury, P. T Jr. (2002). Vesicle permeabilization by protofibrillar alpha-synuclein: implications for the pathogenesis and treatment of Parkinson's disease. *Biochemistry*, **40**, 7812-7819.

van der Brug, M. P., Blackinton, J., Chandran, J., Hao, L. Y., Lal, A., Mazan-Mamczarz, K., Martindale, J., Xie, C., Ahmad, R., Thomas, K. J., Beilina, A., Gibbs, J. R., Ding, J., Myers, A. J., Zhan, M., Cai, H., Bonini, N. M., Gorospe, M., Cookson, M. R. (2008). RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways. *Proc Natl Acad Sci*, **105**, 10244-10249.

Volles, M. J., Lansbury, P. T. Jr. (2002). Vesicle permeabilization by protofibrillar alpha-synuclein is sensitive to Parkinson's disease-linked mutations and occurs by a pore-like mechanism. *Biochemistry*, **41**, 4595-4602.

Wakabayashi, K., Tanji, K., Mori, F., Takahashi, H. (2007). The Lewy body in Parkinson’s disease: molecules implicated in the formation and degeneration od a-synuclein aggregates. *Neuropathology*, **27**, 494-506.

Walle, L. V., Lamkanfi, M., Vandenabeele, P. (2008). The mitochondrial serine protease Htra2/Omi: an overview. *Cell Death and Differentiation*, **15**, 453-460.

Wang, A., Costello, S., Cockburn, M., Zhang, X., Bronstein, J., Ritz, B (2011). *European Journal of Epidemiology*, **26**, 547-555.

Webb, J. L., Ravikumar, B., Atkins, J., Skepper, J. N., Rubinsztein, D. C. (2003). α-synuclein is degraded by both autophagy and the proteosome. *J. Biol. Chem.*, **278**, 25009-25013.

Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., Lansbury, P. T. (1996). NACP, a protein implicated in Alzheimer disease comprehension and learning, is natively unfolded. *Biochemistry*, **35**:43, 13709-13705.

West, A. B., Maidment, N. T. (2004). Genetics of parkin-linked disease. *Hum. Genet.*, **114**, 327-336.

West, A. B., Moore, D. J., Biskup, S., Bugayenko, A., Smith, W. W., Ross, C. A., Dawson, V. L., Dawson, T. M. (2005). Parkinson's disease-associated mutations in leucine-

rich repeat kinase 2 augment kinase activity. *Proc. Natl. Acad. Sci. U S A*, **102**, 16842-16847.

Willingham, S., Outeiro, T. F., DeVit, M. J., Lindquist, S. L., Muchowski, P. J. (2003). Yeast genes that enhance the toxicity of a mutant huntingtin fragment or alpha-synuclein. *Science*, **302**, 1769-7172.

Wood, S. J., Wypych, J., Steavenson, S., Louis, J. C., Citron, M., Biere, A. L. (1999). α-Synuclein fibrillogenesis is nucleation dependent. *The Journal of Biological Chemistry*, **278**:28, 19509-19512.

Wood-Kaczmar, A., Gandhi, S., Wood, N. W. (2006). Understanding the molecular causes of Parkinson's disease. *Trends Mol. Med*., **12**, 521-528.

Wright, P. E., Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol*., **293**, 321-331.

Yoritaka, A., Hattori, N., Uchida, K., Tanaka, M., Stadtman, E. R., Mizuno, Y. (1996). Immunohistochemical detection of 4-hydroxynonenal protein adducts in Parkinson disease. *Proc. Natl. Acad. Sci. USA*, **93**, 2696 -2701.

Zarranz, J. J., Alegre, J., Comez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atares, B., Llorens, V., Gomez Tortosa, E., del Ser, T., Munoz, D. G., de Yebenes, J. G. (2003). The new mutation, E46K, of a □-synuclein causes parkinson and Lewy body dementia. *Annals of Neurobiology*, **55**:2, 164-173.

Zhang, L., Shimoji, M., Thomas, B., Moore, D. J., Yu, S. W, Marupudi, N. I., Torp, R., Torgner, I. A., Ottersen, O. P., Dawson, T. M., Dawson, V.L. (2005). Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis. *Hum. Mol. Genet*., **14**, 2063-2073.

Zhong, N., Kim, C. Y., Rizzu, P., Geula, C., Porter, D. R., Pothos, E. N., Squitieri, F., Heutink, P., Xu, J. (2006). DJ-1 transcriptionally up-regulates the human tyrosine hydroxylase by inhibiting the sumoylation of pyrimidine tract-binding protein-associated splicing factor. *Jour. of Biol. Chem.*, **281**:30, 20940-20948.

Zhu, M., Rajamani, S., Kaylor, J., Han, S., Zhou, F., Fink, A. L. (2004). The flavonoid baicalein inhibits fibrillation of alpha-synuclein and disaggregates existing fibrils. *J. Biol. Chem*.,**279**, 26846-26857.

Zhou, W., Zhu, M., Wilson, M. A., Petsko, G. A., Fink, A. L. (2005). The oxidative state of DJ-1 regulates its chaperone activity toward □-synuclein. *Journal of Molecular Biology*, **356**:4, 1036-1048.

Zimprich, A., Biskup, S., Leitner, P., Lichtner, P., Farrer, M., Lincoln, S., Kachergus, J., Hulihan, M., Uitti, R. J., Calne, D. B., Stoessl, A. J., Pfeiffer, R. F., Patenge, N., Carbajal, I. C., Vieregge, P., Asmus, F., Müller-Myhsok, B., Dickson, D. W., Meitinger, T., Strom, T. M., Wszolek, Z. K., Gasser, T. (2004). Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*., **44**, 601-607.

Zimprich, A., Schulte, C., Reinthaler, E., Haubenberger, D., Balzar, J., Lichtner, P., El Tawil, S., Edris, S., Foki, T., Pirker, W., Katzenschlager, R., Daniel, G., Bruecke, T., Auff,

E., Gasser, T. (2009). PARK11 gene (GIGYF2) variants Asn56Ser and Ans457Thr are not pathogenic for Parkinson's disease. *Parkinsonism Relat. Disord.*, **15**:7, 532-534.

Supplementary Material

I. Primary structure

αS: α-Synuclein

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTN
VGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEE
GYQDYEPEA

NN dimer

MD**C**FMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTN
VGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEE
GYQDYEPEA

MD**C**FMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVT
NVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMP
SEEGYQDYEPEA

CC dimer

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTN
VGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEE
GYQDYEPEA**GC**

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVT
NVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMP
SEEGYQDYEPEA**GC**

NC dimer

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTN
VGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEE
GYQDYEPEA**RS**MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATV
AEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPD
NEAYEMPSEEGYQDYEPEA

DC dimer

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVT
NVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEAAGKTKEGVLYVGSKTKEGVVHGVA
TVAEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPV
DPDNEAYEMPSEEGYQDYEPEA

II. Proteolytic mapping: identification of peptides by ESI-QTOF

Table I. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of aS by proteinase K.

RT[a]	Experimental Mass (Da)[b]	Calculated Mass (Da)[c]	Fragment
10,65	887,46±0,04	887,51	7-15 / 9-17
	700,37±1,06	701,33	24-31
	607,23±0,01	607,25	136-140
	671,38±0,98	672,28	115-120 / 116-121
11,11	729,52±0,05	729,44	10-16 / 95-100
	958,74±0,21	958,54	9-18
	931,69±0,49	931,43	98-105
12,41	1772,64±0,02	1772,67	126-140
13,1	1149,84±0,55	1150,23	116-125
	1625,24±0,53	1625,77	107-121
	3150,72±1,02	3152,31	107-134
13,39	769,34±0,36	769,98	1-6
	638,30±0,02	638,31	2-6
	2256,98±0,46	2256,18	93-113
17,5	5389,81±0,06	5389,74	93-140

RT[a]	Experimental Mass (Da)[b]	Calculated Mass (Da)[c]	Fragment
27,839	5663,25	5662,62	1-56
28,806	8815,5	8815,58	57-140
	7058,85	7059,79	57-125
30,216	11411,98	11412,59	31-140
	12596,29	12596,1	19-140
30,53	9655,9	9656,8	31-125
33,054	14458,8	14460,19	1-140
33,61	12702,8	12704,39	1-125
34,432	11330,01	11327,97	1-113

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.13 upper panels) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of aS.

Table II. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of aS by trypsin.

RT[a]	Experimental Mass (Da)[b]	Calculated Mass (Da)[c]	Fragment
9,88	829,45	829,43	24-32
11,06	1071,6 ±0,01	1072,22	11-21
	872,47	872,47	13-21
11,99	1294,68 ±0,02	1295,46	46-58
12,7	950,52	951,09	35-43
	589,83	1180,36	35-45
13,6	2157,2 ±0,89	2156,18	59-80
	1927,39 ±0,047	1927,04	61-80
13,9	1477,79 ±0,03	1477,78	81-96
16,61	4958,55 ±0,33	4958,2	97-140
16,92	4829,95 ±0,48	4830,03	98-140
17,24	4287,7 ±0,02	4288,43	103-140

RT[a]	Experimental Mass (Da)[b]	Calculated Mass (Da)[c]	Fragment
39,24	12039,3 ±0,4	12039,26	24-140
	11227,51 ±0,27	11227,37	33-140
	13123,54 ±0,51	13123,50	13-140
	13322,94 ±0,54	13322,75	11-140
30,76	10998,39 ±0,07	10998,09	35-140
	6956,44 ±0,41	6956,96	33-102
31,2	6727,32 ±0,31	6727,68	35-102
	8552,39 ±0,84	8552,62	35-119
	6185,27 ±0,69	6186,07	35-97
	6414,93 ±0,64	6415,35	33-97
32,69	14460,53 ±0,22	14460,19	1-140
33,57	6057,28 ±0,04	6057,90	35-96

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.13 upper panels) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of aS.

Table III. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of NN dimer by proteinase K.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
10,63	887,67	888,03	7-15 (147-155) / 9-17 (149-157)
	607,38	607,25	136-140 (276-280)
	644,37	644.35	26-32
11,1	958,80	958,54	9-18
	1029,80	1029,58	9-19
	1183,68 ±0,23	1183,43	121-130 / 126-135
12	1659,21 ±0,64	1659,86 / 1659,79	94-108 / 102-116
	1442,06 ±0,55	1442,42	128-139
	1541,14 ±0,08	1541,68	96-109
12,38	1772,72 ±0,40	1773,81	126-140
	1149,72 ±0,76	1150,23	116-125
	1457,68 ±0,04	1457,59	112-124
13,37	3149,79 ±0,73	3150,22	114-140
17,5	5391,97 ±1,49	5390,88	1-22ss1-18
	4310,77 ±0,55	4310,24	1-48 / 93-140

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
13,34	3798,41	3728,26	19-56
	3727,19	3728,26	20-56
	3870,37	3870,42	18-56
	3149,8	3079,15	114-140
28,8	8816,21	8815,58	57-140
	7059,69	7059,79	57-125
30,14	12597,42	12596,1	19-140
	12526,23	12525,83	20-140
	10841,56	10841,12	19-125
30,58	9464,32	9464,7	19-113

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of NN dimer.

Table IV. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of NN dimer by trypsin.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
11,1	687,85	687,79	97-102
	1071,93	1072,23	11-21
	8712,5	872,97	13-21
	686,85	687,79	97-102
12,03	1296,13 ±0,49	1295,46	46-58
12,7	950,29 ±0,91	951,09	35-43
	2469,26	2469,39	10-34
13,6	1926,75	1927,04	61-80
14	1544,64 ±0,01	1455,95	1-6ss1-6
	1477,85 ±0,01	1477,78	46-61 / 221-236
17	4954,55	4954,7	32-80
	2856,32 ±0,45	2856,31	6-34
17,4	4959,88 ±0,89	4958,2	97-140
	4829,85 ±0,56	4830,03	98-140
	4288,03 ±0,032	4288,43	103-140

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
30,31	12039,18 ±0,93	12039,26	1-24ss1-140
	10994,53	10993,82	1-12ss1-96
	13798 ±1,58	13795,02	1-99ss1-96
30,89	12015,88 ±0,93	12014,96	1-87ss1-32
	12074 ±2,93	12076,46	1-22ss1-96

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.15) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of NN dimer.

Table V. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of CC dimer by proteinase K.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
10,28	1201,12 ±1,04	1202,33	17-28 / 18-29 / 19-30
	1272,14 ±0,31	1272,46	26-38 / 8-20
	644,09	644,64	132-136
11,1	958,20	959,11	9-18 / 23-32
	931,17	931,96	98-105
	730,10	729,87	10-16 / 95-100
12,46	1658,54 ±0,97	1659,79 / 1659,86	102-116 / 94-108
	830,32 ±0,55	830,98	8-15 / 31-38
	1316,48 ±0,64	1316,75	5-17
	1387,50 ±1,09	1388,58	9-22
	1441,46 ±0,08	1441,55	128-139
	1828,6 ±0,18	1828,85	100-116
	1930,6 ±0,72	1930,02	39-57
12,5	2114,98 ±0,22	2115,17	101-111
12,9	1149,56 ±0,22	1149,46	116-125
	3865,06 ±1,49	3864,17	92-126
13,11	2052,78 ±1,49	2052,09	95-113
16,5	3632,46 ±1,49	3633,95	93-125

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
28,172	10928,52	10932,86	57-142ss126-142
28,665	7059,8	7059,79	57-125
	5683,24	5683,37	57-113
28,89	14671,84	14673,09	19-142ss126-142
	14603,49	14602,02	20-142ss126-142
29,876	10842,19	10841,12	19-125
	10770,97	10770,05	20-125
	10912,83	10912,19	18-125
30,498	9465,42	9464,7	19-113
	9394,14	9393,62	20-113
	9536,59	9535,78	18-113
32,12	16555,24	16540,38	1-142ss126-142
32,3	12706,82	12704,4	1-125
33,79	11330,69	11327,98	1-113
34,81	29225,24	29226,76	1-142ss1-142

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of CC dimer.

Table VI. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of CC dimer by trypsin.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
9,9	829,68	829,43	24-32
11,1	1072,43 ±0,01	1072,22	11-21
	872,65	872,47	13-21
12	1295,86 ±0,55	1295,46	46-58
12,7	951,45	951,09	35-43
13,4	2155,86 ±0,55	2156,18	59-80
13,7	1927,59 ±0,08	1927,04	61-80
13,9	1477,04 ±0,12	1477,78	81-96

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
23,8	9566,16 ±0,47	9565,98	6-102
24,1	8894,7 ±0,07	8895,24	103-142ss103-142
	9437,39 ±0,32	9436,86	103-142ss98-142
24,77	9565,21 ±0,01	9565,02	103-142ss97-142
32	19068,49	19067,85	44-142ss58-142
	11685,96	11679,44	98-142ss81-142
36	29226,77	29226,76	1-142ss1-142
	17972,66	17975,67	80-142ss142-114

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.15) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of CC dimer.

Table VII. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of NC dimer by proteinase K.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
10,32	983,52±0,02	983,54	42-50
	1201,28±0,73	1202,33	17-28 / 18-29 / 19-30
11	959,50±0,01	958,54	9-18+equiv
	887,48±0,21	887,51	7-15 +eq / 9-17+eq
12,24	1928,92±0,03	1928,77	126-141
12,4	1772,8±0,45	1772,67	126-140+eq
	2002,22±0,45	2001,79	128-144
	5807±0,30	5806,75	140-196
	2116,32±1,71	2115,17	31-50
12,8	1236,7±1,02	1236,39	3-13+eq
	638,33±0,60	638,78	2-6+eq
	1149,68±0,62	1149,46	116-125
13,16	2272,26±2,54	2272,24 7 / 2271.21	38-59
	1236,67	1236,69	3-13
	1277,66±1,54	1278,51	126-143
	660,3±0,5	659,24	119-124
	778,91±0,55	779,3	134-139
	676,27±0,05	676,31	111-116
	638,31±0,02	638,31	2-6
16,79	5544,12±1,54	5545,93	93-141
	1149,62±0,16	1149,46	116-125
	813,48±0,21	813,32	113-133
	887,46±0,11	887,51	7-15
	958,49±0,6	958,54	9-18
	1108,81±1	1107,65	3-12
	5186,8	5185,51	95-140
	5359,8	5389,74	93-140

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
13,13	3799,22	3728,26	19-56
	3728,19	3728,26	20-56
	2614,58	2615,03	31-56
27,888	5663,25	5662,62	1-56
28,824	7059,21	7059,79	57-125
	8815,5	8815,58	57-140
30,208	12596,24	12596,1	19-140
30,537	10771,33	10841,12	19-125
	10840,24	10770,05	20-125
	12752,41	12753,1	19-141
33,463	14615,83	14616,37	1-141
	14934,1	14834,65	1-143
34,113	12704,27	12704,4	1-125

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of NC dimer.

Table VIII. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of NC dimer by trypsin.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
9,9	828,96 ±0,43	829,43	23-32 / 166-174
11,08	1071,02 ±0,7	1071,59	11-21 / 153-163
	871,91 ±0,44	872,46	13-21 / 155-163
12	1293,89 ±0,65	1294,69	46-58 / 188-200
12,7	1179,67	1180,36	33-43
13,77	768,66	769,41	1-6 / 143-148
	2155,27 ±0,79	2156,18	59-89 / 201-22
14	4281,9 ±0,13	4281,98	2-42 / 144-185
16,4	4985,79 ±0,74	4986,22	98-141
16,9	4287,45 ±0,65	4288,43	245-282
17,3	4829,22 ±0,34	4830,03	98-140

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
27,49	1926,75 ±0,42	1927,04	61-89 / 203-222
28,62	2154,85 ±0,45	2156,18	59-80 / 201-222
30,22	12038,77 ±0,31	12039,26	166-282
	12195,01 ±0,32	12195,45	24-141
	5342,1	5341,7	96-141
	5317	5312,16	137-187
32,66	14616,66 ±0,32	14616,37	1-141
33,72	1777,85	1477,78	81-96 / 223-238
38,09	4953,7 ±0,05	4953,76	138-185
	3213,7 ±0,32	3212,74	27-58 / 169-200

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of NC dimer.

Table IX. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of DC dimer by proteinase K.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
11	1201,76 ±0,12	1201,63	19-30
	1130,73 ±0,80	1130,59	20-30
	983,63 ±0,12	983,54	42-50 / 118-126
	896,61 ±0,09	896,51	43-50 / 119-126
	1272,84 ±0,02	1272,67	18-30
	1114,25 ±0,53	1114,49	177-187
	1040,68 ±0,27	1040,56	41-50 / 117-126
	768,51 ±0,01	768,41	44-50 / 120-126
11,3	1139,86 ±0,33	1139,63	40-50 / 116-126
	1001,72 ±0,32	1001,55	17-26
	887,69 ±0,42	887,51	7-15 / 9-17
11,6	958,55 ±0,02	958,54	9-18
	888,00 ±0,51	887,46	9-17
	1359,74 ±0,29	1359,96	39-51 / 115-117
	743,44 ±0,32	743,29	11-18
	671,90 ±0,50	672,21	11-17 / 12-18
11,7	1183,50 ±0,17	1183,63	202-211
12,4	1128,72 ±0,03	1128,75	7-18
	1057,68 ±0,01	1057,67	7-17
	1037,66 ±0,80	1037,60	46-56 / 122-132
12,6	1829,14 ±0,32	1829,50	55-72
	1539,95 ±0,61	1540,46	1-14
	1387,91 ±0,22	1388,13	5-18
	1316,86 ±0,49	1317,05	5-17
	1440,88 ±0,49	1441,84	204-215
13	1772,89 ±0,62	1773,24	202-216
	2001,85 ±0,49	2001,68	53-72 / 129-148
	1393,73 ±0,22	1393,87	190-201
	1882,31 ±0,15	1881,62	44-61 / 120-137
13,5	1236,91 ±0,08	1236,85	31-42 / 107-118
	1131,85 ±0,51	1131,43	51-61 / 53-64 / 54-63
	1217,92 ±0,39	1217,63	75-87 / 151-163
	2519,91 ±0,51	2520,35	98-121 / 97-120
13,7	1149,9 ±0,46	1149,64	31-41 / 107-117
	3759,97 ±0,31	5760,22	144-182
	975,77 ±0,15	975,53	53-61 / 129-137
14	3149,99 ±0,11	3150,23	190-216
	2640,14 ±0,80	2640,91	172-195
	2880,32 ±0,21	2880,53	41-69 / 117-145
	964,77 ±0,34	964,52	33-41 / 109-117
	2738,61 ±0,68	2738,34	171-195
	2288,25 ±0,58	2288,85	181-201
	1018,63 ±0,16	1018,46	193-201

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
29,5	8818,04 ±0,10	8818,09	32-119
	7059,69 ±0,04	7059,79	133-201
	5684,30 ±0,39	5684,45	30-87 / 106-163
	5797,15 ±0,08	5797,61	31-89 / 107-165
	6063,90 ±0,58	6062,30	57-117
	5744,26 ±0,10	5743,51	57-114 / 37-95 / 113-171
29,9	6086,68 ±0,11	6085,91	53-114
	6406,18 ±0,88	6405,27	53-117 / 61-125
	9157,64 ±0,24	9157,98	129-216
	7402,35 ±0,10	7402,18	129-201
30,6	10280,90 ±0,76	10280,26	118-216
	10192,89 ±0,75	10193,19	119-216
	7211,06 ±0,07	7211,20	*45*-117
	7527,40 ±0,11	7527,56	*42*-117
	8437,18 ±0,21	8437,40	119-201
	8526,36 ±0,21	8424,48	118-201
30,9	8660,06 ±0,48	8659,88	31-117
	7059,89 ±0,29	7059,79	133-201
	7048,37 ±0,59	7148,06	118-*189*
	8342,67 ±0,34	8340,56	31-114
	16325,85 ±0,56	16325,13	57-216
32,2	6933,11 ±0,41	6932,81	120-189
	7470,42 ±0,16	7470,51	52-126
	14570,46 ±0,76	14570,53	43-189
	16503,67 ±0,98	16503,74	29-*194* / 55-210
	20035,01 ±0,57	20035,38	20-216
	20106,02 ±0,38	20106,45	19-216
35,2	18922,64 ±0,58	18922,14	31-216
	17789,17 ±0,22	17789,81	42-216
	21969,71 ±0,67	21969,73	1-216
	20213,48 ±0,54	20213,94	1-201

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of DC dimer.

Table X. Molecular masses of soluble (upper table) and insoluble (lower table) fragments obtained by proteolysis of DC dimer by trypsin.

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
10,8	829,43 ±0,10	829,33	24-32
12	872,51 ±0,11	872,4	13-21
	1071,59 ±0,01	1071,35	11-21
12,8	1294,90 ±0,64	1295,46	46-58 / 122-134
	1523,94 ±0,41	1524,16	44-58/46-60/120-134/122-130
13,5	950,09 ±0,52	950,49	35-43/111-119
	2156,70 ±0,52	2156,12	59-80/135-156
	1605,98 ±0,52	1606,27	81-97/157-173
14,6	1927,11 ±0,51	1927,6	61-80/137-156
	769,38 ±0,31	769,41	1-6
	1477,91 ±0,22	1478,09	81-96/157-172
18,6	4287,4 ±0,46	4287,15	179-216
	4830,0 ±0,66	4829,54	174-216
	4957,2 ±0,42	4957,63	173-216

RT	Experimental Mass (Da)	Calculated Mass (Da)	Fragment
28	3664,	3434,32	*46-80 / 122-156*
	3663,99 ±0,01	3663,4	*44-80 / 120-156*
28,5	4596,75 ±0,38	4597,24	*35-80 / 111-156*
	4826,11 ±0,45	4856,55	*33-80 / 109-156*
	5638,29 ±0,30	5638,46	24-80
29,7	4858,40 ±0,37	4858,49	*59-108*
	5387,01 ±0,01	5388,12	*81*-134
	4288,38 ±0,03	4288,43	179-216
30,6	3516,45 ±0,48	3516,42	137-97
	4058,22 ±0,02	4058,63	137-*178*
31,9	10065,83 ±0,24	10065,01	120-216
	9835,7 ±0,10	9835,74	122-216
33,1	10998,95 ±0,72	10997,49	111-216
	5565,09 ±0,46	5564,75	122-*178*
	5225,95 ±0,02	5124,23	120-*173*
37,8	19548,8 ±0,38	19548,8	24-216

[a] The peptides obtained from the proteolysis of aS with proteinase K were purified by RP-HPLC (see fig. 3.14) and listed in order of retention times (RT).

[b] Experimental molecular masses determined by ESI-MS.

[c] Molecular masses calculated from the amino acid sequence of DC dimer.

III. Main analytical techniques

III.i. Circular Dichroism

Circular dichroism (CD) is the election spectroscopic technique used to determine the presence and the type of secondary structure in peptides and proteins, to evaluate their tertiary structure and to monitor structural transitions during unfolding and folding processes (Woody, 1995; Kelly et al., 2005). CD refers to the differential adsorption of the left and right circularly polarized components of plane-polarised radiation. This effect will occur when a chromofore is chiral (optically active) either (a) intrinsically by reason of its structure, or (b) by being covalently linked to a chiral centre, or (c) by being placed in an asymmetric environment. In practice the plane polarized radiation is split into its two circularly polarized components by passage through a modulator subjected to an alternating (50 kHz) electric field. The modulator usually consists of a piezoelectric quartz crystal and a thin plate of isotropic material (e.g. quartz) tightly coupled to the crystal. The alternating electric field induced structural changes in the quartz crystal which make the plate transmit circularly polarized light at the extremes of the field. If, after the passage through the sample, the left and right circularly polarized components are not absorbed (or are absorbed to the same extent), combination of the components would regenerate radiation polarized in the original plane. However, if one of the components is absorbed by the sample to a greater extent than the other, the resultant radiation (combined components) would now be elliptically polarized, i.e. the resultant would trace out an ellipse. In practice, the CD instrument (spectropolarimeter) does not recombine the components, but detects the two components separately; it will then display the dichroism at a given wavelength of radiation expressed as either difference in absorbance of the two components ($\Delta A = A_L - A_R$) or as the ellipticity in degrees (θ) ($\theta = \tan^{-1}(b/a)$, where b and a are the minor and major axes of the resultant ellipse. There is a simple numerical relationship between ΔA and θ (in degrees), i.e. $\theta = 33 (A_L - A_R)$.

A CD spectrum is obtained when the dichroism, i.e. the variation of θ expressed in mdeg, is measured as a function of wavelength. Once that the spectrum has been acquired, the measurements must be normalized in order to make them independent from the protein concentration and the cuvette pathlength. This step is reached expressing the ellipticity as mean residue ellipticity, $[\theta]_{MRW}$, calculated dividing the ellipticity for the molar concentration (using as molecular weight the mean molecular weight for residue) and for the pathlength: $[\theta] = \theta \cdot MRW / 10 \cdot c \cdot d$, where θ is expressed in mdeg, the protein concentration c in mg/ml, the pathlength d in cm and MRW is the mean residue molecular weight, obtained dividing the protein molecular weight for the number of its amino acid residues. The value $[\theta]$ is expressed in deg $\cdot$ cm^2 $\cdot$ dmol^{-1}.

Every chiral chromophore or belonging to a chiral molecule presents a characteristic activity of circular dichroism, i.e. it is active at certain wavelength values. As far as proteins are concerned, interesting information can be obtained evaluating the region between 250 and 180 nm (far-UV CD). In this region the chromophore responsible for the dichroic signal is the amidic bond that presents a different ellipticity pattern depending on the conformation of the adjacent bond angles , i.e. depending on the kind of secondary structure in which it is embedded (alpha-helix, beta-sheet or random coil). Each kind of secondary structure possesses a typical spectrum, characterized by specific signals. Alpha-helix is characterized by two intense negative bands at 222 and 208 nm and one positive band at 192 nm. Beta-sheet structure presents a weak negative band at 218 nm and a positive band at 198 nm, while random coil structure shows a weak positive band at 218 nm and an intense negative one at ~ 198 nm (Fig. III.1).

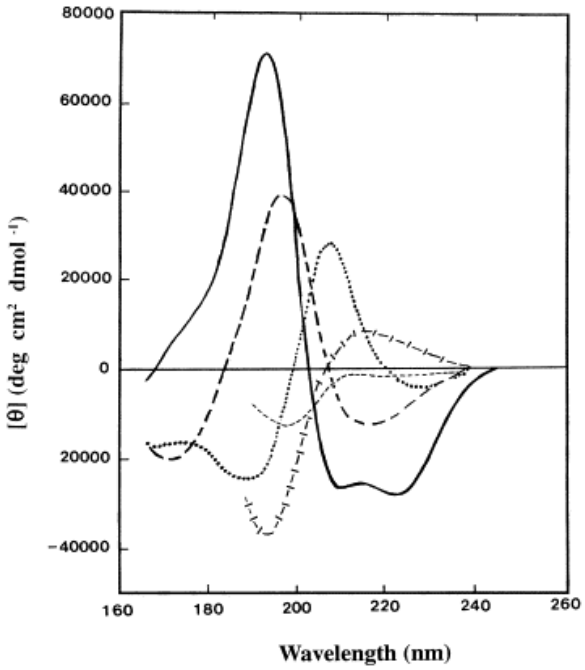


Fig. III.1. Far UV CD spectra associated with various types of secondary structure. Solid line, α-helix; long dashed line, anti-parallel β-sheet; dotted line, type I β-turn; cross dashed line, extended 31-helix or poly (Pro) II helix; short dashed line, irregular structure (reprinted from Kelly et al., 2005).

III.ii. Fluorescence

Fluorescence emission is observed when an excited electron returns from the first excited state back to the ground state. As some energy is always lost by non-radiative processes, such as vibrational transitions, the energy of the emitted light is always less than that of the absorbed light. Hence the fluorescence emission is shifted to longer wavelengths compared with the absorption of the respective chromophore. Fluorescence spectroscopy is one of the most powerful methods to study protein folding, dynamics, assembly, and interactions, as well as membrane structure. It has been successfully applied to investigate the complex mechanisms of protein aggregation including amyloid fibril formation. Almost all proteins have natural fluorophores, tyrosine and tryptophan residues, which allow study of changes in protein conformation. Also site-specific labeling with external fluorophores is easily achievable by mutagenesis and chemical modifications. Fluorescence spectroscopy requires a small amount of material (pM–nM range) and has a high signal-to-noise ratio.

Thioflavin T assay

ThT is a fluorescent dye used as a non-covalent extrinsic fluorescent probe in studies of amyloid fibril formation (LeVine, 1993). It has a two-ring structure: the conjugated benzothiazol and aminobenzol rings are arranged in an almost planar orientation (ϕ 30°) in the minimum energy conformation (Fig. III.2) (Munishkina & Fink, 2007). During excitation, the rings rotate in order to obtain the most stable exited state conformation (ϕ 90°) which has low fluorescence efficiency. ThT may bind between the betasheets of the fibril; however, no experimental data exist to support this assumption. ThT has two excitation (~335 nm and 430 nm) and two corresponding emission peaks (425–455 nm and 483 nm). Excitation at ~350 nm results in emission at ~438 nm, whereas excitation at 440 nm leads to emission at ~483 nm. The binding to amyloid fibrils induces a blue shift of the ThT λmax from 483 nm to ~478 nm and probably stabilizes the planar form of the molecule and leads to a 10–500-fold increase in ThT fluorescence intensity.

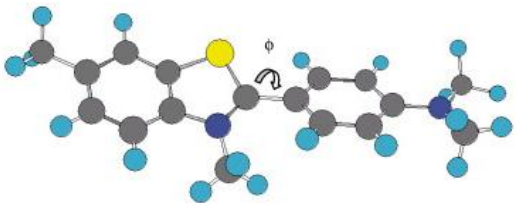


Fig. III.2. Ball-and-stick model of ThT: the rotational angle is around 30°. Carbons are dark gray, hydrogens are cyan, nitrogens are blue, and sulfur is yellow. (Reprinted from Munishkina & Fink, 2007).

III.iii. Fourier Transform Infrared Spectroscopy (FT-IR)

Infrared spectroscopy is a valuable tool for the investigation of protein structure, of the molecular mechanism of protein reactions and of protein folding, unfolding and misfolding. It works by shining infrared radiation on a sample and seeing which wavelengths of radiation in the infrared region of the spectrum are absorbed by the sample. For any given transition between two states, the light energy (determined by the wavelength) must exactly equal the difference in the energy between the two states, usually ground state (E_0) and the first excited state (E_1): $\Delta E = E_2 - E_1 = h f = h c v$, where E_1 is the initial Energy, E_2 is the final energy, h is the Planks constant, f is the vibrational frequency (sec^{-1}), c is the speed of light and v is the wavelength (cm^{-1}). The energy corresponding to these transitions between molecular vibrational states is generally 1-10 kilocalories/mole which corresponds to the infrared portion of the electromagnetic spectrum. Since vibrational frequency and probability of absorption depend on the strength and polarity of the vibrating bonds, they are influenced by intra- and intermolecular effects. The strength of absorption increases with increasing polarity of the vibrating bonds. Characteristic bands found in the infrared spectra of proteins and polypeptides include the Amide I and Amide II. The Amide I band is due to carbonyl stretching vibrations (C=O bond) while the Amide II is due primarily to N—H bending vibrations. Because both the C=O and the N—H bonds are involved in the hydrogen bonding that takes place between the different elements of secondary structure, the locations of both the Amide I and Amide II bands are sensitive to the secondary structure content of a protein.

When a protein is dissolved in water the hydrogens attached to the amide nitrogen can readily exchange for hydrogens attached to water molecules. For unfolded polypeptides this exchange can occur thousands of times a second. When a polypeptide folds to form the well-defined 3dimensional structure of a protein, many of the amide hydrogens become buried in the interior of the protein and no longer come into direct contact with the solvent water. This causes a marked decrease in the rate of exchange for these hydrogens. One drawback of infrared spectroscopy in aqueous solutions is the strong absorbance of water in the mid-infrared spectral region (near 1645 cm^{-1}) which overlaps the important amide I band of proteins and some side chain bands. The strong water absorption demands a short path length for aqueous samples, which is typically around 5 µm, and in turn relatively high concentrations. Using D_2O, the pathlength can be increased to 50 µm and the concentration lowered because the water band is downshifted to 1210 cm^{-1}. As a matter of fact, on going from C-H to C-D, the mass increases by a factor of approximately 2, thus the vibrational frequency for a C-D bond is smaller than that of the corresponding C-H bond. The stretching vibration of water, instead, appears near 3400 cm^{-1}, so the spectral region to analyze proteins in aqueous solution is between 3000 and 1000 cm^{-1}. The peptide bound gives rise to 9 characteristic bands, named amide A, B, I, II, III, IV, V, VI e VII band. The aminoacid side chain absorbs

in the spectral region between 1800 and 1400 cm^{-1}. The absorption of a side chain in a protein may deviate significantly from the environment, which modulates strength and polarity of bonds. The strongest bands are those of Asp, Asn, Glu, Gln, Lys, Arg, Tyr, Phe and Hys.

The secondary structure absorbs predominantly in a specific range of the amide I region, as specified in the following table:

Secondary Structure	Band position in D_2O / cm^{-1}	
	Average	Extremes
α-helix	1652	1642-1660
β-sheet	1630	1615- 1638
	1679	1672-1694
Turns	1671	1653-1691
Disordered	1645	1639-1654

III.iv. Mass Spectrometry

Thanks to recent technological progresses, mass spectrometry has become one of the most powerful tool in protein analysis (Aebersold & Mann, 2003). A further incentive to the development of such techniques has come also from the huge amount of data deriving from the sequencing of entire genomes. Indeed, mass spectrometry (MS), allowing the identification of thousands of proteins starting from complex mixtures, has given significant insights into biological and medical matters. MS can be exploited in many fields, ranging from chemistry and biology to pharmaceutical analysis. It can be used to measure the accurate molecular weight of peptides or proteins, both chemically synthesized or produced by recombinant methods. Other applications are the evaluation of the purity of a protein sample, the identification of post-translational modifications, the monitoring of chemical/enzymatic reactions, the sequencing of proteins or oligonucleotides. Lately, MS has been used also to study the folding and unfolding processes of proteins, or the formation of sovramolecular complexes and the determination of macromolecular structures.

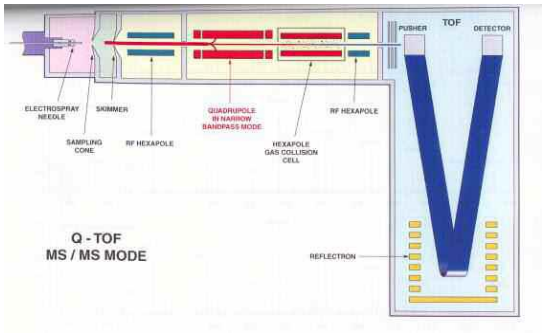


Fig.III.3 A schematic representation of an ESI Q-TOF instrument.

Analysis of peptides and proteins using MS has been allowed by the development of ionization techniques not inducing protein fragmentation. The instrument used in the work presented in this Thesis is an ESI-Q-TOF (*Electrospray-Quadrupole-Time of Flight*) from Micromass (Manchester, UK), an instrument (see the schematic below) having an electrospray sorgent and an analyzer system composed of two quadrupoles and a time of flight, in series. Between the two analyzers there is a collision chamber, used, when needed, for ions fragmentation.

In a generic MS experiment, a peptide/protein sample, dissolved in a solvent mixture containing H_2O/Acetonitrile/formic acid, is injected in the electrospray (ESI) ion source through a very small, charged and usually metal capillary. In electrospray ionization, a liquid is pushed though the capillary, with the analyte, dissolved in a large amount of solvent, which is usually much more volatile than the analyte. Volatile acids, bases or buffers are often added to this solution too. The analyte exists as an ion in solution either in its anion or cation form. Because like charges repel, the liquid pushes itself out of the capillary and forms an aerosol, a mist of small droplets about 10 μm across. An uncharged carrier gas such as nitrogen is sometimes used to help to nebulize the liquid and to help evaporate the neutral solvent in the droplets. As the solvent evaporates, the analyte molecules are forced closer together, repel each other and break up the droplets. This process is called Coulombic fission because it is driven by repulsive Coulombic forces between charged molecules. When the analyte is an ion free of solvent, it moves to the mass analyzer.

In the following phase the ion produced is analyzed, in the instrument, by a Q-Tof Micro (Micromass, Manchester), that connects an ESI source with two combined analyzers quadrupole Tof (*time of flight*) (Fig. III.4). Mass analyzers separate ions according to their mass-to-charge ratio, following the dynamic properties of charged particles in electric and magnetic fields in vacuum. The quadrupole mass analyzer uses oscillating electrical fields to selectively stabilize or destabilize ions passing through a radio frequency (RF) quadrupole field, acting as a mass selective filter. The time-of-flight (Tof) analyzer uses an electric field to accelerate the ions through the same potential, and then measures the time they take to reach the detector. If the particles all have the same charge, then their kinetic energies will be identical, and their speed will depend only on their masses. Lighter ions will reach the detector first.

The data produced are represented in a mass chromatogram of total ion current (TIC), measured in the ion source. The instrument acquires a mass spectrum of the injected peptide/protein, based on an intensity vs. *m/z* (mass-to-charge ratio) plot. Afterwards, the instrument can determine the sequence of the peptides though a Tandem mass spectrometry analysis.

III.v. Transmission Electron Microscopy

Negative staining has been a useful specimen preparation technique for biological and medical electron microscopists for almost 50 years, following its introduction as an established procedure by Robert (Bob) Horne (Brenner et al., 1959).

Preparation of carbon supports

A solution of 0.5% w/v butvar was prepared in a 9% v/v glycerol–91% v/v chloroform emulsion. This emulsion is stable at room temperature and needs only to be vigorously shaken before use. Clean glass microscope slides were inserted into the glycerol–chloroform emulsion to approx. two thirds of their length and withdrawn vertically. Excess emulsion was removed by touching the end onto a tissue paper and one side of the slide/mica was wiped to remove the surface film. The slide was then positioned horizontally and the fluid film on the upper glass surface allowed to dry. The small glycerol–water droplets penetrate the drying film of butvar, thereby creating small holes. After wiping the edges of the slide with a tissue, the perforated butvar film was floated onto a clean water surface. An evenly opaque appearance indicates the presence of a suitable array of small holes. EM 400 mesh copper grids (shiny side up) were then placed on the floating film, and a piece of white paper over-layered without moving the grids. The paper slowly became completely wet and was then removed with forceps, along with the attached grids and perforated butvar film, and dried in a dust-free environment. The sheets of paper+grids and perforated butvar film were then carbon-coated. Before use, the butvar was dissolved by spraying the grids with chloroform and grids were briefly glow-discharge treated (20 s) immediately before use. The hole size ranges from *ca*.1 to 10 µm.

Preparation of negatively stained specimens on holey carbon supports

To prepare the specimens using carbon support films we used the single-droplet Parafilm procedure. Specimens were individually negatively stained with 1 % w/v uranyl acetate. Briefly, 25 µl sample droplets and 20 µl droplets of negative stain were placed in rows on a clean Parafilm surface. The sample was applied to the holey carbon support film by touching a grid to the droplet surface and most of the fluid removed by touching to the edge of a filter paper wedge. Then, depending upon the salt concentration of the sample solution, the negative stain solution was applied and removed as a single or multiple droplets (e.g. ×2 or ×3), with the intention of sequentially washing away the buffer and other salts or solutes that may interfere with the production of an amorphous stain-sugar film. At the final stage, maximal removal of negative stain was necessary, so the filter paper was held in contact with the grid edge for ca. 10–20 s, thereby leaving only a thin film of aqueous negative stain+sample. If this final precaution

is not adhered to, the stain film will often tend to be too thick. Specimen grids, still held by fine forceps, were then positioned horizontally and the sample+stain allowed to air-dry at room temperature (22°C).

III.vi. References

Aebersold, R. and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature,* **422**, 198-207.

Brenner, S., Horne, R. W. (1959). A negative staining method for high resolution electron microscopy of viruses. *Biochim. Biophys. Acta*, **34**, 103-110.

Creighton, T. E. (1997). Protein Structure: A Practical Approach (Practical Approach Series).

Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., Zambonin, M. (1999). Limited proteolysis in the study of protein conformation. In Proteolytic Enzymes: Tools and Targets. Sterchi EE, Stocker W, eds, 257-284. Springer Verlag, Heidelberg.

Fontana, A., Polverino de Laureto, P., Spolaore, B., Frare, E., Picotti, P., Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim. Pol.*, **51**, 299-321.

Hubbard, S. J. (1998). The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta*, **1382**, 191-206.

Kelly, S. M., Jess, T. J., Price, N. C. (2005). How to study proteins by circular dichroism. *Biochim. Biophys. Acta*, **1751**, 119-139.

LeVine, H. (1993). Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci*., **2**, 404-410.

Munishkina, L. A. and Fink, A.L. (2007). Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochim. Biophys. Acta*, **1768**, 1862–1885.

Woody, R. W. (1995). Circular dichroism. *Methods Enzymol.,* **246**, 34-71.

Imperial College London

Department of Life Sciences
Biopolymer Mass Spectrometry Laboratory

N-Glycan analysis of *Mgat5* knock-out mice kidney and study of the effect of GlcNAc food administration in kidney and spleen tissues

&

Pilot study on Mixed Link Glucan emicellulose biosynthesis in barley coleoptile tissue

Supervisors:
Dr. Stuart M. Haslam
Dr. Maria Panico
Prof. Anne Dell

PhD student:
Micaela Pivato

May-October 2011

FOREWORD

During the third year of my PhD I spent six months at the Biopolymer Mass Spectrometry Laboratory of Imperial College in London, being involved in two projects: a glycomic analysis of mice tissues and a pilot study on expression and biosynthesis of mixed linked glucans emicellulose. This research experience gave me the possibility to learn and apply advanced techniques in mass spectrometry analysis of small organic compound, using GC-MS and MALDI-TOF spectrometers.

In this report I'm focusing on the glycomic project, giving a detailed description of experimental methods and results, and I enclose a synthetic report about the experiments conducted on the pilot study.

I would like to thank Dr. Anne Dell for the precious opportunity that she gave me to work in her lab, Dr. Maria Panico for being my external supervisor and Dr. Stuart Haslam for the support and guidance through the project. Moreover, special thanks to Dr. Paola Grassi and Dr. Poh Choo Pang for the technical teaching, assistance and friendship.

N-Glycan analysis of *Mgat5* knock-out mice kidney and study of the effect of GlcNAc food administration in kidney and spleen tissues

Pilot study on Mixed Linked Glucan emicellulose biosynthesis in barley coleoptile tissue

ABBREVIATION

CAD	Collision Activated Decomposition
CFG	Consortium of Functional Genomics
EI	Electron Impact
ES	Electrospray ionization
GC-MS	Gas Chromatography-Mass Spectrometry
GlcNacT-V	N-acetylglucosaminyltransferase V
ko	knock out
MALDI	Matrix assisted Laser Desorption Ionization
MS	Mass spectrometry
TOF	Time of Flight
wt	wild type

Amino Acids

Asn	Asparagine
Pro	Proline
Ser	Serine
Thr	Threonine

Sugars

Fuc	Fucose
Gal	Galactose
GalNAc	N- acetylgalactosamine
Hex	Hexose
Man	Mannose
NeuAc	N-acetylneuraminic acid
NeuGc	N-glycolylneuraminic acid

SUMMARY

N-acetylglucosaminyltransferase V (GlcNacT-V), encoded by the *Mgat5* gene, is a medial Golgi enzyme which catalyzes the addiction of a β-1,6-linked GlcNAc to the α-1,6 mannose of the trimannosyl N-glycan core. GlcNacT-V plays a pivotal role in the formation of tri- and tetra-antennary N-glycans on newly synthesized glycoprotein. This branch provides the preferred substrate for the enzymatic subsequent synthesis of polylactosamine chains and terminal modification including the Lewis antigens.

In this study, glycomic analyses were performed to investigate possible changes in protein N-glycosylation in wild type conditions and in the absence of *Mgat5* gene in C57B5 mice kidneys.

In parallel, N-glycan profile of kidneys and spleens coming from mice treated with high fat diet GlcNAc supplementation were analyzed. Previous results demonstrate that the effects of GlcNAc salvage appear to increase flux to UDP-GlcNAc. Therefore we were interested to know whether this implementation affects N-glycan branching.

Results show that *Mgat5* deficient mouse kidney display less amount of tri-antennary and tetra-antennary structure compared to controls. However, GlcNAc dietary salvage has no apparent effect on N-linked glycosylation in the kidney and spleen, even if the experiments conducted on cell lines demonstrate that increased influx of UDP-GlcNAc resulted on increased N-glycan branching. Moreover, the performance of optimized glycome procedure allowed the identification of more tri-antennary glycan structures than the one reported on CFG database.

1. INTRODUCTION

1.1 Protein Glycosylation

Among the co- and post-translational protein modifications, glycosylation is the most common and highly diverse (Apweiler et al, 2009). Glycosylation is the enzymatic process in which glycans are attached to proteins, or to lipids, and it is reported both in prokaryotic and eukaryotic systems (Taylor and Drickamer, 2010). The importance of glycosylation is indisputable, and can be highlighted by the fact that about two third of known human protein include the N-glycosylation consensus sequence, and that complete deficit in N-glycosylation is expected to be incompatible with life (Apweiler et al, 2009; Dennis et al., 2001). The sugar moieties strongly affect protein structure and function, and exert their actions both intracellularly and extracellularly. Glycans are involved in crucial steps of tissue development, cell-cell communication, immunity, but also protein folding, localization, activity and half-life. Heterogeneity is guarantee by the presence of different monosaccharides that can combine in a variety of ways that differ per sequence, chain length, anomery (α and β), position of linkages and branching points. Despite this diversity, infinite combinations are not possible, since the glycosylation process is strictly regulated by glycosyltransferases, glycosidases and carbohydrate-modifying enzymes. Moreover, based on the different specificity and expression pattern of these classes of enzymes, glycosylation results a species- and cell-specific process (Lis and Sharon, 1993; Coutinho et al. 2003).

Protein glycosylation can be divided into two classes, *N*-linked and *O*-linked, depending on whether glycans are attached to amide nitrogen of Asn or to oxygen of hydroxyls groups of Ser and Thr, respectively.

1.2 N-Glycosylation

Humans N-glycans contain a common core structure ($Man_3GlcNAC_2$) and differ in the terminal elaborations that extend from this core. Sugar composition and structure of N-glycans vary during the several steps of the glycosylation process, and depending on composition, sugars are classified in high-mannose, hybrid and complex glycans. Each of these three types of N-glycan can be present in the final mature glycoprotein (fig 1.1).

N-glycan biosynthesis begins in the endoplasmic reticulum (ER) with the formation of a lipid-linked precursor oligosaccharide, which is then transferred *en bloc* to a nascent polypeptide during its translation. A consensus sequence, Asn-X-Ser/Thr (where X is any amino acid except Pro) allows the linkage. The oligosaccharide is further processed by specific glucosydases and glycostltransferases in the ER and the Golgi apparatus. A series of mannosidases remove some or all the mannose residues that constitute the high mannose oligosaccharides and complex glycans are built on the core structure. Bi-, tri- or tetra-antennary glycans are built in a cell-and tissue-specific way, five or more branches have been documented, but not in human (Taylor and Drickamer, 2010). The expression of complex glycans is also regulated temporally during

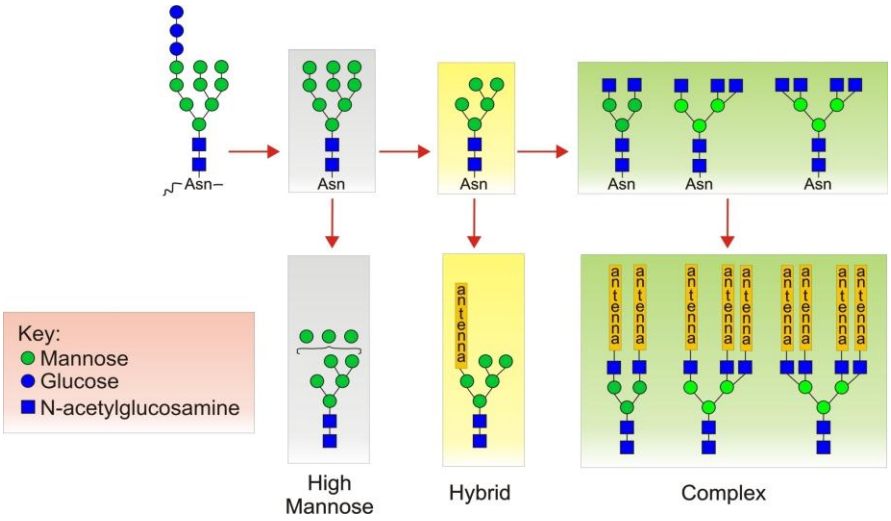


Fig. 1.1 Processing of initial high mannose N-linked glycan to generate complex glycan (kindly provided by GlycoTRIC, The Glycobiology Training course).

development, especially of the immune system, and in certain disease situation, including some cancers (Dennis et al., 1999a).

1.3 Mgat5: N-acetylglucosaminyltransferase V

N-acetylglucosaminyltransferase V (GlcNacT-V), encoded by the *Mgat5* gene, is a medial Golgi enzyme which catalyzes the addiction of a β-1,6-linked GlcNAc, from UDP-GlcNAc, to the α-1,6 mannose of the trimannosyl N-glycan core (Morgan et al., 2004). GlcNacT-V plays a pivotal role in the formation of tri- (2,2,6) and tetra- (2,4,2,6) antennary N-glycans on newly synthesized glycoprotein, and this branch provides the preferred substrate for the enzymatic subsequent synthesis of poly-N-acetyllactosamine chains (i.e. Galβ-1,4-GlcNAcβ-1,3 repeating units of 2 to 10 in length) and other terminal modification including the Lewis antigens (Dennis and Laferte, 1989; Alverez et al. 2002). Immunohistochemistry detection of Mgat5 is mainly performed indirectly, by staining with leukoagglutinin (L-PHA), which selectively binds GlcNacT-V products (i.e. the β-1,6-linked lactosamine antenna). Actually, the control of polylactosamine and Lewis antigens expression emerges mainly as a consequence of substrate preference, enzyme compensation and tissue-specific regulation of glycosyltransferase and GlcNacT-V appears to be a rate limiting enzyme for these processes (Dennis et al. 1999b).

GlcNacT-V is expressed in several tissues at significantly different levels. It is expressed in different stages of mouse embryogenesis, in neuroepithelium of CNS, vascular endothelium, and specialized basal epithelia of gastrointestinal tract, kidney, endocrine system, skin and respiratory tracts. In the adult mice it is detected in follicular and testicular cells (Granovsky et al., 1995). Moreover, *Mgat5* gene transcription was reported to be induced by cell viral transformation, and by oncogene transfection with RAS or v-FPS. Elevated expression of GlcNacT-V has been also related with a number of different tumors including breast, colon and melanomas. Staining with L-PHA also revealed that the presence of GlcNacT-V is proportional to the tumor aggressiveness, and required for metastasis processes. Deficient *Mgat5* mice model confirms these observations, displaying a suppression of tumor growth and metastasis. Furthermore,

they demonstrate that GlcNacT-V glycan products stimulate membrane ruffling and phosphatidylinositol 3 kinase-protein kinase B activation, fueling a positive feedback loop that amplify oncogene signalling and tumor growth *in vivo* (Granovsky et al. 2000).

Common feature of embryonic basal epithelia and neoplastic cells is the capacity to migrate, a cellular function which requires GlcNacT-V-dependent glycoconjugates. β-1,6 GlcNAc branched N-glycans enhance focal adhesion turnover and cell motility, with positive feedback amplifying intracellular signalling pathway. In particular, glycosylated integrins are substratum of adhesion receptor, and this binding is mediated by branched N-glycans (Dennis et al. 1999b).

N-acetyllactosamine chains are also ligand for galectins, proteins involved in the modulation of T-cell proliferation and apoptosis. Knock out *Mgat5* mice display kidney autoimmune disease, due to and enhanced T-cell receptor (TCR) clustering. Indeed, galectin-3 is associated with the TCR complex at the cell surface, and the interaction depends on the expression of *Mgat5* gene: GlcNacT-V product glycoprotein interaction with galectins limits TCR recruitment to the site of antigen presentation (Demetriou et al., 2001).

1.4 Mass Spectrometry determination of glycans

First mass spectrometry (MS) analyses on glycoprotein were performed about 30 years ago. In the last decades, significant improvements in glycan characterization came with technological advantages in MS. However, several key aspects were not outsmarted. The derivatization of chemical moieties is still a preferred strategy that improves result throughput and sensitivity, and in particular permethylation survives as one of the favourite techniques. Actually soft ionization MS approaches, such as electrospray ionization (ES) or matrix laser desorption ionization (MALDI), are commonly used for the identification of glycans and glycocopeptides, while electron impact (EI) is used in association with gas chromatography for the definition of linkage and configuration of sugars (Dell and Morris, 2001).

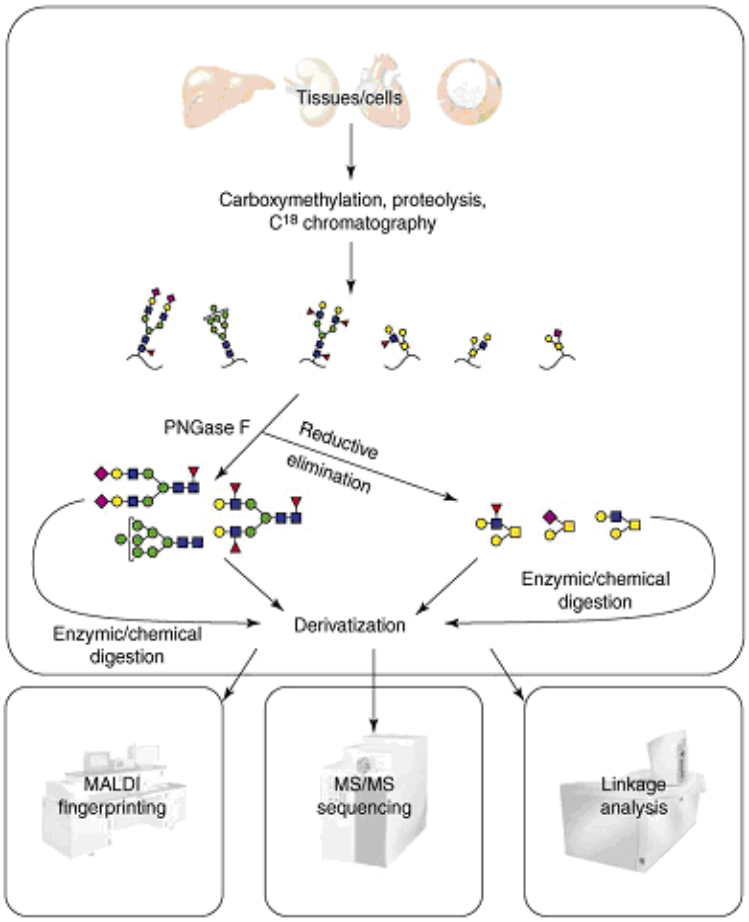


Fig. 1.2 Schematic workflow of a glycomic experiment. The sample is digested and purified before the release of glycan pools,that can be prepared for analysis directly, or subjected to a series of further enzymatic and chemical digenstions. The native or processed glycan pools are then derivatised before analysis. The sample can then be analysed using a range of MS techniques, including MALDI fingerprinting; ES-MS/MS or MALDI-MS/MS, and GC-MS (reprinted from Haslam et al.,2006).

Figure 1.2 summarizes the typical steps of a glycomic experiment, which comprises: sample handling before MS analysis, MALDI mass fingerprinting, ES-MS/MS or MALDI-MS/MS sequencing and supplementary experiments, such as linkage analysis (Haslam et al., 2006). In brief, biological samples, tissues or cells, are homogenized and proteolytically digested, and the glycome pool is subsequently released by enzymatic (i.e. N-glycisidase F hydrolyze the link between N-glycans and Asn residues) or chemical (i.e. alkaline β-elimination for O-glycans) methods. These glycan pools, which comprise protein N- and O-glycan, and polar and non-polar glycolypid-derived glycans, are then derivatized by permethylation to optimise the separation and detection of the constituent glycans prior to MS analysis. MALDI ionization time-of-flight spectrometry (MALDI-TOF) in generally used for profiling glycans pools (sample fingerprinting), wherein the overall carbohydrate content of the sample and relative abundance of specific moieties can be observed. MS/MS sequencing experiments, conducted either with ES or MALDI ionization, are performed to confirm the structure of each individual glycan, while ES-LC-MS/MS is preferred for the identification of peptide glycosylation sites (glycoproteomic). Supplementary experiments, such as GC-MS linkage analysis,

which requires further sample handling, provide specific data on individual linkages within the glycan structure (North et al., 2010).

1.5 Consortium for Functional Glycomic

Consortium for Functional Glycomics (CFG) is a large international research initiative, founded in 2001, that coordinates international glycobiology research. The goal of CFG is to define the paradigms by which protein-carbohydrate interaction mediate cell communication. It provides a networking forum, where participant research groups contributes to reveal functions of glycans and glycan-binding proteins (GBPs) that impact human health and disease. The CFG offers also offer access to data analysis tools and large glycomics databases, which provide detailed information about glycan structures, GBPs, and glycosyltransferases. The extensive data sets that have resulted from the CFG founded researches are public through the website (http://www.funtionalglycomics.org), and include results from glycan array screening, glycogene microarray screening, mouse phenotyping, and glycan profiling experiments.

1.6 Aim of the study

In this study, mass spectrometry glycomic analyses were performed to investigate possible changes in protein N-glycosylation in the absence of *Mgat5* gene in C57B5 mice kidneys. Wild type (wt) C57B5 mice kidneys were, to this aim, analyzed in parallel.

We were also interested into the N-glycan characterization of mice receiving a high fat diet GlcNAc supplementation to their drinking water. Treated mice show differences in weight gain and liver pathology. Previous results demonstrate that the effects of GlcNAc salvage appear to increase flux to UDP-GlcNAc, directly affecting N-glycosilation. In particular, immunocytochemistry analyses on NMuMG cells and have shown that GlcNAc salvage results in increased N-glycan branching (Lau et al., 2007; Lau et al. 2008). Moreover, Mgat5 null mice were demonstrated to be resistant to weight gain on an enriched diet (Cheung et al, 2007). Therefore we were interested to know whether this affects N-glycan branching in kidney and spleen tissues.

In total, six spleens samples and four kidney samples were analyzed. Spleens consisted on two wt controls, and four wt spleens from mice treated with different amount of GlcNAc supplementation: one with 0.5 mg/ml, two with 5mg/ml and one with 15 mg/ml per day additional GlcNAc. Wt and knock out (ko) *Mgat5* kidneys, both from implemented and from control GlcNAc diet, were analyzed. All the mouse tissues were provided by the group of Prof. J. W. Dennis within the Consortium of Functional Glycomics.

2. EXPERIMENTAL PROCEDURES

Experimental procedures reported here were described in details by North and colleagues (North et al., 2010).

2.1 Tissue homogenization, glycolipid and glycoprotein extraction

Mouse tissues and organs analyzed by the CFG were harvested from 6- to 8-week old C57B5 mice. The excised samples were snap-frozen immediately and stored at -80° C .

The 4 kidneys (wt+, wt-, null+, null -, were +/- refers to GlcNAc supplementation on the diet added to mice drinking water) and 6 spleens (wt-, a spleen treated with 0.5 mg/ml, two with 5 mg/ml and one with 15 mg/ml per day GlcNAc addiction on the diet) were weighed, resuspended in 4 volumes of ice cold water and homogenized. The total water volume of the sample was calculated, assuming that 80% of the tissue weight is water, and according to this 2.67 volumes of methanol were added before vortexing at room temperature. Then, 1.33 volumes were added followed by vigorous mixing and centrifugation at 3000 rpm for 10 min. The supernatant containing glycosphingolipids (GSLs) was carefully separated from the protein-containing pellet.

Polar and non-polar GSLs were separated by extraction with water, i.e. 0.175 water volumes were added to the supernatant and mixed. Centrifugation for 15 minutes at 3000 rpm allowed the separation of two phase: polar glycolipid (upper layer) and non-polar glycolipid were collected and stored for potential further GSL analyses, that were not performed in this study.

Excess organic solvent was removed from the pellet under a nitrogen stream without allowing the samples to dry completely. 50 µL of 0.6 M Tris buffer was added to the samples which were put under a nitrogen stream but not dried completely. The protein fraction were subsequently reduced and carboxymethylated.

2.2 Reduction and carboxymethylation

The protein pellets from each homogenized tissue were reduced by incubation with 2 mg/ml dithiothreitol in 2 mL 0.6 M Tris-HCl buffer, pH 8.5 for 1 hour at 37° C. For carboxymethylation equal volume of 0.6 M Tris-HCl buffer were added containing 12 mg/ml of iodoacetic acid. The reaction was carried out in the dark at room temperature for 90 min and terminated by dialysis in 4.5 L of 50 mM ammonium bicarbonate buffer for 48 h at 4°C.

The buffer was replaced regularly and after completion of dialysis the samples were lyophilized.

2.4 Proteolytic digestion

The reduced and carboxymethylated proteins were digested using TPCK pre-treated bovine pancreas trypsin (Sigma). About 1.5 mg of trypsin in 50 mM ammonium bicarbonate buffer pH 8.4, were added to each sample for 16 h at 37 °C. The reaction was terminated by heating the sample to 100°C for 5 minutes.

Digestion products were purified by chromatography with Oasis HLP Plus cartridges (Waters). The columns were first conditioned with 5 ml of 5 % (v/v) acetic acid, 5 ml propanol and another 15 ml of 5% acetic acid. Subsequently samples were loaded and a washed with 20 ml of 5% acetic acid. Then peptides and glycopeptides were eluted with 20%, 40% and 100% (v/v) propanol in 5% acetic using 5 ml of each. The volumes of all the elutions were reduced in a SpeedVac before combining the 20% and 40% propanol fractions, followed by overnight lyophilisation.

2.5 Release of N-glycans from glycopeptides

PNGase F (Roche) digestion of the peptide fragments was performed in 250 µL of 50 mM ammonium bicarbonate in which 3U of enzyme were added. The reaction was left at 37°C for 24 h. The products were then lyophilized.

Separation of N-linked glycans from O-linked glycopeptides was performed by chromatography with C_{18} Sep-Pak short body classic cartridges (Waters). The column was conditioned sequentially with 5 ml of methanol, 5 % (v/v) acetic acid, propanol and 15 ml 5 % (v/v) acetic acid. The samples were dissolved in 200 µL of 5% acetic acid and loaded dropwise into the column. Glycans were collected with 5 ml 5% acetic acid and 4 ml each of 20%, 40% and 100% propanol in dilute acetic acid. The volume of the collected fractions was then reduced in a SpeedVac. The 20% and 40% propanol fractions were combined and all the samples were lyophilized.

2.6 Permethylation

Permethylation was performed using sodium hydroxide method. Sodium hydroxide pellets were crushed into fine powder before adding 2-3 ml anhydrous dimethyl sulphoxide. 0.5-1 ml of the mixture was pipetted into the dried samples. About 0.5 mL of methyl iodide was then added before subjecting the samples to vigorous shaking for 30 min at room temperature. The reaction was quenched by dropwise addition of water followed by shaking, until no more effervescence was observed. More water was then added to reach 5 mL follow by 2 mL of chloroform. After vigorous mixing, the mixture was centrifuged for about 30 sec for separation into two layers. The top aqueous layer was discarded and the whole washing step with addition of water was repeated four more times. The chloroform layer was then dried under a nitrogen stream.

The final step was purification by another chromatography with C_{18} Sep-Pak short body classic cartridges (Waters). The column was conditioned sequentially with 5 ml methanol, water, acetonitrile and 15 ml of water. The samples were dissolved in 200 µl of methanol/water and loaded dropwise. The column was washed with 5 ml of water

before elution with the 4 ml each of 15%, 35%, 50% and 75% acetonitrile (v/v) in water. The samples were lyophilized prior to MS analysis.

2.7 Mass spectrometry

The samples were dissolved in 10 µL of methanol from which 1 µL was mixed with 1 µL of matrix solution containing 20 mg/mL DHB (2,5-dihydroxibenzinoic acid, Sigma) in 70% methanol 30% water. The 2 µL mix resulting from each sample was spotted onto the MALDI plate and dried prior to analysis.

MALDI-TOF (Voyager, Applied Biosystem) analysis was performed for samples mass-mapping and thus obtaining N-glycan fringerprinting. MS spectra were manually analyzed prior to perform MSMS analyses to confirm the glycan structures.

Sequence information were then obtained by fragmentation with collisionally activated decomposition (CAD) of parental molecular ion carried out by tandem mass spectrometry equipment with MALDI-TOF/TOF (4800, Applied Biosystem). MS/MS allows the fragmentation of sample individual molecular ion, each of one corresponds to single or multiple glycan species defined by a precise molecular mass.

3. RESULTS

N-glycans were analyzed from wild type and GlcT-V *Mgat5* knock out (ko) mouse kidney tissues. Wild type and mutant mice kidney glycosylation profiles are reported in figure 3.1. Spectra were obtained from the 50% MeCN fraction of a C_{18} Sep-Pak column performed after permethylation (see section 2.6). All molecular ions are $[M+Na]^+$, since the derivatization method was carried out using sodium hydroxide procedure.

m/z values reported on the spectra refer to the monoisotopic peak m/z. The glycan structures have been predicted based on composition, involved biosynthetic pathways, and tandem mass spectrometry analyses. The cartoon structures were drawn according to the Consortium of Functional Glycomics guidelines. Structures that show sugars outside a bracket have not been unequivocally defined. For clarity ions with higher intensity are shown in figure 3.1, while figure 3.2 reports also less abundant ions, splitting the mass spectra of wt mouse kidney in three panels.

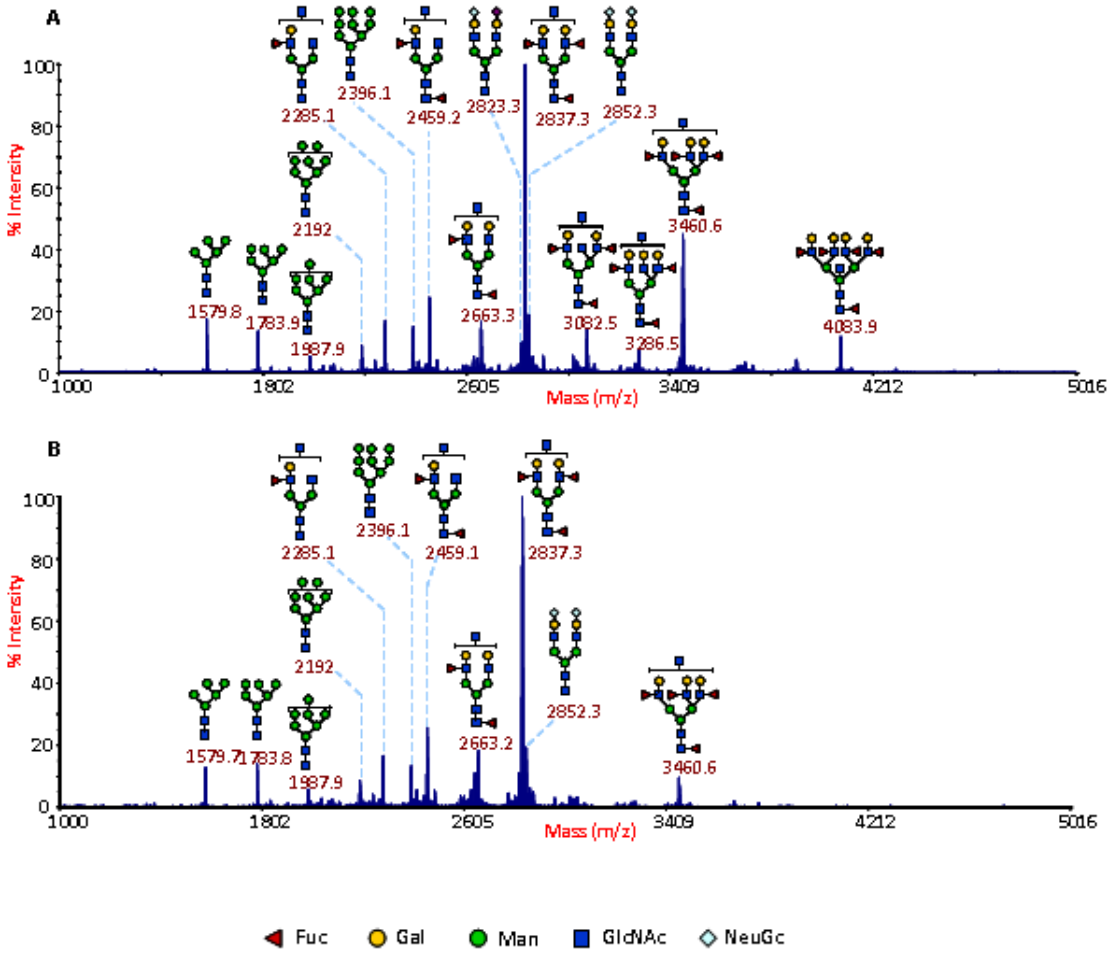


Fig. 3.1 MALDI-TOF spectra of wild type (A) and GlcT-V deficient (B) mouse kidney N-glycan. Major molecular ion glycan structures, predicted depending on literature, composition, and MS/MS fragmentation are shown.

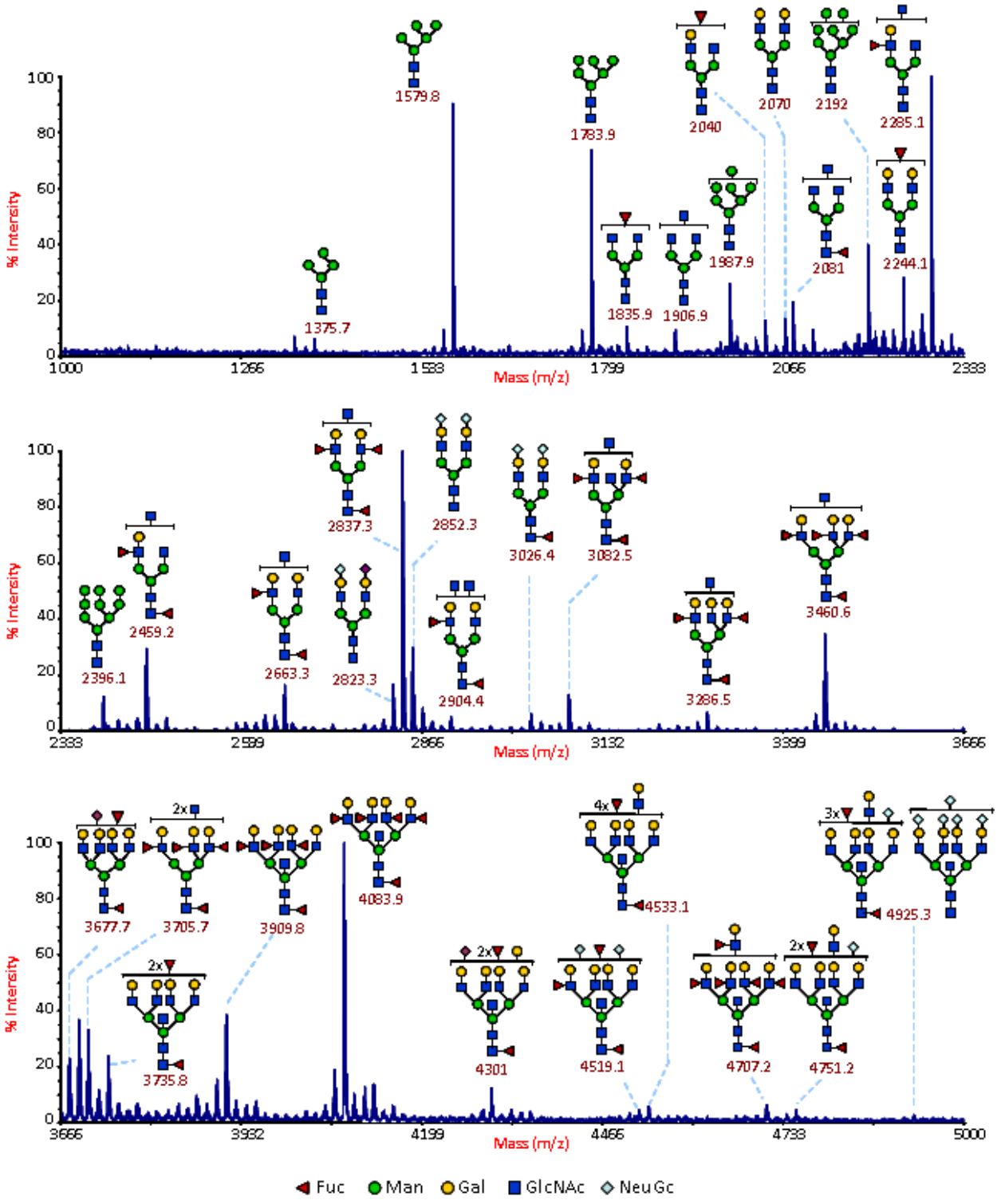


Fig. 3.2 MALDI-TOF spectra of wild type mouse kidney N-glycan. Low intensity molecular mass ion glycan structures, predicted depending on literature, composition, and MS/MS fragmentation are shown.

Both wt and *Mgat5* ko kidney tissues present molecular ions corresponding to high mannose glycans (1578.9, 1783.9, 1987.9, 2192 and 2396.1 m/z, with an error of 0.1 Da). Complex glycans are characterized by Lewis X terminal structure (Gal-Fuc-GlcNAc), and core fucosilation. The most abundant peak is 2837.3 m/z, which could correspond both to a bi-antennary structure, with terminal Lewis X antigens, fucorolated core and bisecting GlcNAc (non-elongated GlcNAc branch), or to a tri-antennary glycan, where the GlcNAc residue is not bisectin but constiture a new andenna. Moreover, N-glycans of wt mouse present tri- and tetra-antennary structures with both elongatied antennae with two Lewis X antigens (es. 4707.2 m/z) and antennae sialylated with NeuGc (i.e. 4925m/z) or NeuAc (i.e. 4301 m/z). N-glycans of *Mgat5* ko kidney present less abundance of tri-antennary structures (i.e. 3460.6 m/z) and no presence of tetra-antennary ones (i.e. 4083.9 m/z).

Kidney tissues of wt and ko mice fed with supplementary GlcNAc (daily added on mouse water supply) were processed in the same way and the spectra profiles are shown in supplementary material (fig. 6.1). N-glycan fingerprinting does not reveal any difference in composition and relative abundance of glycan structures of fat diet mice compared to the respective standard fed ones (fig 6.1, supplementary material).

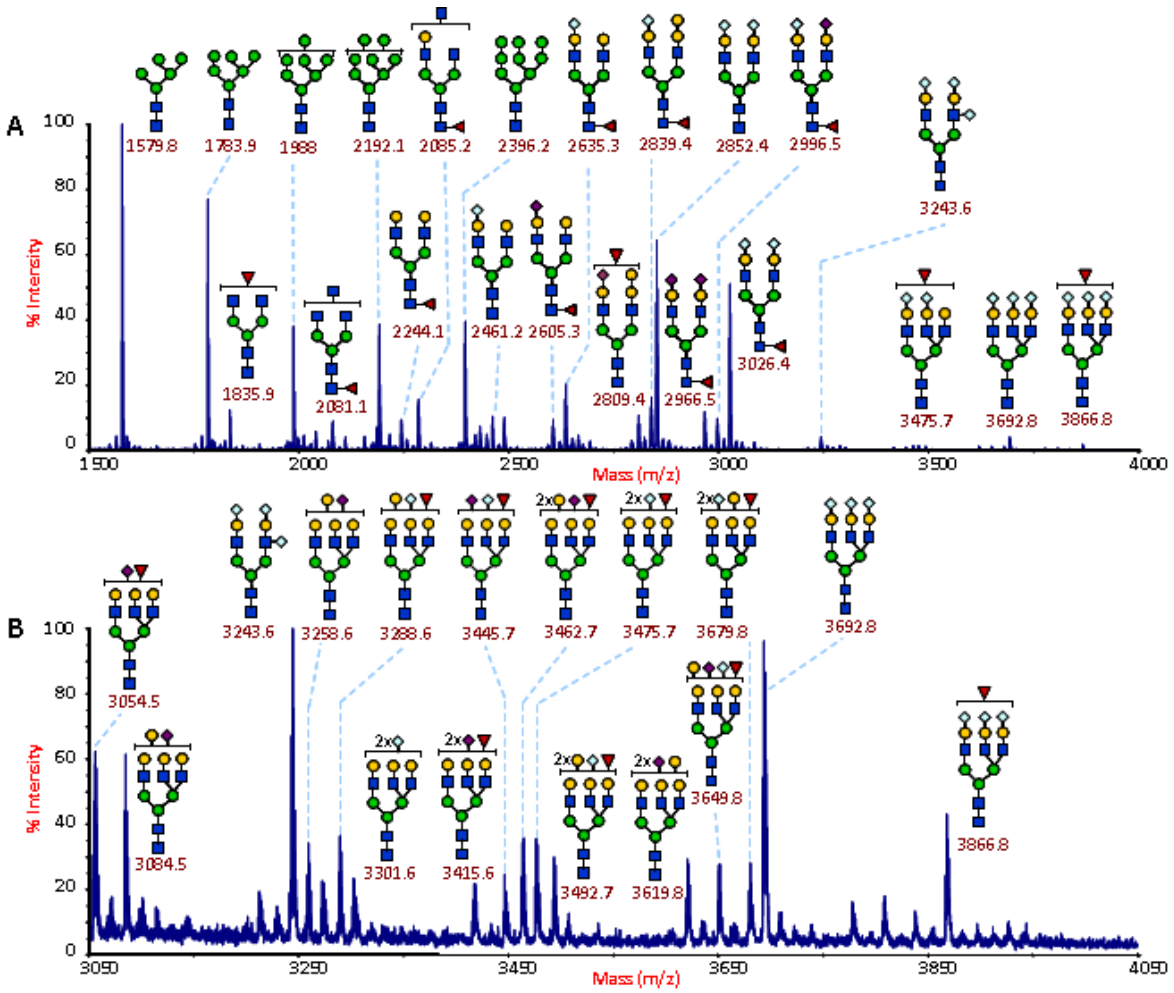


Fig. 3.3 MALDI-TOF spectra of wild type mouse spleen N-glycan (A). The mouse was fed with standard amount of GlcNAc. Glycan structures, predicted depending on literature, composition, and MS/MS fragmentation are also shown. Low intensity molecular mass ion assignment ranging between 3050 and 4050 m/z values are shown in panel B.

N-glycomes of several spleen organs coming from wt mice fed with different amount of additional GlcNAc were also analyzed. The following samples were analyzed: two spleen from mice fed with standard diet, one with 0.5 mg/ml per day supplement of GlcNAc, two with 5 mg/ml per day supplement of GlcNAc and one with 15 mg/ml per day supplement of GlcNAc on the water supply. MALDI-TOF spectrum of the glycomic profiles of a control mouse (standard alimentation) is reported in figure 3.1, while spectra derived from the other tissues analysed are shown in supplementary material (fig. 6.2). Again, spectra were obtained from the 50% MeCN fraction from a C_{18} Sep-Pak after permethylation and the glycan structures have been predicted based on composition, literature, and tandem mass spectrometry analyses.

No main differences were detected in the composition and relative abundance of glycan structures of fat diet mice compared to the standard fed mice. Difference in the relative intensity of 2635.3 m/z and 2839.4 m/z ions were detected within the control mouse (standard alimentation), and within high mannose glycans of all the analyzed tissues, probably due to sample handling. All the tissues analyzed display consistent amount of molecular ions corresponding to high mannose glycans (1578.8, 1783.9, 1988, 2192.1 and 2396.2 m/z, with an error of 0.1 Da). Complex glycans are bi- and tri-antennary, and terminal Galactose residues are mostly syalilated, both with NeuAc and with NeuGc (i.e. 2996.5 m/z), but predominantly with this last one, which is a sugar residue present in mammals with the exception of humans. Gal-α-Gal antenna decoration was also identified (2839.4 m/z, and outside the bracket of 3258.6, 3288.6, 3492.7, 3649.8, 3679.8 m/z), and also this decoration is not reported in humans. Core fucosilation is also present (i.e. 2635.3 m/z). The most abundant peaks are the high mannose glycans, and peaks 2635.3, 2852.4 and 3026.4 m/z, correspondent to bi-antennary structures with terminal syalilation and core fucosilation (2635.3 and 3026.4 m/z).

Tandem mass spectrometry

A consistent number of molecular ions from N-linked glycans were subjected to collision activated decomposition (CAD) for MS/MS analysis. The fragmentation of glycan molecular ion provided structural information used for the assignments reported in previous figures of this section. Tandem MS analysis can also give indication on the abundance of the different structures with the same m/z.

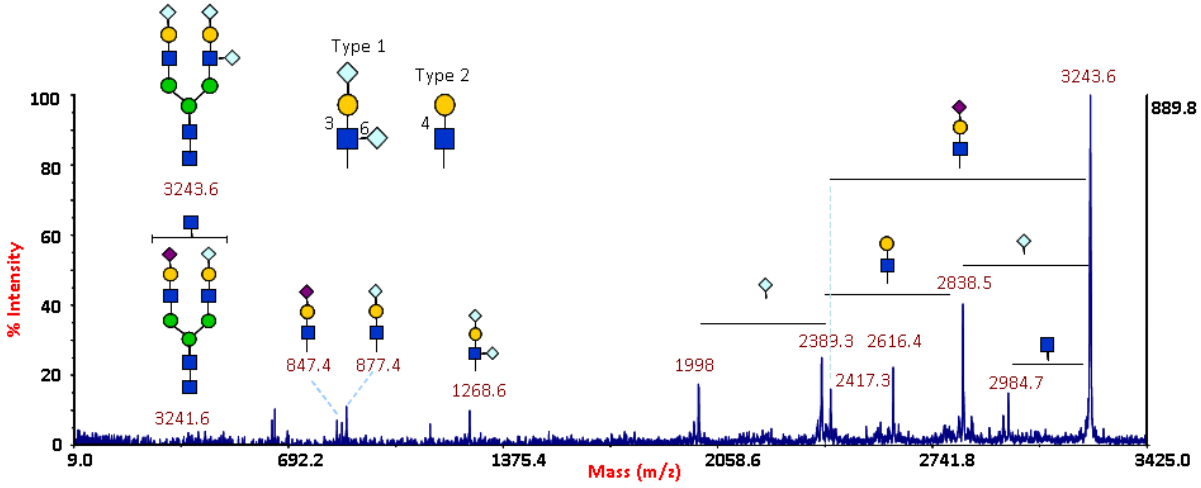


Fig.3.4 MALDI-TOF-TOF spectrum of permethylated N-glycans at 3246.6 m/z from wild type mouse spleen, fed with standard GlcNAc amount. The ion was selected for fragmentation, and the fragments ions corresponding to the two structures are indicated on the spectrum. Type 1 and Type 2 of Gal-GlcNAc linkage are also shown.

Fragmentation patterns were the same for organs of fat diet animals compared to the relative reference and for kidney of wt and ko mice, suggesting that the same structures are present in both types of animals. An example of data obtained from MS/MS and how the information can be interpreted is shown in figure 3.4, where the fragmentation pattern of the molecular ion at 3243.6 m/z present in spleen tissues is shown. As indicated in the previous spectra, this m/z value is referred to the monoisotopic peak (each molecular ion is represented by more peaks that differ for Δ = 1 Da, due to the varied isotopic composition), and the molecular ions are $[M+Na]^+$. The fragmentation pattern revealed the presence of a second glycan structure at m/z 3241.6, which is involved in the collision event since its isotopic composition overlaps m/z values of the 3243.6 m/z glycan. Based on the knowledge of the biosynthetic pathway and the composition of the molecular ion these two structures were identified: $NeuGc_1NeuAc_1Hex_5HexNAc_4$ and $NeuAc_3Hex_5HexNAc_4$ for 3241.6 and 3243.6 m/z glycan, respectively. The fragmentation pathway that commonly occurs during CAD is known and so the fragments ions can be predicted. The detected molecular ions correspond to fragments expected to be observed from both structures. The relative intensities of the fragmentend ion peak confirm that the $NeuAc_3Hex_5HexNAc_4$ glycan is prevalent. A GlcNAc residue was drawn outside the bracket of $NeuGc_1NeuAc_1Hex_5HexNAc_4$ glycan because, on the base of this MS/MS spectra, it is not possible to assign whether it is a bisecting GlcNAc or if it is linked to a core mannose residue, to initiate a tri-antennary structure. $NeuAc_3Hex_5HexNAc_4$ contains three sialyc acids residues, two of which are located on the same antenna. This structure is unusual, because it requires Type 1 Gal-GlNAc linkage (on carbon 3 of GlNAc residue), which is less common respect to Type 2 linkage (Gal residue linked on carbon 4 of GlcNAc) (Type 1 and 2 structures are shown in figure). This structure have been reported in mass liver tissues (CFG data). The same fragmentation profile was observed for all the spleen tissues analyzed (not shown). A similar approach was followed for interpretation of the other MS/MS data (not shown).

4. DISCUSSION

In this project a glycomic analysis was performed on wild type and *Mgat5* knock out mice kidney tissues. This was done in an effort to determine whether *Mgat5* lack has an effect on the overall N-glycosylation of cells, which would result in cell-cell communication deficit and would have consequence on signalling pathways. Moreover, kidney tissues were chosen for this characterization because *Mgat5* ko mice display kidney autoimmune disease (Demetriou et al., 2001).

Mgat5 gene code for a GlcNAc transferase, GlcNAcT-V, which catalyzes the addiction of a β-1,6-linked GlcNAc to the α-1,6 mannose of the trimannosyl N-glycan core, therefore it is responsible of the building of 2,2,6-tri-antennary structure and 2,4,2,6-tetra-antennary structures starting from 2,4,2-tri-antennary glycans. These branches provide the preferred substrate for the enzymatic subsequent synthesis of poly-N-acetyllactosamine chains and other terminal modification including the Lewis antigens (Dennis and Laferte, 1989; Alverez et al. 2002). Beyond the kidney autoimmune disease, *Mgat5* ko mouse display enhanced delayed-type hypersensivity, increased susceptibility to experimental autoimmune ecephalomyetitis and minor propensity to metastatic processes in tumors. Autoimmune disorders have been explained by the fact that GlcNAc glycosylation is a limiting step for the synthesis of glycoconjugates that interacts with galectins. Galectins, in turns, play a fundamental role in the recruitment of TCR on T-cell surface. The lack of interactions between galectins and glycoprotein binding partners, whose glycosylation is mediated by GlcNAcT-V, results in enhanced TCR clustering, that leads to T-cell activation (Demetriou et al., 2001). Reduction of tumor growth and metastatis process in *Mgat5* ko mouse, and the consistent presence of the enzyme in basal epithelia during development, are due to the fact that β-1,6-GlcNAc branched N-glycans enhance focal adhesion turnover and cell motility (Dennis et al. 1999b).

N-linked glycans of wt and ko mice kidneys were analyzed by following an established protocol for sample preparation and MS data acquisition (North et al., 2010). Glycan structures have been assigned based on composition, N-glycan biosynthetic pathways and tandem mass spectrometry. N-glycan fingerprinting revealed a minor presence of high molecular mass ions in ko tissue. Data obtained from wt tissue were consistent with previous N-glycan kidney characterization studies (Takamatsu et al., 2010; supplementary material). Ko mice show less presence of tri-antennary and traces of glycan structures with m/z values > 4000 which could correspond both to tetra-antennary structure and to elongated tri-antennary structures. Due to the small intensity, it was not possible to perform MS/MS analyses on these peaks. To determine whether tri-antennary structures detected in ko mice are 2,2,6 or 2,4,2, linkage analyses, which were not performed in this study, should be performed. Completely lack of β-1,6 GlcNAc transferase should result in the absence of 2,2,6 tri- and tetra-antennary structures. However, another β-1,6-GlcNAc transferase, GlcNAcT-Vb or GlcNAcT-IX, has been recently described, with preferential expression in testis and brain (Abbott et al., 2006). In order to determine if compensation events are taking place, or if GlcNAcT-V is

the only β-1,6-GlcNAc transferase present in kidneys, linkage analysis has to be performed to identify which tri-antennary structure (2,2,6 or 2,4,2) is present in ko sample.

Glycomic analyses were also performed on the same tissues, and on spleen tissues of wt mice, in order to characterize the effect of GlcNAc fat diet on N-glycosylation. Salvage of GlcNAc on the diet was shown to increase UDP-GlcNAc metabolite flux, directly affecting N-glycosilation. Analyses on cell lines have shown that GlcNAc salvage results in increased N-glycan branching (Lau et al. 2008). Moreover, Mgat5 null mice were demonstrated to be resistant to weight gain on an enriched diet, in a contrary behaviour respect to wt controls (Cheung et al, 2007).

N-glycosylation fingerprinting of analyzed tissues does not reveal remarkable differences on the glycome profile neither of wt and *Mgat5* ko kidney of fat diet treated mice compared to controls, nor of spleens from wt mice treated with different amount of supplementary diet (no, 0.5, 5 and 15 mg/ml per day GlcNAc addition). This result is in contrast with immunocytochemistry results previously described (Lau et al. 2008). Other tissues should be analyzed to determine whether UPD-GlcNAc increased influx affect global N-glycan branching. However, speen analysis revealed the presence more types od tri-antennary glycan structures (variously syalilated, fucosilated or containing Gal-α-Gal motif) than the one reported on CFG database, including a tri-syalilated bi-antennary glycan structure (fig. 3.4) which was previously reported on liver. This structure suggest the presence of 1,3-Galactose transferase in spleen.

In conclusion, results from the N-glycomic analysis carried out on wild-type and *Mgat5* deficient mouse kidney, indicate that ko mice show less presence of tri-antennary. It was not possible to determine whether high m/z values correspond to tetra-antennary structure or to elongated tri-antennary structures. A linkage analysis should be performed to identify which tri-antennary structure (2,2,6 or 2,4,2) is present in ko sample, therefore establishing whether GlcNAcT-V is or not the only β-1,6 GlcNAc transferase expressed in kidney.

GlcNAc salvage on the diet has no apparent effect on N-linked glycosylation in the kidney and spleen, even if experiments conducted on cell lines demonstrate that increased influx of UDP-GlcNAc resulted on increased N-glycan branching. Moreover, the performance of optimized glycome procedure allowed the identification of more tri-antennary glycan structures than the one reported on the CFG database.

5. REFERENCES

Abbott, K. L., Troupe, K., Lee, I., Perce, M. (2006). Integrin-dependent neuroblastoma cell adhesion and migration on laminin is regulated by expression levels of two enzymes in the O-mannosyl-linked glycosylation pathway, PomGnT1 and GnT-Vb. *Experimental Cell Research*, 312:**15**, 2837-2850.

Alverez, K., Haswell, C., St. Clair, M., Perng, G. S., Shorebah, M., Pierce, M., Fregien, N. (2002). Sequences of the ouse N-acetylglucosaminyltranferase V (*Mgat5*) mRNA and an mRNA expressed by an *Mgat5*-deficient cell line. *Glycobiology*, **12**:7, 389-394.

Apweilwe, R., Hermjakob, H., Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta*, **1473**:1, 4-8.

Cheung, P., Pawling, J., Partridge, E. A., Sukhu, B., Grynpas, M., Dennis, J. W. (2007). Metabolic homeostasis and tissue renewal are dependent on β1,6GlcNAc-branched N-glycans. *Glycobiology*, 17:**8**, 828-837.

Coutinho, P. M., Deleury, E., Davies, G. J., Henrissat, B. (2003). An evolving hierarchical family classification of glycosyltransferases. *Journal of Molecular Biology*, **328** :2, 307-317.

Dell, A., Morris. H. (2001). Glycoprotein structure determination by mass spectrometry. Science, **291**:5512, 2351-2356..

Demetriou, M., Granovsky, M., Quaggin, S., Dennis, J. W. (2001). Negative regulation of T-cell ativation and autoimmunity by *Mgat5N*-glycosylation. *Nature*, **409**, 733-739.

Dennis, J. W., Laferte, S. (1989). Oncodevelopmental expression of –GlcNAcβ1-6Manα1-6Manβ1—branched Asparagine-linked oligosaccharides in murine tissues and human brest carcinomas. *Cancer Research*, **49**, 945-950.

Dennis, J. W., Granovsky, M., Warren, C. E. (1999a). Protein glycosylation in development and disease. *BioEssays*, **21**, 412-421.

Dennis, J. W., Granovsky, M., Warren, C. E. (1999b). Glycoprotein glycosylation and cancer progression. *Biochimica et Biophysica Acta*, **1473**, 21-34.

Dennis, J. W., Warren, C. E., Granovsky, M., Demetriou, M. (2001). Genetic defects in N-glycosilation and cellular diversity in mammals. *Current Opinion and Structural Biology*, **11**:5, 601-607.

Granovsky, M., Fode, C., Warren, C. E., Campbell, R. M., Marth, J. D., Pierce, M., Fregian, N., Dennis, J. W. (1995). GlcNac-transferase V and core 2 GlcNac-transferase expression in the developing mouse embryo. *Glycobiology*, **5**:8, 797-806.

Granovsky, M., Fata, J., Pawling, J., Muller, W. J., Khokha, R., Dennis, J. W. (2000). Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nature Medicine*, **6**:3, 306-312.

Haslam, S. M., North, S. J., Dell, A. (2006). Mass spectrometric analysis of N- and O-glycosylation of tissues and cells. *Current Opinion in Structural Biology*, **16**, 584-591.

Lau, K. S., Partridge, E. A., Grigorian, A., Silvescu, C. I., Reinhold, V. N., Demetriou, M., Dennis, J. W. (2007). Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell*, **129**, 123-134.

Lau, K. S., Khan, S., Dennis, J. W. (2008). Genome-scale identification od UDP-GlcNAc-dependent pathways. *Proteomics*, 8, 3294-3302.

Lis, H. and Sharon, N. (1993). Protein glycosilation: structural and functional aspects. *European Journal of Biochemistry*, **218**, 1-27.

Morgan, R., Gao, G., Pawling, J., Dennis, J. W., Demetriou, M., Li, B. (2004). N-Acetylglucosaminyltransferase V (Mgat5)-mediated N-glycosylation negatively regulates Th1 cytokine production in T cells. *The Journal of Immunology*, **173**, 7200-7208.

North, S. J., Jang-Lee, J., Harrison, R., Canis, K., Ismail, M. N., Trollope, A., Antonoupoulos, A., Pang, P. C., Grassi, P., Al-Chalabi, S., Etienne, A. T., Dell, A., Haslam, S. M. (2010). Mass spectrometry analysis of mutant mice. *Methods in Enzymology*, **470**, 27-77.

Takamatzu, S., Antonopoulos, A., Ohtsubo, K., Ditto, D., Chiba, Y., Le, D. T., Morris, H. R., Halslam, S. M., Dell, A., Marth, J. D., Taniguchi, N. (2010). Physiological and glycomic characterization of N- acetylglucosaminyltransferase-IVa and –Ivb double deficient mice. Glycobiology, **20**:4, 485-497.

Taylor, M. E. and Drickmar, K. (2010). Introduction to Glycobiology, third edition. *Oxford University Press*.

CFG databases (http://www.funtionalglycomics.org)

6. SUPPLEMENTERY MATERIAL

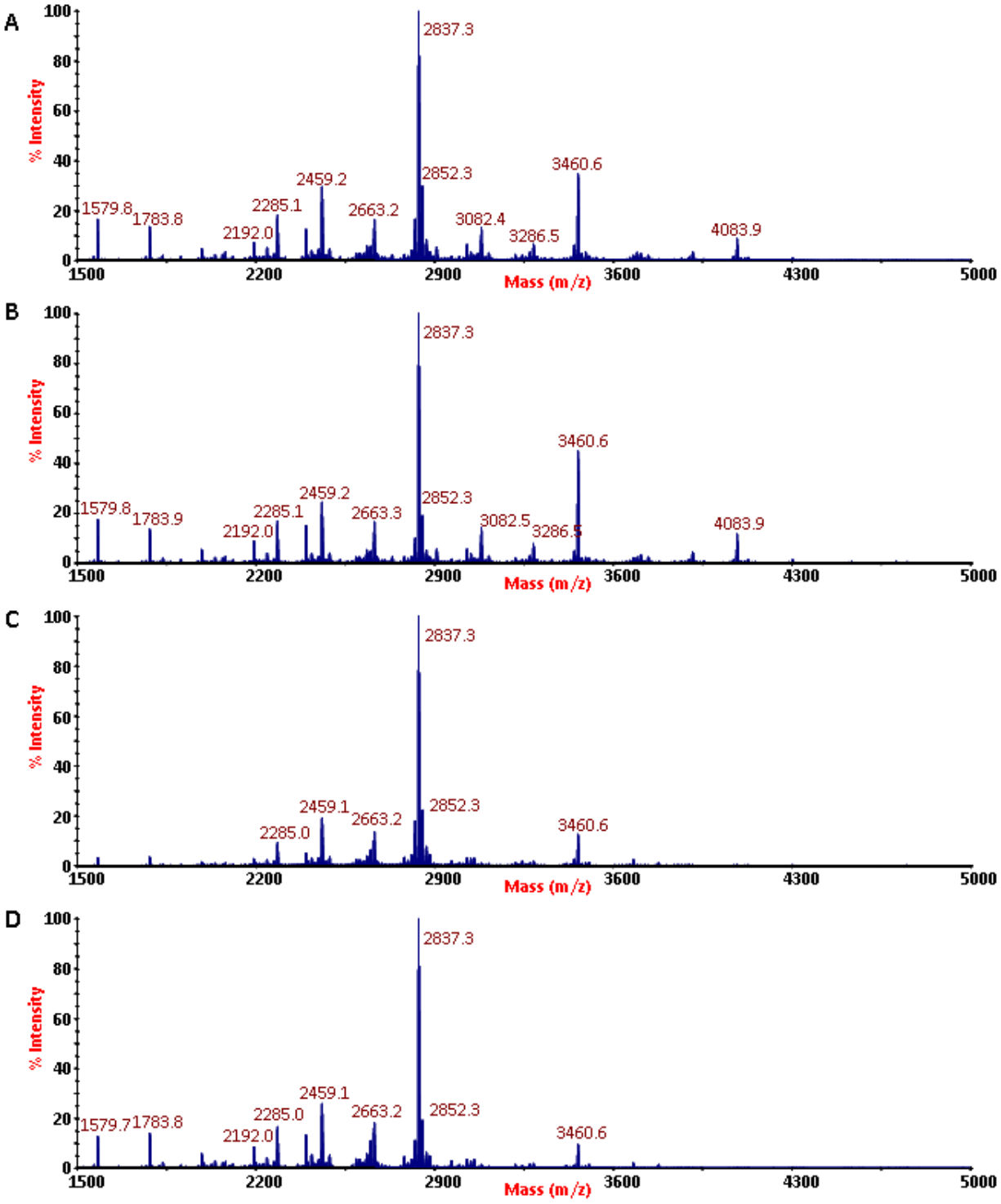


Fig. 6.1 MALDI-TOF spectra of wild type (A) and GlcT-V deficient (B) mouse kidney N-glycan and of wild type (C) and GlcT-V deficient (D) mouse kidney fed with additional GlcNAc on the daily diet. Molecular mass ions of the most intense peaks are indicated. To make the spectra comparison more immediate, glycan annotation was not shown here.

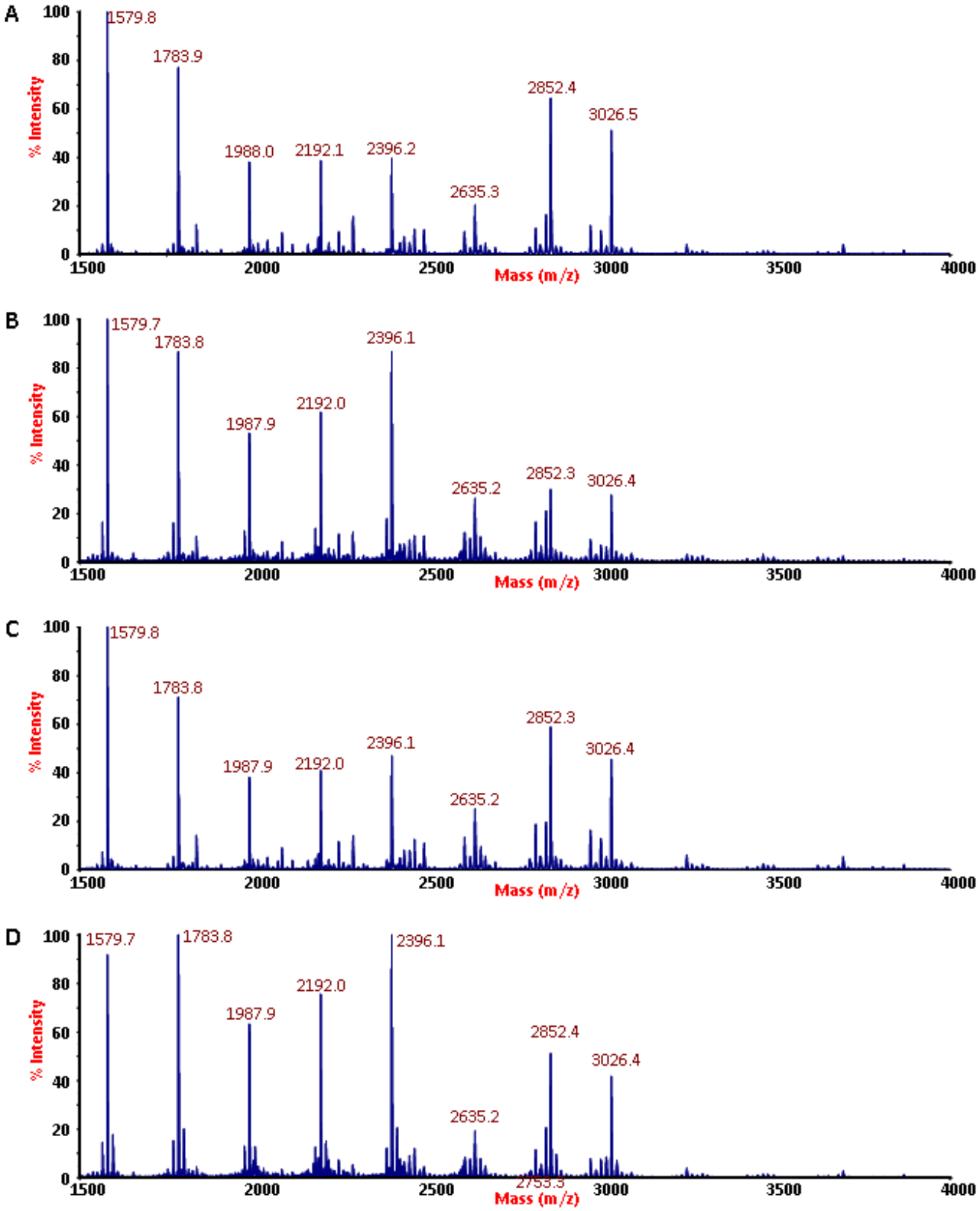
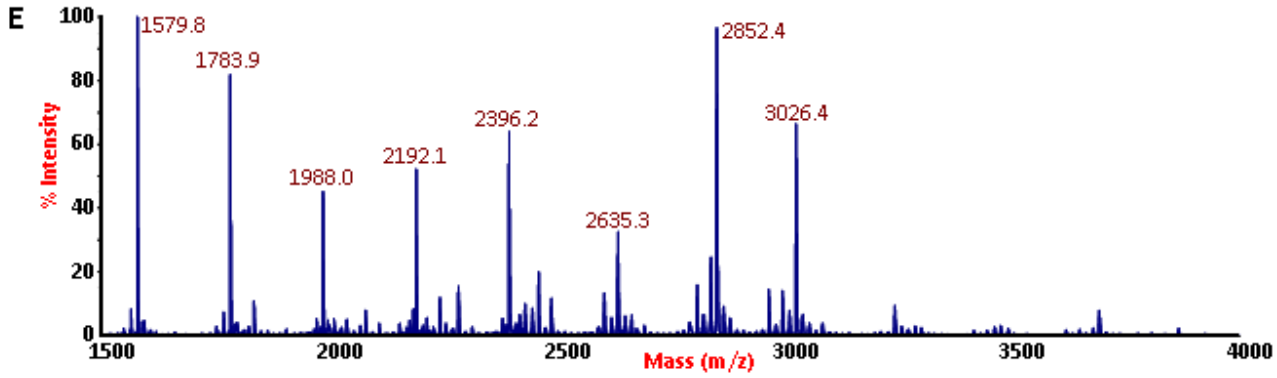
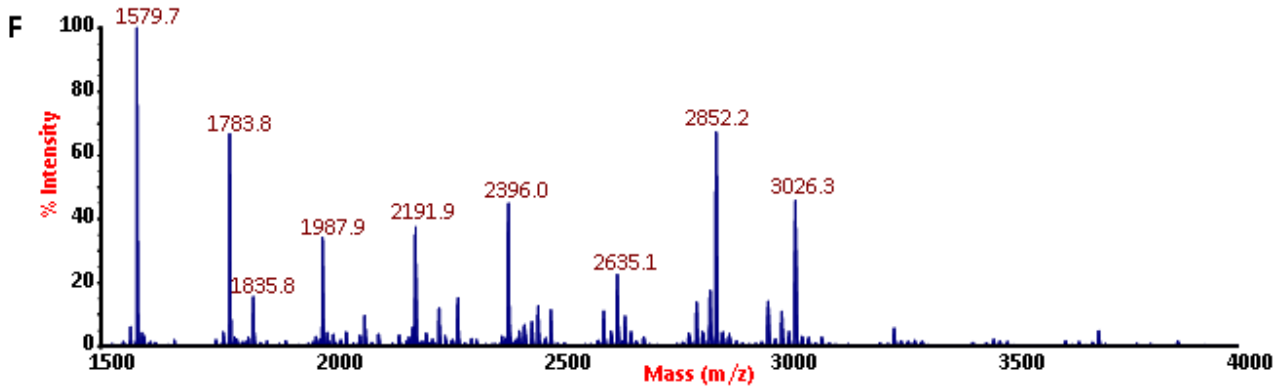


Fig. 6.2 MALDI-TOF spectra of spleens from: control mouse, standard fed (A and B), supplementary 0.5 mg/ml per day GlcNAc fed mouse (C), supplementary 5 mg/ml per day GlcNAc fed mouse (D and E), supplementary 15 mg/ml per day GlcNAc fed mouse (F). Molecular mass ions of the most intense peaks are indicated. To make the spectra comparison more immediate, glycan annotation was not shown here.


E
100
80
60
40
20
0
% Intensity
1579.8
1783.9
1988.0
2192.1
2396.2
2635.3
2852.4
3026.4
1500
2000
2500
3000
3500
4000
Mass (m/z)


F
100
80
60
40
20
0
% Intensity
1579.7
1783.8
1835.8
1987.9
2191.9
2396.0
2635.1
2852.2
3026.3
1500
2000
2500
3000
3500
4000
Mass (m/z)

Division of Molecular Biosciences
Department of Life Sciences
Faculty of Natural Sciences
Imperial College London

Biochemistry Building Level 1
South Kensington Campus
London SW7 2AZ
Tel: +44 (0)207 594 5219 Fax: +44 (0) 207 225 0458
Professor Anne Dell CBE FRS FMedSci

Report prepared by Micaela Pivato

Precursor of (1,3;1,4)-beta-D-Glucans in cell walls

AIM: identification of putative lipid linked glucosides associated with the production of mixed linked glucan (MLG).

The starting hypothesis is that putative lipid linked glucosides could be relatives of the sterol glycosides previously reviewed in Kovganko and Kashkan (Chemistry of Natural Compounds, 1999), although we aimed to use methodologies that potentially would reveal other families of glycolipid.

Samples, which arrived on the 6th of June, were the extracted glycolipidic fraction (divided into polar and non-polar lipidic extracts), of the following tissues:

- *Nicotiana benthamiana* leaf tissue infiltrated with agrobacterium strain AGL1 (negative control);
- *Nicotiana benthamiana* leaf tissue infiltrated with agrobacterium strain AGL1 containing barley CslF6, shown to contain mixed link glucan;
- Barley coleoptiles tissue, shown to contain high levels of mixed link glucan.

According to our working hypothesis, we are looking for the m/z values listed in the following table.

	m/z value depending on ionization		
Chemical	H^+	Na^+	K^+
Beta-Sitosterol	415	437	453
sitosterol + 1 glucose	577	599	615
sitosterol + 2 glucose	739	761	777
sitosterol + 3 glucose	901	923	939

Considering that molecular masses of phytosterols range from 396-414 Da, more values than the listed ones were taken in account.

According to the literature, sterol glycosides should be found in the polar lipidic fraction.

It is essential to note that sterol glycosides are important component of plants, both for membrane permeability and as intracellular metabolic actors. We should expect to find them in all the samples, with differences in the ratio between MGL expressing and non expressing tissues.

Preliminary tests

Several solutions (made of water, methanol and chloroform in different ratios) were tested to solubilize the samples, and water/methanol 1:1 solution was proved to be the best. Since non polar fractions were unexpectedly insoluble in chloroform, further separation into polar and non-polar fractions was performed. Different MALDI matrices were then tested, to compare analyte compatibility, ionization and co-crystallization ability. For the early analyses, DHB and HABA matrices were used in parallel, whilst HABA matrix was later preferred.

All the 9 samples (polar extracts of the 3 samples, and polar and non-polar fraction of the non-polar extracts) were analyzed on the MALDI-tof spectrometer without any further purification (fig. 1, 2 and 3). A test run using the spectrometer in negative mode was also performed.

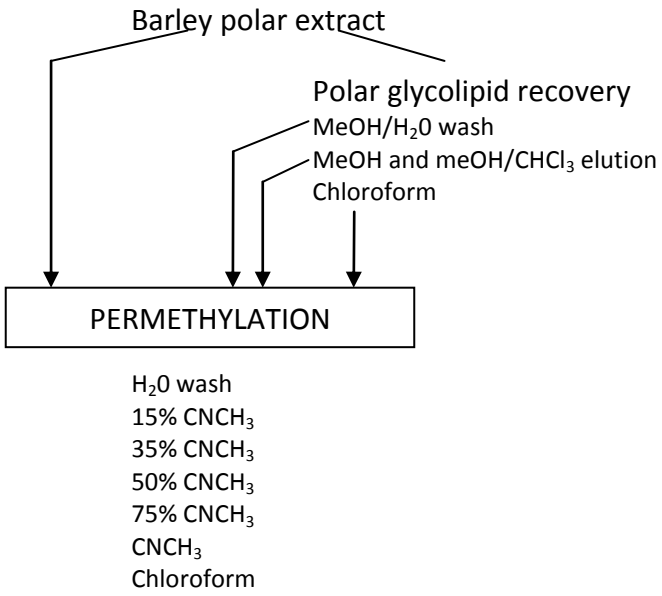
Spectra profiles verify that non-polar extracts contain polar components (polar fractions of non-polar extracts spectra showed similar profiles to the corresponding polar extracts). Moreover, *N. benthamiana* negative control and *N. benthamiana* MGL expressing tissue spectra profiles were extremely similar and shown the same pattern of peaks. The polar extract of barley tissues showed high molecular weight components that differ by 162 Da which is the mass of hexose. The molecular ions correspond to non-substituted hexose chains that differ in the number of units (from 5 up to 15 units). (N.B. MS data cannot distinguish between glucosyl and other hexosyl oligomers).

Regrettably, none of the values listed in the table were found in the MGL expressing tissues. Moreover we did not observe any molecular ions attributable to lipid-like molecules.

Chemical modification

It is well established that derivatization of samples improve MS results in terms of analyte ionization and sensitivity. The protection of functional groups prevents unwanted reactions.

Permethylation was performed on the native polar extract and on polar fractions purified by hydrophobic chromatography (C18 Sep-pak column). Purification of permethylated samples was also performed using hydrophobic chromatography. Elution steps are listed below.



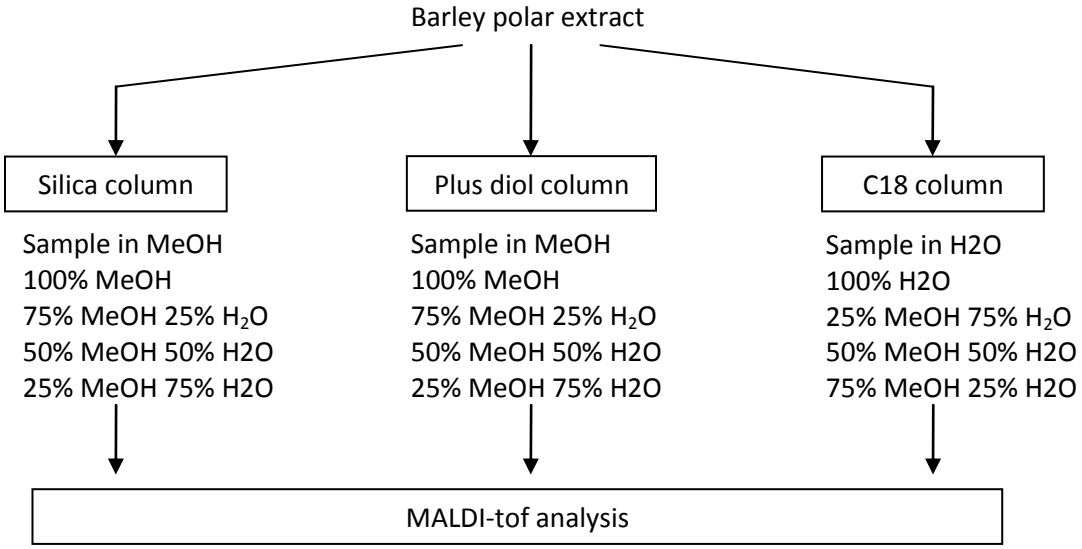
Permethylation substitutes the hydrogen of a hydroxyl group with a methyl group, so that we expect a +14 Da imcrement at each hydroxyl group. (+14 for single sterol, +14 x5 for a free glucose, +14 x4 for glucose with one oxygen linked to sterol, and so on). Each of the 7 elution steps of the 4 samples (native and purified fractions of polar barley) were analyzed on the MS. Unfortunately, none of the expected m/z values were found. At high molecular weight, peaks with nice Gaussian profiles (fig. 4), with higher intensity compared to the non permethylated sample, indicated the presence of hexose oligomers, as seen in the native samples. However, once again their were no peaks attributable to glycolipids.

Other experiments tried were acetylation and trimethylsilylation derivatization methods with comparable results.

Trimethylsilylation was performed after a different chromatographic purification of the native sample (Silica column; elution steps: chloroform, 1:1 chloroform/acetone, acetone, 3:1 acetone/methanol, 1:1 acetone/methanol, methanol, 1:1 methanol/water). None of the elution fractions analyzed by MS showed the m/z value expected for phytosterol or sterol glycosides. We also performed MS/MS of ions with higher intensity, looking for the loss of a glucose unit but none was observed.

Sample fractionation

At this point, since we believed that sterol glycosides should be present in our samples (at least as membrane and signalling pathways, if not as MGL biosynthesis precursors), we focused our efforts on fractionation, in order to possibly enrich a fraction containing sterol glycosides. To this aim we used chromatography on different sep-pak cartriges: Silica (polar), Plus diol (polar) and C18 (non-polar) working in the polarity range of methanol and water (according to literature of sterol glycosides). Once again, despite carrying out many different experiments we did not obtain meaningful data.



GC-MS

Gas chromatography-MS (GS-MS) combine gas liquid chromatography with in tandem mass spectrometry. The ionization approach is not "soft" as the laser mediated, but "harder". Ionization with electron impact (EI) actually results in a direct fragmentation of analytes (more powerful than a MS/MS experiment).

GS-MS is commonly used for the definition of linkage and configuration of sugars (Dell and Morris, 2001).

TMS analyses of standard and barley sample

β-sitosterol standard (Sigma) was spotted on a MALDI plate both native and after parmethylation. Spectra profile did not reveal a unique/prevalent m/z peak value corresponding to the attended mass, but a variety of peaks (data not shown). Therefore, we conclude that MALDI-MS was not the suitable technique for the identification of sterols (probably due to deficient ionization ability).

GS-MS analyses ware then performed on TMS-modified β-sitosterol and TMS-derivatized barley sample. Results are shown in figure 5 and 6, respectively. Standard profile shows a main peak at RT (retention time) 28.59 min, which was identified as β-sitosterol by the correspondent MS spectra profile, according both with literature (*J Agric Food Chem*, 1983), and with the instrument chemical database.
TMS-derivatized barley sample results in a nice resolved chromatogram. Each elution peak was analyzed one by one in order to find correspondence with TMS β-sitosterol standard, whitout positive output.

(Further experiments: it will be nice to try to remove glycans from lipids, and use another GC colum just to study the lipidic composition of the sample)

Linkage analysis of barley sample

Linkage studies consist in the analysis of PMAA (partial methylated alditol acetate) derivatives of monosaccharide constituents that allow the identification of glycosidic bond position. In brief, the protocol plans sample methylation of the hydroxyl groups, hydrolysis of glycan chain into monosaccharides, distruption of the monosaccharide ring and acetilation of the free hydroxyl groups (the ones which were not methylated, and get free after ring disruption). Identification of monosaccharides is achieved by comparison of GC RT and EI-MS spectra with standards (North et al., *Medods of Enzymology*, 2010).

75% and 100% MeCN fractions of permethylation were precessed for linkage analysis. From MALDI-MS data we can observe polydisperse Hexose chains in the range of 1-40 units, with a Gaussian profile (fig. 4). Figure 7 report the GC chromatograms of barley sample (A) and 1,4-cellobiose (B) and 1,3-linked-laminarilbiose (C) standard.

Comparing MS profile of each sample elution peak with standard here reported and standards of previous reported in our lab, we were able to identify several sugars: t-Glu

(fig. 8), which represent the 61% of the total recognized sugars component, 1,4-Glu and 1,3-Glu (fig. 9 and 10), present in small amount, 0.3 and 0.1% of the total recognized sugars component, 1,6 and 2,6-Hexose, putatively Glucose (fig. 11), 4 and 5% of the total recognized sugars component. Peaks at RT 17.6-17.7 and RT 19.4-19.5, that in total correspond to 29% of the total recognized sugars component, were difficult to assign: their mass profile resemble hexofuranose structures differently linked.

The following table summarize the results, and percentage reported are calculated on total elution area, which include peak at RT 13 (which is not sugar) and other smaller peaks.

Characteristic fragment ions	Assignment	Retention Time (min)	Relative Abundance
84, 110 111, 153, 154, 213	Other	13.87	30.3 %
72, 87, 101, 102, 129, 161, 162	Hexofuranose ?	17.7	11.8 %
87, 88, 101, 102, 118, 129, 145, 161, 162, 205	t-Glc	18.4	41.8 %
87, 129, 188, 190, 234	Hexofuranose ?	19.4-19.5	8.4 %
59,71, 74, 87, 101, 118, 129, 161, 243	1,3-Glc	19.55	0.1 %
87, 99, 102, 113, 118, 129, 162, 187, 233, 277	1,4-Glc	19.68	0.2 %
87, 88, 99, 102, 118, 129, 162, 189	6-Hex	19.91	2.7 %
87, 88, 129, 130, 188, 190	2,6-Hex	21.06	3.6 %

Such a small presence of putative MLG (1,3- and 1,4-Glu) could be explained by the fact that MLG is not present in the glycolipid polar extraction, and could probably have more affinity for the polar extraction procedures that peceed glycolipid extraction.

FIGURES

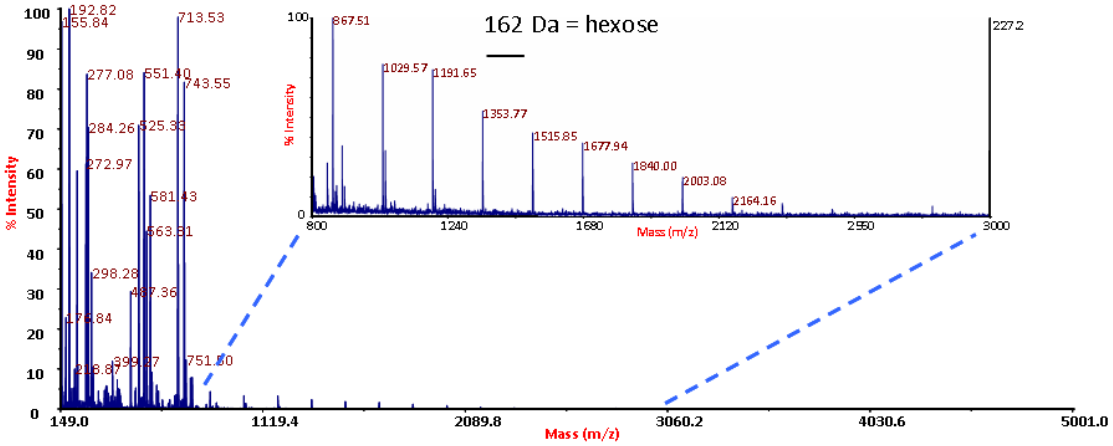


Fig. 1 MALDI-tof MS profile of Barley, polar fraction (native sample).

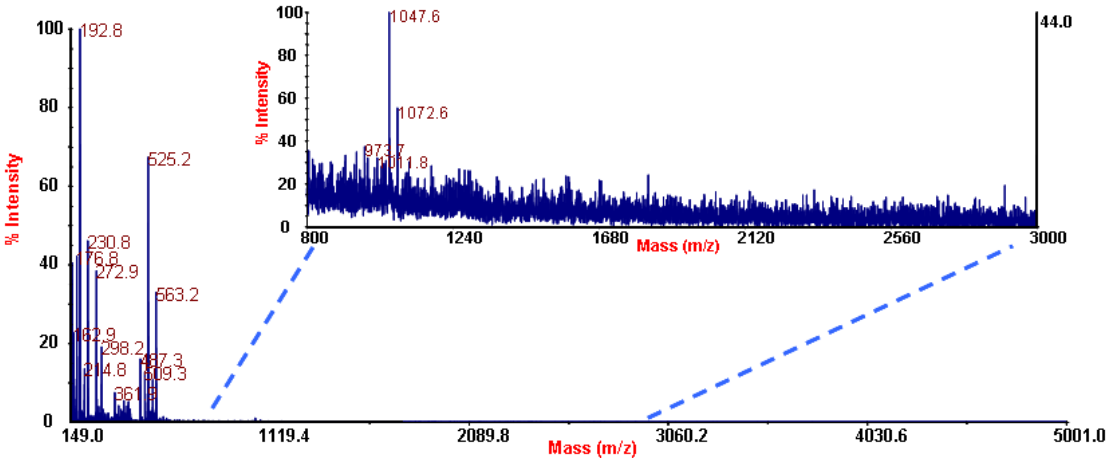


Fig. 2 MALDI-tof MS profile of *N. Benthamiana*, negative control, polar fraction (native sample).

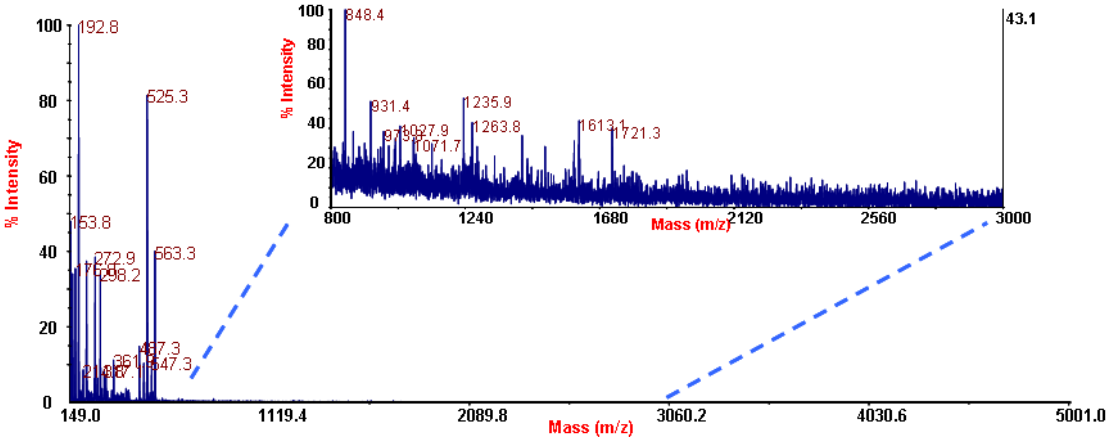


Fig. 3 MALDI-tof MS profile of *N. Benthamiana*, MGL expressing, polar extract (native sample).

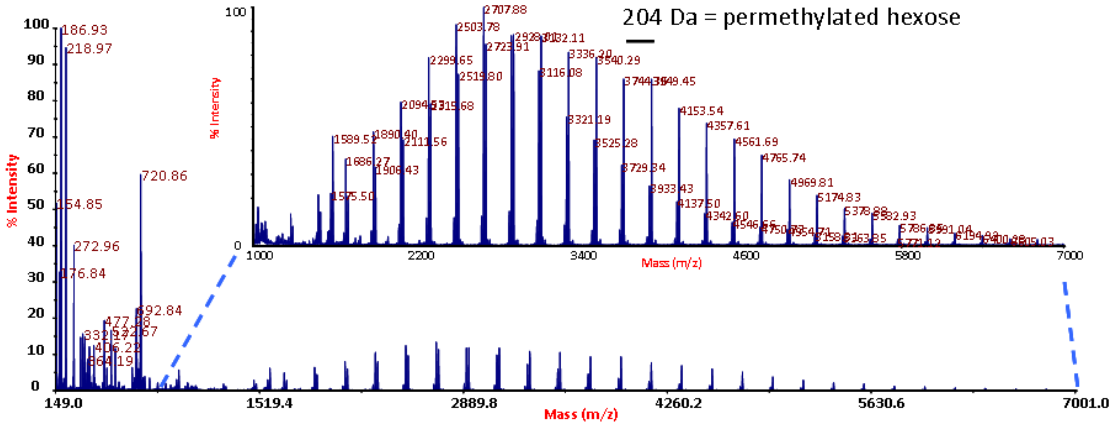


Fig. 4 MALDI-tof MS profile of permethylated Barley polar extract (100% acetonitrile fraction).

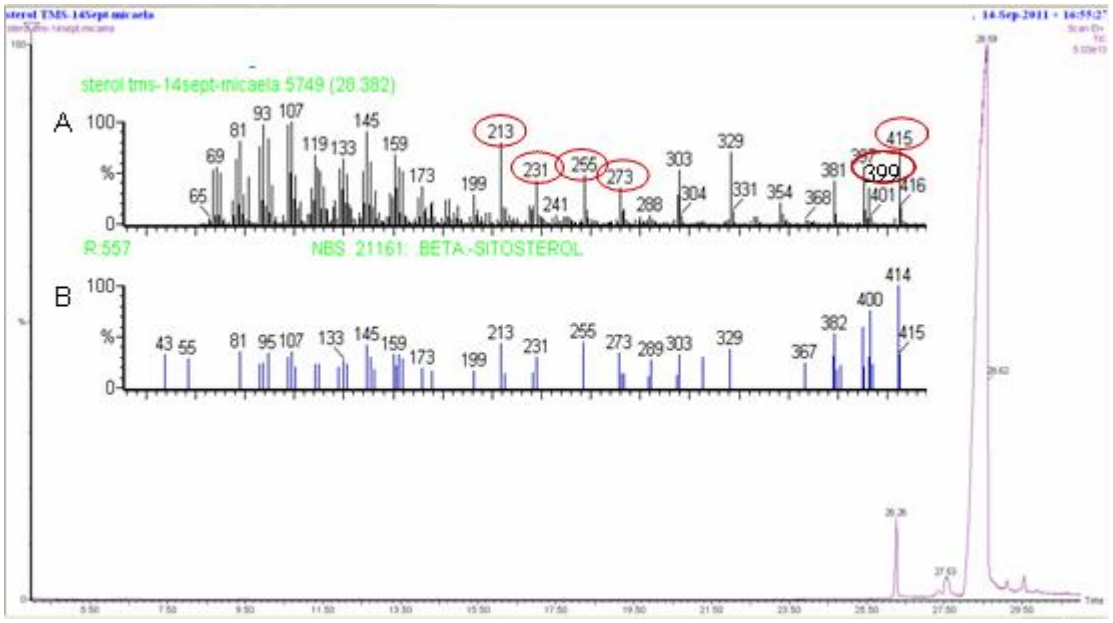


Fig. 5 GC elution of TMS derivatized β-sitosterol standard. β-sitosterol elutes at RT 28.59 min, MS spectra of the chromatogram peak is shown in panel A. Highlighted m/z values refers to MS peak values indicated in Daily and colleagues (*J Agric Food Chem*, 1983) for β-sitosterol, while panel B refers to β-sitosterol MS spectra reported in our database. MS of peak correspondent to RT 26.26 of GC has shown that it could be a chemical modified species of β-sitosterol.

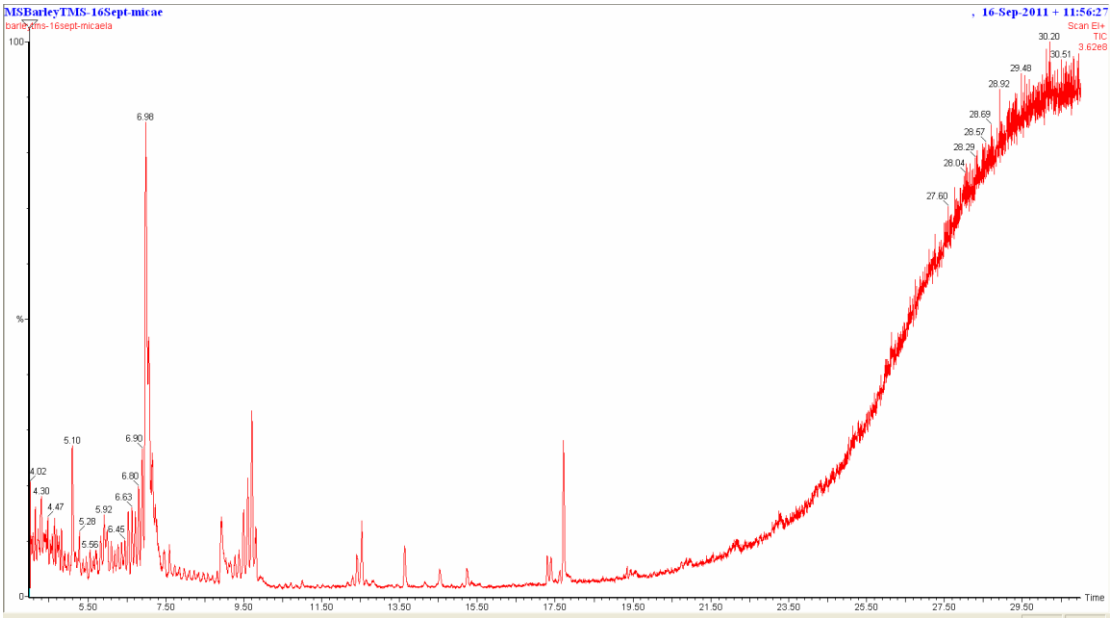


Fig. 6 GC elution profile of TMS barley sample. MS profiles of each elution peak were analyzed one by one in order to find correspondence with TMS β-sitosterol standard, whitout positive output.

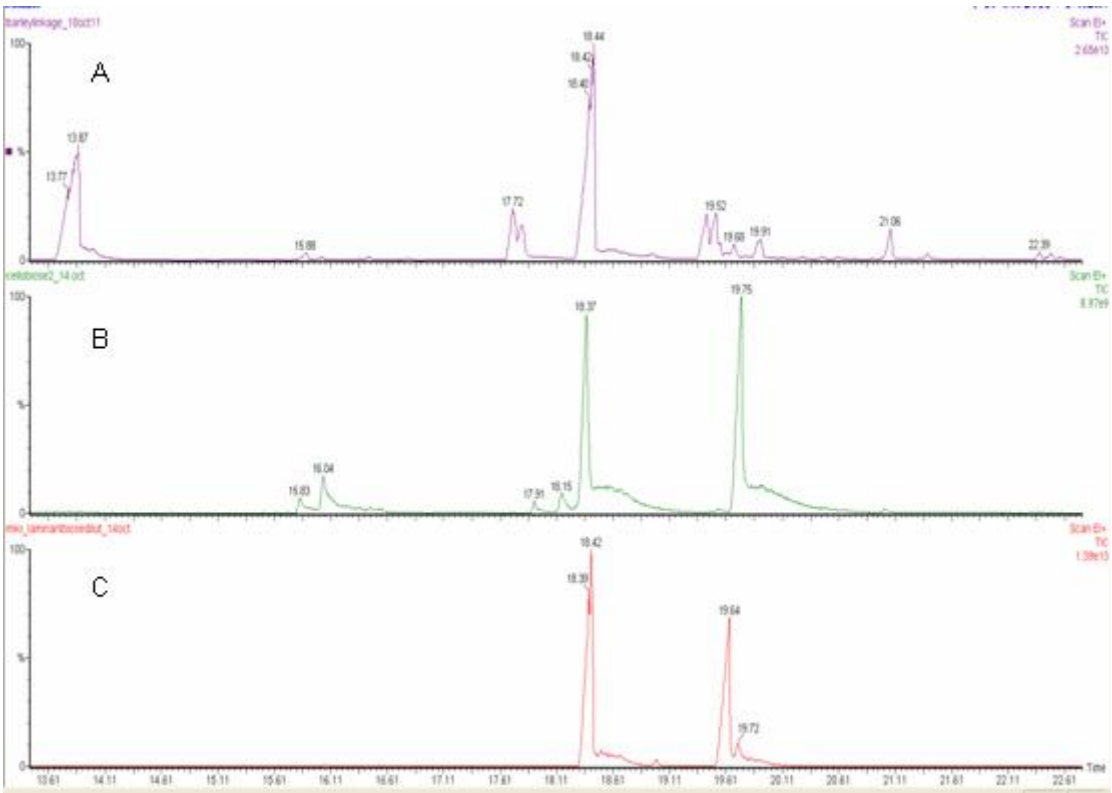


Fig. 7 GC chromatogram of linkage analysis of barley sample (A), 1,4-linked-cellobiose (B) and 1,3-linked-laminarilbiose (C).

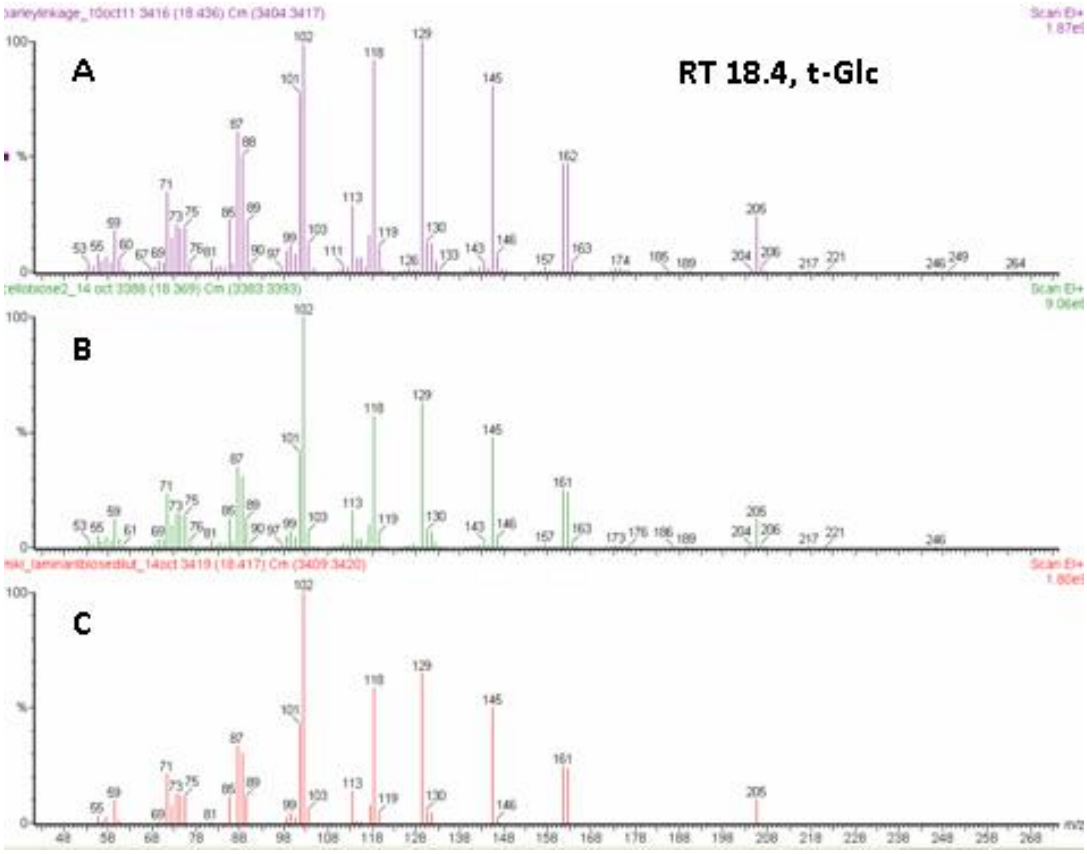


Fig. 8 MS spectra of sample RT 18.4 (A) , 1,4-cellobiose RT 18.3 (B) and 1,3-laminarilbiose RT 18.4 (C), corresponding to teminal-Glucose.

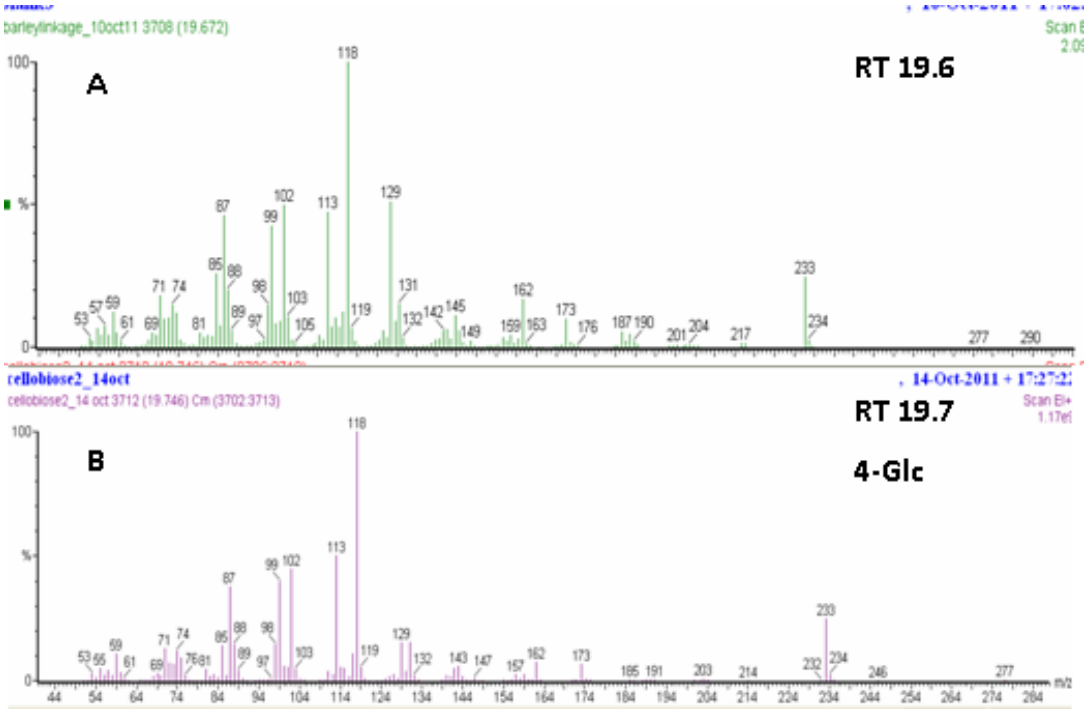


Fig. 9 MS spectra of sample RT 19.6 (A) and 1,4-cellobiose RT 19.7 (B), corresponding to 4-linked Glucose.

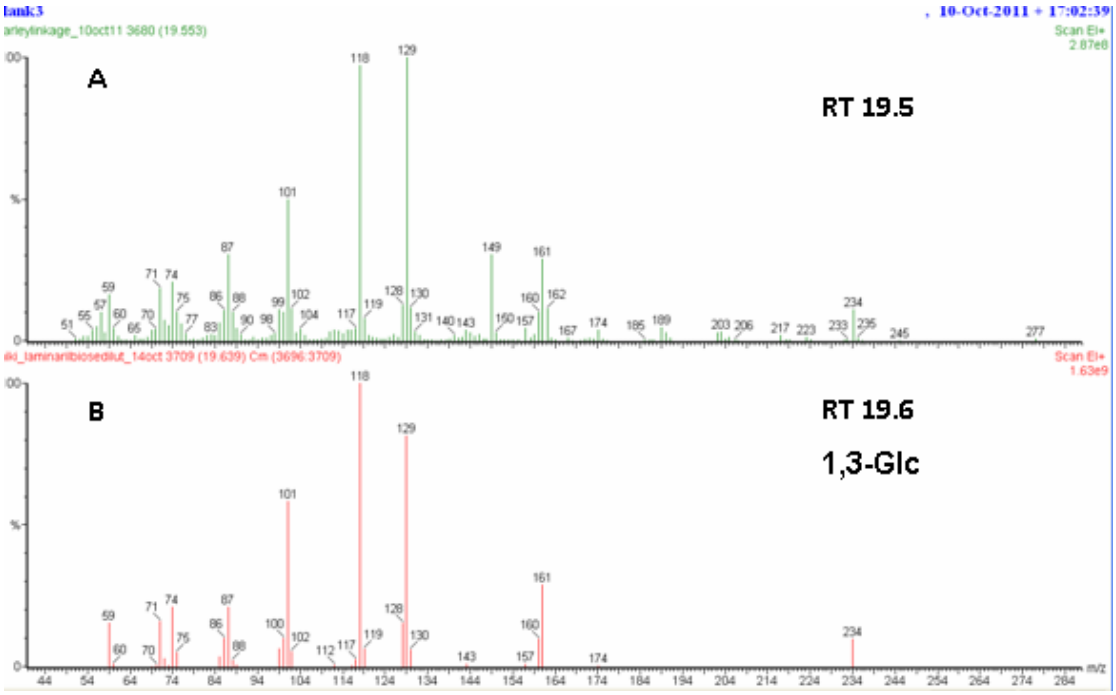


Fig. 10 MS spectra of sample RT 19.55 (A) and 1,3-laminarilbiose RT 19.64 (B), corresponding to 3-linked Glucose.

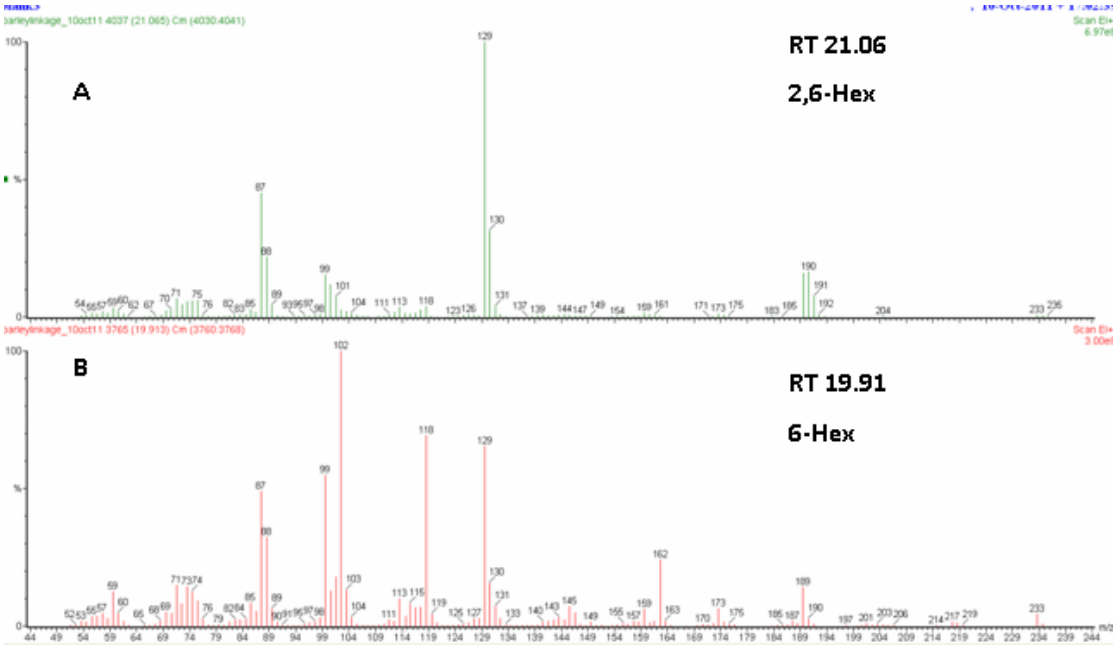


Fig. 11 MS spectra of sample RT 21.06 (A) and RT 19.91 (B), corresponding to 2,6-Hex and 6-Hex linked Glucose (spectra compared with standards that are not shown here).

Acknowledgments

Many thanks to my supervisor Prof. Patrizia Polverino de Laureto for having accompanied me in my PhD and for the guidance through the project. Thanks to Dr. Giorgia de Franceschi and Dr. Erica Frare for the technical support and assistance. I also have to thank Prof. Luigi Bubacco and Dr. Isabella Tessari for the conception of the dimers and the support in the molecular biology section. Thanks to Dr. Marco Bisaglia for the NMR experiments and to Dr. Marco Brucale for the AFM data reported in this thesis. I would like to mention Mr. Beppe Tognon that introduced me to electron microscopy with enthusiasm.

I would like to thank Dr. Anne Dell for the opportunity to work in her lab, and Dr. Stuart Haslam for having encouraged me to always formulate new hypothesis and test them. Moreover, special thanks to Dr. Paola Grassi and Dr. Poh Choo Pang for the technical teaching, assistance and friendship.

Un'esperienza coinvolgente e di forte impatto come il dottorato di ricerca è destinata inesorabilemente a deviare il percorso di vita di chi la intraprende, sia per quanto riguarda l'ovvia formazione professionale che tale strada implica, ma anche per l'evoluzione personale che ne deriva. Le scelte dell'Università in cui compiere tale esperienza, Padova o Zurigo, o del progetto di ricerca nel quale essere coinvolta, in se, si sono compiute in pochi, frenetici, giorni, senza che mi rendessi conto di quanto avrebbero condizionato la mia vita.

Vorrei qui ringraziare alcune persone che ho incontrato SOLO grazie a questo dottorato, e che hanno, in diversi modi, contribuito alla mia evoluzione in questi ultimi, intensi, tre anni. Grazie alle mie coinquiline, Dania, Susi e Marta, alle mie colleghe e amiche Odra, Samanta e Chiara, e agli amici (e colleghi) della porta accanto, Andrea, Elisa, Laura, Lisa, Robin e Fabrizio. Anche non trovando gli aggettivi adatti a descrivervi, voi rappresentate il più stimolante team di sostegno ed energia a cui poter ricorrere sia nel momento del bisogno (in particolare grazie per il più improbabile viaggio di ritorno from UK) che, fortunatamente, nei momenti di svago. ***Grazie****.*

E grazie a chi, in questo ultimo pezzo del mio viaggio, è rimasto forse un po' dietro le quinte, registrando con la costante empatia ogni mio mutamento: alla mia famiglia, ***Grazie****.*