

**Alma Mater Studiorum
Università di Bologna**

**Dottorato di Ricerca in
Scienze Biochimiche e Biotecnologiche**

XXVI ciclo

Settore Concorsuale di afferenza: 05/G1

Settore Scientifico disciplinare: BIO14

**Characterization of the Caspr2 and NLGN2
ligands: a proteomic and biochemical approach**

Presentata da:

Dott. Gabriele Vincelli

Coordinatore Dottorato

Chiar.mo Prof.
Santi Mario Spampinato

Relatore

Chiar.mo Prof.
Santi Mario Spampinato

Correlatore

Prof. Davide Comoletti

ANNO ACCADEMICO 2012-2013

INDEX

	Page
1. INTRODUCTION	1
1.1 AUTISM SPECTRUM DISORDERS	1
1.1.1 Symptoms of ASD	2
1.1.2 Causes of ASD	4
1.1.3 Neuropathology of ASD	8
1.2 CASPR2	10
1.2.1 The CNTNAP2 gene and ASD	10
1.2.2 Caspr2 is involved in other diseases	12
1.2.3 Caspr2, a member of the Neurexin superfamily	13
1.2.4 Role of Caspr2 in myelinated axons	15
1.2.5 Other Caspr2 localizations	21
1.3 NEUREXINS AND NEUROLIGINS	22
1.3.1 Neurexins and Neuroligins in ASD	22
1.3.2 Neurexins and Neuroligins: two synaptic adhesion proteins	23
1.3.3 The Neurexin-Neuroligin complex	26
1.3.4 Biological role of the complex	28
1.4 OTHER PROTEINS IDENTIFIED IN THE WORK	32
1.4.1 Secreted Frizzled Related Protein 1	32
1.4.2 Clusterin	34
1.4.3 Apolipoprotein E	35
1.4.4 Tenascin R	36
2. AIM OF THE RESEARCH	39
3. MATERIALS AND METHODS	42
3.1 MOLECULAR AND CELL BIOLOGY	42
3.1.1 Eukaryotic Cell Culture	42
3.1.2 Bacterial culture	42
3.1.3 Plasmids used	43
3.1.4 cDNAs	45
3.1.5 Cloning	45
3.1.6 Bacterial transformation	47
3.1.7 Plasmid DNA purification	47
3.1.8 HEK cells' transfection	48
3.1.9 Selection of stable transfected cells	48
3.1.10 Protein extraction from HEK cells	48
3.1.11 Protein purification	48
3.1.12 Size exclusion chromatography	50
3.2 BIOCHEMISTRY	50
3.2.1 SDS-PAGE and Western Blot	50
3.2.2 Co-immunoprecipitation experiments	51
3.3 PROTEOMICS	52
3.3.1 Fishing experiments	52
3.4 BIOPHYSICS	54
3.4.1 Isothermal Titration Calorimetry	54

4. RESULTS	56
4.1 CHARACTERIZATION OF THE BINDING BETWEEN CASPR2 AND TAG-1	56
4.1.1 Biophysical analysis of the complex	56
4.1.2 Co-immunoprecipitations using full-length and truncated proteins	59
4.1.3 Validation of the co-IP protocol	63
4.2 IDENTIFICATION OF NOVEL BINDING PARTNERS FOR CASPR2 AND NLGNs' FAMILY MEMBERS	68
4.2.1 Fishing experiments	68
4.2.2 Secreted Frizzled Related Protein 1	75
4.2.3 Clusterin	78
4.2.4 Apolipoprotein E	83
4.2.5 Contactin 1	86
4.2.6 Tenascin-R	91
5. CONCLUSIONS	93
6. BIBLIOGRAPHY	97
ACKNOWLEDGMENTS	109

1. INTRODUCTION

1.1 AUTISM SPECTRUM DISORDERS

Autism spectrum disorders (ASD) describes a range of early-onset conditions classified as neurodevelopmental disorders, characterized by deficits in social interactions and communication, as well as by restricted interest and repetitive behaviors. The term “spectrum disorder” expresses the concept that, although a classification is generally done, the diseases are a continuum of the same condition, with different causes, not well-defined symptom boundaries, as well as major differences in developmental course, with different children manifesting sign of the disorder with different timing (Nolen-Hoeksema et al., 2009) (fig.1.2). The classification of ASD includes different pathologies like:

- Autism
- Asperger Syndrome (merged with autism in the new Diagnostic and Statistical Manual of Mental Disorders (DSM-5) of May 2013 [Lauritsen, 2013])
- Pervasive Developmental Disorder - Not Otherwise Specified (PDD-NOS).

Rett Syndrome and Childhood Disintegrative Disorder are also listed among ASDs in the current version of the DSM-5, even if they might have different causes (Abrahams and Geschwind, 2008).

The prevalence of ASD has been estimated in one affected child every 88 (fig.1.1). ASDs are almost 5 times more common in boys (1 in 54 children) than in girls (1 in 252 children) (Autism and Developmental Disabilities Monitoring Network Surveillance, 2012), an effect that becomes even more pronounced in so-called “high-functioning cases” (Abrahams and Geschwind, 2008).

Due to the communication and social deficits, affected children have problems in making friendships and in social relationships in general. These disorders present very high clinical heterogeneity, with very different pervasiveness, severity and onset, thus affecting life outcome in very different ways. In a study published in 2004 the authors followed 68 individuals that during childhood were diagnosed with ASD: in adult age 1/5 earned an academic diploma, five attended college and two had a post-baccalaureat degree, almost 1/3 had a job and 1/4 managed to have important friendships; however, 58% didn't manage to live alone or to keep a job, having persistent communication and social deficits (Howlin et al., 2004). The poor quality of life and the limited understanding of the biology of autism make render the research on ASD of high importance for improving the lives of affected individuals.

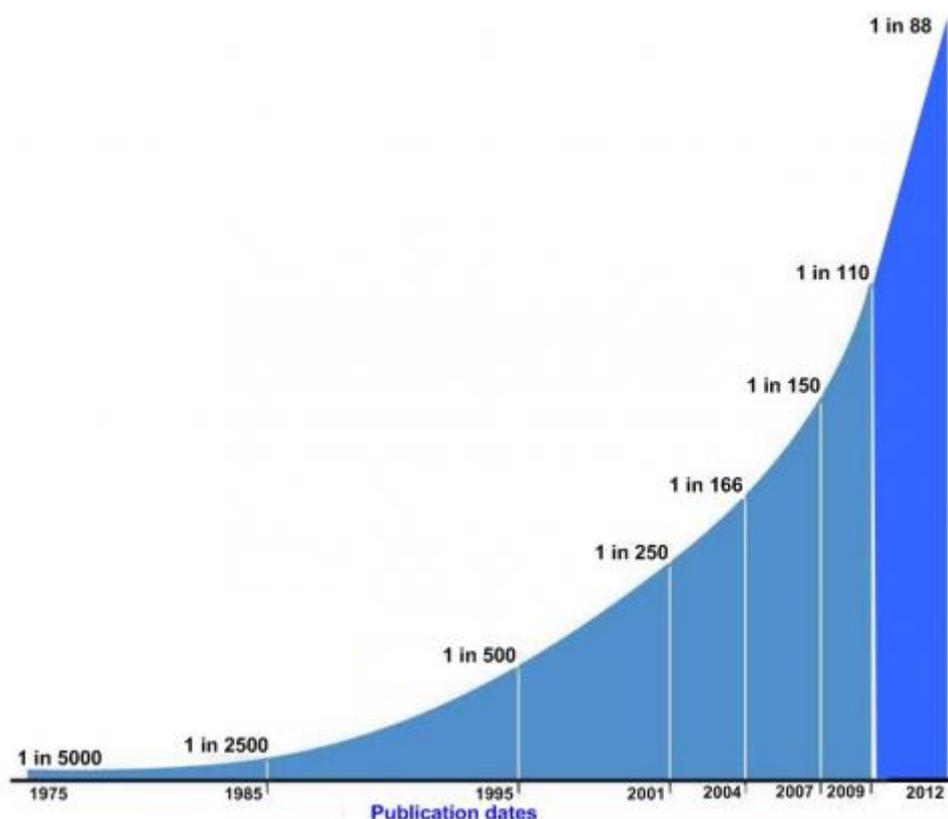


Figure 1.1: Prevalence of ASD affected child from 1975 to 2012 as reported from the Centers of Disease Control.

1.1.1 Symptoms of ASD

The new-latin word “autismus” was first used in 1910 by the Swiss psychiatrist Eugene Bleuler, while describing the symptoms of schizophrenia. He

derived this word from the Greek “αὐτός” (that means “self”) to describe a morbid self-admiration, referring to “autistic withdrawal of the patient to his fantasies, against which any influence from outside becomes an intolerable disturbance”. In the 40’s, Leo Kanner of the Johns Hopkins Hospital and Hans Asperger of the Vienna University hospital used this word for the first time in its modern sense: the first was describing 11 children, whose conditions were distinguishable from mental retardation on the basis of their social isolation; the second described similar patients affected by an “autistic psychopathy”.

To make an ASD diagnosis, the symptoms have to appear in the first three years of life. Parents usually notice the signs in the first 18 months of life, but 25%-40% of children with ASD show near-normal development till 24 months and then regress to a form generally indistinguishable from early-onset autism (Werner and Dawson, 2005). Even if the criteria to define autism were refined since their formulation by Kanner, the symptoms can still be included in three core categories (Folstein and Rosen-Sheidley, 2001; Lam, 2007):

1. **Abnormalities in social development.** During the first year of life, autistic children fail to show attachment to their primary caregivers. Later they show attachment, but little interest in other children. Adolescents and adults have awkward social interaction, like avoiding eye contact, awkward greetings and lack of social small talks. During conversations they usually insist in topics of little interest to others and don’t know when to stop speaking;
2. **Abnormalities of language development.** In one-third of children with autism, speech never develops past the occasional word, and most of these children do not attempt to use other means of communicating, such as gesture or eye contact. The remaining two-thirds develop a varying amount of speech, from a few stereotyped phrases to entirely normal in structure, although not pragmatic (i.e. modulating speech on the basis of the context and connecting it to non-verbal communication). Autistic people use speech to communicate needs or to provide information, but rarely to exchange information, chat or socialize. Their speech has a monotone quality, without the natural intonation that aids in communication. Because language delay is not specific of ASD and is present in other unrelated diseases, this trait is no more used in the new

criteria to define ASD in DSM-5 and this core domain has been merged with the previous one in a social-communication domain (Lauritsen, 2013);

3. **Rigid and repetitive behaviors.** Autistic children display repetitive movements like hand clapping, head rolling or body rocking (**stereotypy**). Sometimes these movements can also injure the person, like in the case of eye poking, skin picking, hand biting and head banging (**self injury**). They appear to follow rules, such as arranging objects in a certain order in a stack or line (**compulsive behavior**). The children also become upset at being interrupted (**sameness**) and they prefer an unvaried daily routine (**ritualistic behavior**). Also, they have overly focused interests, such as playing with a specific toy during childhood or studying bus or train timetables, baseball statistics or weather forecasts during adult age (**restricted behavior**).

Other features that are associated to ASD, but don't belong to the core symptoms, have been reported. In fact, children with ASD can present mental retardation (in less than 50% of cases) and epileptic seizures (about 15%).

While all the ASDs present the social impairments and the repetitive behaviors, Asperger syndrome affected children usually use language in an age-appropriate manner and are not mentally retarded. Individuals with PDD-NOS, instead, show marked impairments in each core domain, but do not meet diagnostic criteria for autism. Unfortunately, diagnosis is based solely on behavior, as currently there are no biological markers that can be used to simplify diagnosis. Several diagnostic instruments are available. Two are commonly used in autism research: the Autism Diagnostic Interview-Revised (ADI-R) is a semi-structured parent interview, and the Autism Diagnostic Observation Schedule (ADOS) uses observation and interaction with the child. The Childhood Autism Rating Scale (CARS) is used widely in clinical environments to assess severity of autism based on observation of children (Volkmar et al., 2005)

1.1.2 Causes of ASD

When Leo Kenner described Autism in 1943 little was known about the causes. For decades psychiatrists thought that autism was partly due to a social reticence observed in parents. The mothers were labeled as “refrigerator mothers” even though social reticence was more common in fathers. The symptoms of an

autistic child were then seen as the result of a withdrawal in him/herself, preferring a secret fantastic world to reality (Nolen-Hoeksema et al., 2009).

It took 20 years before researchers understood the importance of genetics in the aetiology of autism. The first lines of evidence are provided by the higher risk for a child to be diagnosed with ASD if a sibling is affected: a recent estimation addresses this in a 25-times higher probability than in normal population. Second, siblings and parents of affected children are more likely to show subtle cognitive or behavioral features similar to ASD-affected people. Third, independent twin studies indicated that the concordance rate for monozygotic twins is around 70-90%, while the one for dizygotic twins is 0-10% (Abrahams and Geschwind, 2008). Moreover, a number of other genetic syndromes manifest ASD at higher than expected frequencies compared to the general population: these syndromes account for more than 10% of all ASD cases and include tuberous sclerosis, fragile X, Down, neurofibromatosis, Angelman, Prader-Willi, Williams and Duchenne (Zafeiriou et al., 2013). Overall, these results indicate that autism is a genetic disease to such an extent that now it is accepted as the most genetic of all the neuropsychiatric syndromes. The causes of ASD then have to be searched in genes, even if some of them have to be attributed to the environment.

Albeit the genetic component in the aetiology of ASD is clear, this disease cannot be connected to a single gene mutation. On the contrary, up to 100 genes and chromosomal abnormalities have been linked to it. Different theories have been postulated: some scientists have proposed a model in which up to ten loci are involved, with at least three of them epistatic (interacting); others instead estimated three to six loci, all epistatic. Even if these are just theories, what is clear is that the genetics of autism can't be explained with a simple model (Folstein and Rosen-Sheidley, 2001).

In the last years a big emphasis has been placed on genetic research on ASD and several studies have been then conducted in order to find all the genes associated to autism, also with the hope that knowing which processes are affected can shed some light in the complicate pathology of autism. Although many genes have been associated to autism (table 1.1 for a quick overview), none of them can account for more than the 1-2% of all the ASD cases. There are three main groups of causes in syndromic autism (Abrahams and Geschwind, 2008) (fig. 1.2):

1. **Known genetic conditions** such as Fragile X or tuberous sclerosis: a substantial proportion of individuals with these conditions show comorbid ASD;
2. **Cytogenetic abnormalities** can be found in 6%-7% of children with ASD: several have been found associated to ASD (nearly on each chromosome). One of the most common is a duplication on chromosome 15q11–13, where the gene encoding for the GABA receptor is located. Other regions that are implicated in the ASDs by chromosomal abnormalities in multiple patients include 22q13, that contains SHANK3 (a synaptic adaptor protein), 5p15, 17p11 and Xp22, that contains NLGN4X;
3. **De novo copy number variations or single nucleotide polymorphisms (SNP)** have been reported in 2–10% of children with ASD: among these important are the mutations found in NRXN1 (neurexin-1), NRXN3, NLGN3 (neuroligin-3) and NLGN4 genes, that are proteins involved in synaptic function. Mutations have also been found in the CNTNAP2 (contactin associated protein-like 2) gene, that is a member of the Neurexin superfamily. Another important gene associated with autism is RELN (Reelin).

When considering the genetics of ASD another important aspect has to be taken into account: even though there has been a strong presumption that the three cores of symptoms are strongly intertwined and proceed from a common cause at the genetic, cognitive and neural levels, this may not always be the case. There is evidence that each defect composing this disease can have a different independent aetiology. This would explain not only the heterogeneity of ASD, but also the high correlation of this disease with other neuropsychiatric syndromes (Happè and Ronald, 2008).

Finally, as already stated, environmental factors have a role, even if secondary, in the causes of ASD. Recently, media has given a considerable attention to these risk factors, starting from the assumption that the incidence of autism is growing: even if the latter consideration was true, it has to be taken into consideration that this is due to the broadening of the diagnostic criteria, that lead to the inclusion in the ASD population of children that before would have had a different diagnosis. Much of the available information on the environmental risk factors reports association between autism and pre-natal conditions, like maternal hypothyroidism,

Gene	Syndrome or mutation(s)	Replicated association	Analysis of variant	Mouse model	Other evidence	Total score
<i>Promising</i>						
AVPR1A	0	0	0	1	0	1
DISC1	0	0	0	1	0	1
ITGB3	0	1	0	0	0	1
AHI1	2	0	0	0	0	2
EN2	0	1	0	1	0	2
GRIK2	0	1	0	0	1; homozygous mutation results in non-syndromic mental retardation	2
NRXN1	2	0	0	0	0	2
SLC25A12	0	1	0	0	1; associated with neurite outgrowth, expression is upregulated in ASD brain	2
<i>Probable</i>						
CACNA1C	2	0	1	0	0	3
CNTNAP2	2	1	0	0	0	3
MET	0	1	1	0	1; expression reduced in brains of cases versus controls	3
OXTR	0	1	0	1	1; expression reduced in blood of cases versus controls	3
SHANK3	2	0	0	0	1; modulates glutamate-dependent reconfiguration of dendritic spines	3
SLC6A4	0	1	1	0	1; clinical benefit from inhibitors, variation linked to gray-matter volume	3
CADPS2	2	0	1	1	0	4
DHCR7	2	0	1	0	1; hypocholesterolaemia in a proportion of probands	4
FMR1	2	0	1	1	0	4
NLGN3	2	0	1	1	0	4
NLGN4X	2	0	1	1	0	4
PTEN	2	0	0	1	1; mutations result in abnormal structure and function of the synapse	4
TSC2	2	0	1	0	1; regulates dendrite morphology and function of glutamatergic synapses	4
GABRB3	2	1	0	1	1; expression is dysregulated in pervasive developmental disorders	5
MECP2	2	0	1	1	1; MECP2 deficiency causes reduced expression of UBE3A and GABRB3	5
TSC1	2	0	1	1	1; regulates dendrite morphology and function of glutamatergic synapses	5
UBE3A	2	0	1	1	1; expression is dysregulated in pervasive developmental disorders	5
RELN	2	1	1	1	1; levels reduced in brains of cases versus controls	6

Table 1.1: Genes associated with ASD. A score was assigned to express how likely is the association and genes were then sorted in promising or probable, depending on the total score. AHI1, Abelson helper integration site 1; AVPR1A, arginine vasopressin receptor 1A; CACNA1C, calcium channel voltage-dependent L type alpha 1C subunit; CADPS2, Ca²⁺-dependent activator protein for secretion 2; CNTNAP2, contactin associated protein like 2; DHCR7, 7-dehydrocholesterol reductase; DISC1, disrupted in schizophrenia 1; EN2, engrailed homeobox 2; FMR1, fragile X mental retardation 1; GABRB3, gamma-aminobutyric acid (GABA) A receptor beta 3; GRIK2, glutamate receptor ionotropic kainate 2 precursor; ITGB3, integrin beta 3; MECP2, methyl CpG binding protein 2; MET, met proto-oncogene; NLGN3, neuroligin 3; NLGN4X, neuroligin 4 X-linked; NRXN1, Neurexin 1; OXTR, oxytocin receptor; PTEN, phosphatase and Tenascin homologue; RELN, Reelin; SHANK3, SH3 and multiple ankyrin repeat domains 3; SLC25A12, solute carrier family 25 (mitochondrial carrier, Aralar) member 12; SLC6A4, solute carrier family 6 (neurotransmitter transporter, serotonin) member 4; TSC1, tuberous sclerosis 1; TSC2, tuberous sclerosis 2; UBE3A, ubiquitin protein ligase E3A (Abrahams and Geschwind, 2008).

maternal use of thalidomide or valproic acid, maternal use of alcohol and congenital cytomegalovirus, but the most compelling association is the causal link between congenital rubella syndrome and ASD. Also the measles-mumps-rubella (MMR) vaccine has been proposed as a risk factor, but the available data didn't confirm the association. Moreover, a study on children born in the UK from 1979 to 1992 didn't reveal any increase in ASD occurrence after the introduction of this vaccine in 1988. However, no single environmental factor has been shown to be a major contributor

and, given also their small predicted contribution, it is likely that they are only a “second hit” in individuals with genetic susceptibility (Folstein and Rosen-Sheidley, 2001).

1.1.3 Neuropathology of ASD

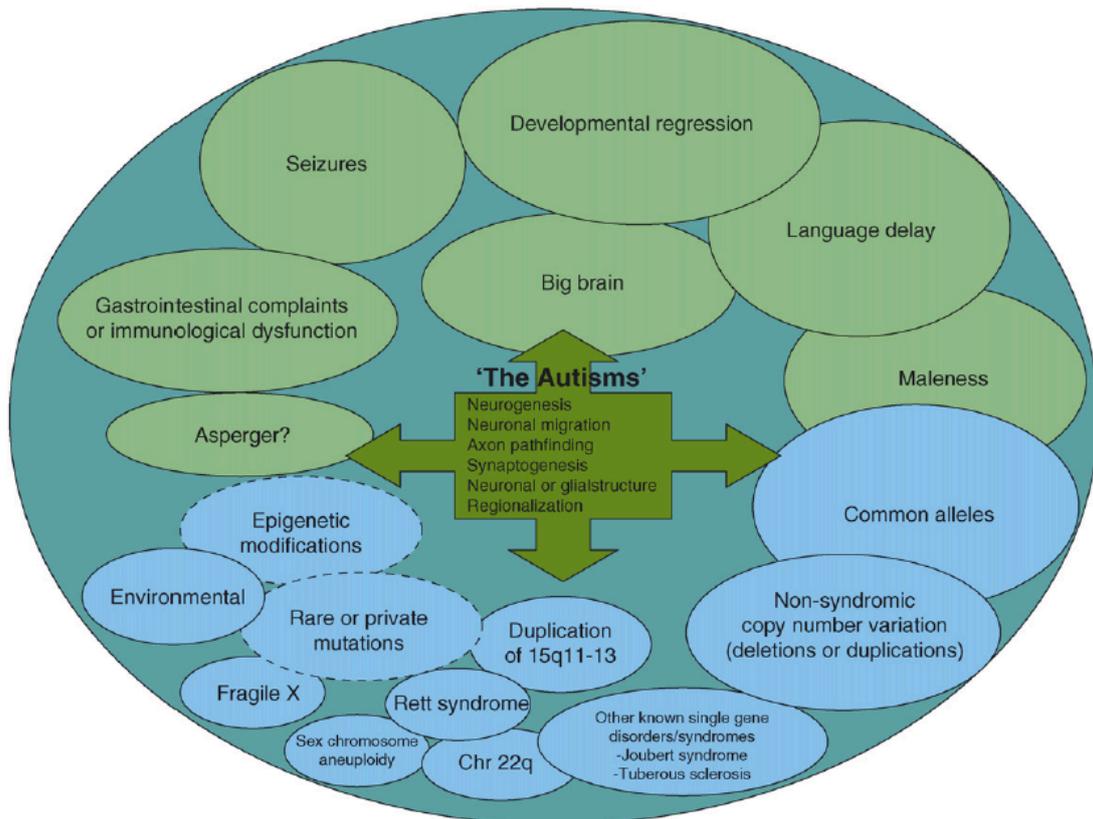
Several differences have been noted in autistic patients’ brains that, even if based on small numbers, indicate subtle, but widespread abnormalities throughout the cortex. MRI in young children with ASD (ages 18 months to 4 years) shows a 5%-10% abnormal enlargement in total brain volume, increasing the size of both the grey and white matter. Even if also the parietal and temporal lobes seem affected, the most consistent increase has been reported in the frontal lobes. In these regions the space between the columns is reduced, a result that, coupled with the higher neuronal density, led to the hypothesis that there may be a greater number of neurons in the autistic cortex. Also the cerebellar region has been found to be bigger in autistic patients in an extent proportional to the brain volume. Finally, the amygdala shows an enlargement of 15% in 8-to-12 years old boys with autism, but doesn’t look affected in older subjects, maybe due to the normal growth of this brain region between 8 and 18 years old in normal developing persons. Surprisingly, both in cerebellum and amygdala the neuron number seems to be lower in autistic patients compared to controls. Nevertheless, even if all these data show an abnormal growth of the brain, some studies revealed that what is disturbed in autistic patients is the time course of brain development rather than the final condition: the brain seems to undergo a period of precocious growth during early postnatal life followed by a deceleration in age-related growth (Amaral et al., 2008).

So far, as suggested by the genetic findings, four major pathways seem to be involved in ASD pathology (Persico and Bourgeron, 2006; Geschwind and Levitt, 2007) (fig. 1.2):

1. **Disturbed neuronal migration** during the first phases of brain development, that can interfere with the normal development of the brain;
2. **Unbalanced excitatory to inhibitory synapses ratio**: normal levels of inhibition are crucial for cortex development and insufficient inhibition has been proposed to underlie deficits in cognitive process and motor control, in addition to an increased susceptibility to epilepsy (Rubenstein and Merzenich, 2003);

3. **Axon pathfinding;**
4. **Synapse formation, maintenance and pruning,** as well as neuron-glia interaction, in which structural proteins like the Neuroligins' and Neurexins' family members can be involved.

One of the most prominent hypotheses to explain the pathobiology of ASD suggests that all these aberrant processes lead to an abnormal development of brain's connections. In typical non-affected children the increase in brain size is due to a complicated and time-regulated set of events that lead to a growth of gray matter followed by pruning of synaptic structures and myelination. In affected children, instead, different sets of malfunctions, either prenatally or postnatally, lead to an increased local and a decreased long-distance connectivity between different areas of the brain, especially those involving the frontal lobe. This reduced cortical–cortical reciprocal activity and coupling impairs the fundamental frontal function of integrating information from widespread and diverse systems (emotional, language, sensory, etc...) that are required for a big subset of higher skills, thus favoring processes that require local connectivity, like visual imagination (Courchesne and Pierce, 2005; Geschwind and Levitt, 2007; Frith, 2004). For example, a study published in 2006 showed that ASD-affected children were recruiting the visuospatial areas of the brain for the comprehension of sentences, even when imagery was not required, while their not affected counterpart presented activation of the same areas only in the sentences that required high imagination (Kana et al., 2006). Another example comes from a study that scanned using functional MRI the brains of volunteers while they were reading a passage and answering some questions: affected patients were using more prominently Wernicke's area instead of Broca's area that is activated in the controls, a result that led the author to conclude that volunteers with ASD were concentrating on single words rather than on the sentence as a whole (Just et al., 2004). Due to the different subset of deficits that two affected children might have and to the big role that experience (and then the environment) have even on a normal developing brain, this developmental disconnection theory would explain why autism form a spectrum of diseases, with a broad range of different symptoms and developmental courses (fig. 1.2). Moreover, it would explain the high rate of other neurogenic syndromes that are often associated with ASD (Geschwind and Levitt, 2007).



Current Opinion in Neurobiology

Figure 1.2: Scheme that summarizes some of the many possible etiological features (genetic and environmental heterogeneity; blue circles) and the clinical or syndromic heterogeneity (green circles) within autism spectrum disorder, in addition to the myriad of neurodevelopmental processes that might be disrupted to cause disconnection (central arrowed box). Broken lines emphasize areas where the extent of the contribution to the disorder remains to be defined relative to the contributions made by other factors (Geschwind and Levitt 2007).

1.2 CASPR2

1.2.1 The CNTNAP2 gene and ASD

With the emerging knowledge on the genetic basis of ASD, several studies have observed that CNTNAP2 is one of the strongest candidate genes associated to this disease.

In a 2006 study on an Amish family affected by CDFE (Cortical Dysplasia-Focal Epilepsy) syndrome Strauss et al. found a single-base deletion at nucleotide 3709 of the CNTNAP2 gene in all the affected individuals. This frameshift mutation generates a premature STOP codon, truncating the protein before the trans-membrane domain, leading to a non-functional product. CDFE is a rare neuronal migration disorder resulting in epileptic seizures, language regression and, in 67% of cases,

ASD. Jackman et al. in 2009 reported the same mutation in an Amish girl affected by cortical dysplasia, focal epilepsy and autism with language regression.

In 2008, three other separate studies reported a correlation between the CNTNAP2 gene and ASD, supporting the strong role of this gene in this pathology. Arking and colleagues performed a linkage and association two-stage whole genome scan on 72 families with at least two affected children, finding a significant signal in region 7q35, identifying a common polymorphism in the 5' end of CNTNAP2 gene. Bakkaloglu et al., instead, based on the identification of a chromosomal inversion on chromosome 7 that interrupts CNTNAP2 between exons 10 and 13 in a child with autistic features, resequenced the gene in 635 affected individuals and 942 uncharacterized controls: they found 13 rare variants present in patients with autism, but not in controls, of which 8 were predicted to be deleterious or were found in conserved regions. Finally, Alarcòn et al. performed a study using 2758 SNPs across 10Mb of the chromosome 7q35, finding a contribution of CNTNAP2 to the increase of the age of the first word that reveals a language delay in children carrying a mutation. Moreover, they noted an enrichment of its mRNA in brain structures known to be involved in language, such as the frontotemporal cortex during fetal development.

The study of Alarcòn and colleagues also lead to the hypothesis that CNTNAP2 can have a specific role in the development of language in the etiology of ASD. This idea is also supported by a study that confirmed these results using a different marker of language deficits, such as non-sense word repetitions, in which the authors excluded from the analysis any subject with autistic spectrum disorders (Vernes et al., 2008). Other evidence comes from the significant percentage of non-autistic siblings of affected people that have a history of speech and language delay (Bolton et al., 1994). Moreover, in-situ hybridization experiments performed in brains of humans and rodents revealed marked differences in the expression of this gene in the anterior temporal and prefrontal regions of brain, where expression of CNTNAP2 is high in humans and low or absent in rodents, reflecting the differences in the development and connection of these areas among these species, consistent with the higher language skills of humans (Abrahams et al., 2007). On the contrary, an expression analysis conducted on Zebra finch, a songbird that shares vocal learning behaviors with humans, revealed an enrichment of CNTNAP2 in cortical regions of the brain critical for vocal learning (Condro and White, 2013; Panaitof et al. 2010).

Together these data are compatible with the hypothesis that different components of ASD are independent and different genes can influence only some of them (Happé et al., 2008), with CNTNAP2 being responsible of the speech and language delays (Peñagarikano et al., 2012).

Finally, in an association study of 2010, subjects bearing variation in CNTNAP2 gene presented a diminished long-range anterior-posterior connectivity as opposed to individuals with non-risk alleles; on the contrary they showed an increased local frontal connectivity. The normal brain maturation, instead, is characterized by the weakening of short-range and strengthening of long-range cortical connectivity (Dosenbach et al., 2010). This relationship between CNTNAP2 genotype and functional connectivity was independent of autism diagnosis, indicating that this gene modulates a continuum of normal brain functions (Scott-Van Zeeland et al., 2010). Altogether these data are in agreement with the underconnectivity model proposed for autism (Geschwind and Levitt, 2007).

1.2.2 Caspr2 is involved in other diseases

Mutations in the CNTNAP2 gene are found in diseases other than ASD. As already discussed, a frameshift mutation has been found associated with CDFE in an Amish family (Strauss et al., 2006). Other works confirmed this data (Jackman et al., 2009) and demonstrated an association between this gene and schizophrenia (Clarke et al., 2012).

Moreover, a complex insertion/translocation involving chromosome 2 and 7, disrupting the CNTNAP2 gene, has been associated with Gilles de la Tourette syndrome (GTS) (Verkerk et al., 2003). GTS is an inherited neuropsychiatric disorder characterized by repeated involuntary motor and vocal tics. It is noteworthy that this syndrome is also often accompanied with obsessive-compulsive disorder (OCD), ASD and attention deficit hyperactivity disorder (ADHD) (Clarke et al., 2012); the latter has also been associated to rare variants of this gene (Elia et al., 2010). It is also interesting that also NRXN1, a member of the Neurexin superfamily as CNTNAP2, is found to be associated with GTS (Clarke et al., 2012).

Mutations in NRXN1 and CNTNAP2 have also been reported in patients with severe mental retardation with seizures and breathing anomalies, resembling Pitt-Hopkins syndrome (Zweier et al., 2009). Although mental retardation is often

reported in children affected by ASD (<50%), its presence doesn't belong to the three core symptoms (Geschwind and Levitt, 2007).

1.2.3 Caspr2, a member of the Neurexin superfamily

The CNTNAP2 gene is unusually large, spanning approximately 2.3 Mb on chromosome 7q35.1-q35.2; currently it is the largest known gene of human chromosome 7. It encodes Caspr2, a member of the Neurexin (NRXN) superfamily (see also par. 1.3.2). Interestingly, while the other NRXN genes have at least two promoters and multiple splice variants, currently there are not alternative forms of CNTNAP2 known (Nakabayashi and Scherer, 2001). Poliak and colleagues (1999) observed two bands in a Northern Blot analysis from rat brain and described two bands in a Western Blot analysis from the same tissue, with the lightest disappearing in the first postnatal week. The significance of this second form is currently unknown. An overview of the exons that compose this gene is provided in table 1.2.

Exon-Intron Structure of the CNTNAP2 Gene

Exon ^a	Position (nt)	Size (bp)	Intron-exon boundary ^b	Exon-intron boundary ^b	Element
1	1 .. 237	237		TCCACGTCCCgtaagtagcc	Initiation codon (141 .. 143)
2	238 .. 348	111	ttgttttcagAAAAATGTGA	AAGAGAGGAGgtaagccaaa	
3	349 .. 542	194	ccatcttcagGTGCTGGGGG	GAATATCTGGgtaagtcatt	
4	543 .. 690	148	ccttttccagGCATTCCCG	TGTTCTTACTgtgagtatcg	
5	691 .. 894	204	ttgttttagGGGCTGATGT	TTAAACTTAGgtgtgttctg	
6	895 .. 1079	185	gtgtacgcagGAAGCAACCA	GGACTATGAGgtacatgtga	
7	1080 .. 1223	144	tttctcaaagATAACCTTTG	CTCAAATGTGgtaaggattt	
8	1224 .. 1488	265	tattttacagGGAATTTGA	ATTTCTCAGgtcagtgaaa	
9	1489 .. 1638	150	tacttaccagGTTCTGGGTT	TTTTTTGGAGgtaagaatgc	
10	1639 .. 1810	172	tgtttcacagGTTTCTGAA	TCATAGACAGgtaaatgatc	
11	1811 .. 1917	107	tctctgacagATGTGTGCC	TGCCCAACTgtgagtgcca	
12	1918 .. 2037	120	tttgttctagCTATCTACGA	AACATGACAGgtaactgtgt	
13	2038 .. 2238	201	ctccccacagAGGACAAAGT	AACACCCAGgtaggctgag	
14	2239 .. 2395	157	tttctttagATGGAAGCCC	ACAAGCAATGgtgagtgcc	
15	2396 .. 2523	128	tgtgatccagGAGGAAGGAT	CAAGGAGACAgtaagtttgc	
16	2524 .. 2694	171	cttccccacagGAATTTATG	GAGCTGAAGTgtgagtataa	
17	2695 .. 2913	219	gctcctccagCTGCCACAGA	TTATTTGTGGgtaagtaatg	
18	2914 .. 3150	237	tcattcccagGTGGTGTCTGG	TGCAACAAAGgtaagggtgga	
19	3151 .. 3387	237	tttataacagATGTTGGTGC	AAACCCACTGgtaaggacaa	
20	3388 .. 3521	134	tcttctatagGAAGCTTACA	CTTCTCAAGgtaacatac	
21	3522 .. 3615	94	cctcctgcagCTCGATCATT	AAAGTTATAGgtaagaatgt	
22	3616 .. 3855	240	tttcttttagAAACAGGGAA	CTGGATTCAGgtaagtcct	
23	3856 .. 3936	81	ttctctacagCCAGTGCAGA	ATCATTGGAGgtaggtgatg	
24	3937 .. 7802	3866	tctccgtcagGCGTCATTGC	GAATTCAGTaaagccagcca	Stop codon (4134 .. 4136)
25	7803 .. 8107	305	caagaaacagTCTTAGATT		Poly(A) signal (8076 .. 8081)

Table 1.2: Exon-Intron structure of the CNTNAP2 gene as published by Nakabayashi and Scherer in 2001.

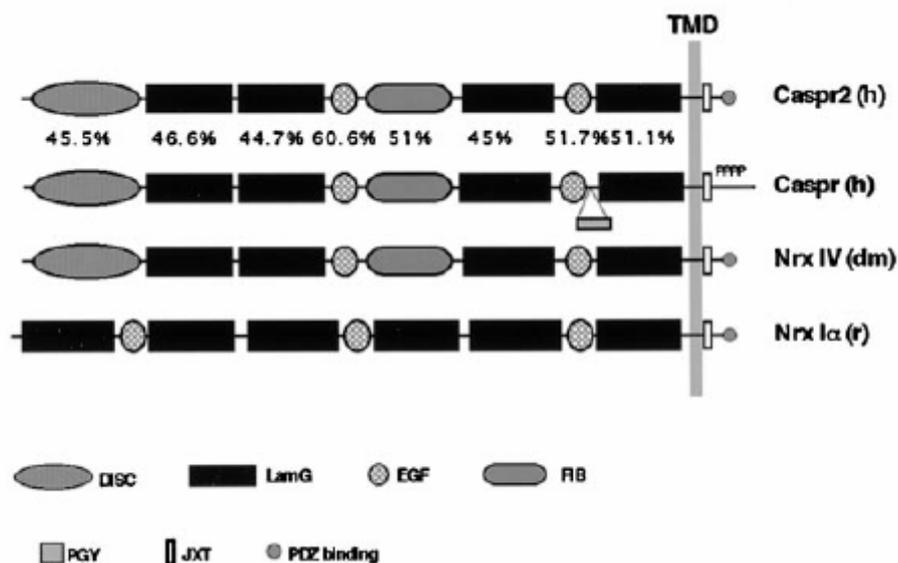


Figure 1.3: Domain organizations of Caspr2, Caspr1, Drosophila NrX-IV and rat α NRXN1. h: human; dm: drosophila melanogaster; r: rat; DISC: discoidin-like domain; LamG: laminin-G//LNS domain; FIB: fibrinogen similar region; PGY: A proline, glycine, tyrosin repeat; JXT: a shared juxtamembrane sequence, also responsible for protein 4.1 binding; PDZ: a PDZ containing protein binding domain (Poliak et al., 1999).

Caspr2 was identified in 1999 by Poliak and colleagues while searching for Caspr1 homologs in an ESTs database. The Caspr2 open reading frame encodes for 1331 amino-acids, with a predicted molecular weight of 130 kDa. The extracellular region of the protein contains 12 potential N-linked glycosylation sites, probably explaining the apparent molecular weight of 180 kDa that a Western Blot analysis gives. This region also contains 36 cysteines, likely forming 18 disulfide bonds, with a total of 8 independently folded domains.

The primary sequence of Caspr2 shares 45% of identity with Caspr1; moreover, the protein was described as the mammalian homolog of Drosophila Neurexin IV (NrX-IV), with which it shares 34% of identity. The extracellular region of Caspr2 is a mosaic of domains, including discoidin/neuropilin and fibrinogen-like domains, two EGF domains and four domains similar to a region in laminin A, referred to as the LamG/LNS domain. The structural organization of Caspr2 is very similar to both Caspr1 and NrX-IV, all having the hallmarks of type I trans-membrane proteins. In fig. 1.3 a comparison between the domain organizations of these proteins is reported. As shown, the domain organization among Caspr1, Caspr2 and NrX-IV is very similar, although the last two lack the PGY repeat found near the trans-

membrane domain of Caspr1. But the biggest difference is found in the cytoplasmic domain: Caspr2 and Nrxx-IV, like all the neuroligins, have a short amino-acid sequence at their C-terminus that serves as a binding site for type II PDZ domains; this consensus is not present in Caspr1. Caspr1 and Caspr2 share instead the juxtamembrane motif, responsible for binding to protein 4.1 (Poliak et al., 1999).

1.2.4 Role of Caspr2 in myelinated axons

Myelinating Schwann cells and oligodendrocytes ensheath the axon in segments separated by the nodes of Ranvier. This arrangement allows saltatory propagation of action potentials from node to node, improving their efficiency and rapidity. Myelinated axons appear thus organized in distinct domains (fig 1.4 and 1.5): the node, the paranode, the juxtaparanode and the internode, each with a different function and a different set of proteins. The formation and maintenance of these is driven by glia. In fact, the accumulation of their specific proteins appears during myelination and disperses following demyelination (Peles and Salzer, 2000).

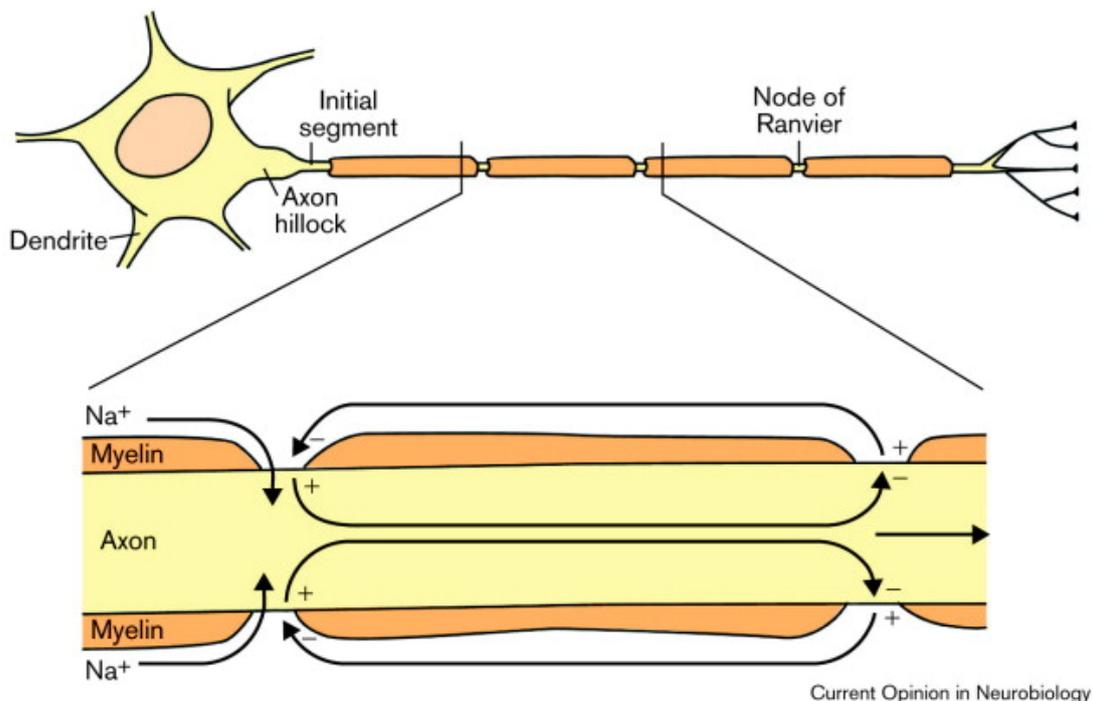


Figure 1.4: Schematic representation of action potential propagation in axon. Ensheathing glial cells promote saltatory propagation, providing high resistance and low capacitance, thus allowing the current flowing only at the nodes of Ranvier, where the action potential can be then regenerated (Peles and Salzer, 2000).

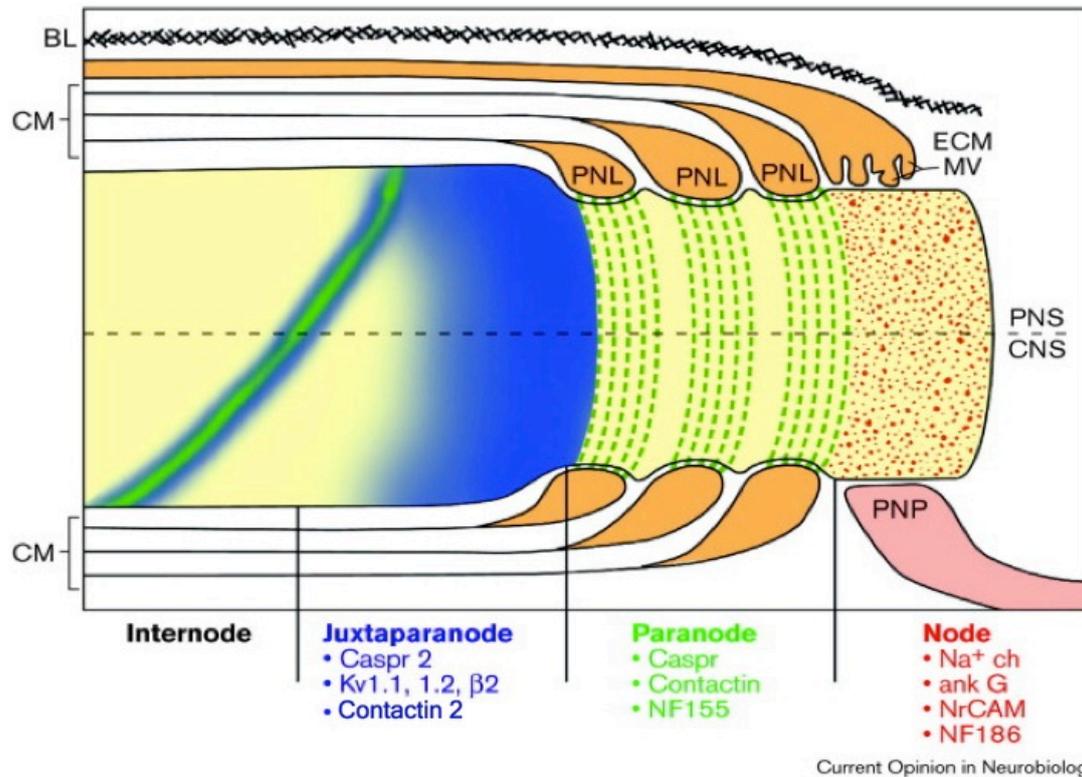


Figure 1.5: Longitudinal domains of myelinated axons in PNS and CNS. BL: basal lamina; CM: compact myelin sheath; PNL: paranodal cytoplasmic loops; MV: microvilli; PNP: perinodal astroglial processes (Peles and Salzer, 2000).

The node of Ranvier is devoid of ensheathing glial cells and enriched in voltage-gated Na⁺ channels, that are responsible of inward current flow and, thus, the regeneration of the action potential (Peles and Salzer, 2000).

The paranode is the region next to the node; its function is to form the septate junction to anchor the myelin loops to the axon and to form a diffusion barrier to limit lateral diffusion of membrane components, confining the Na⁺ channels to the node and the K⁺ channels to the juxtaparanodes. One of the components of these junctions is Caspr1, that binds contactin 1 (CNTN1), a glycosyl-phosphatidil-inositol (GPI)-anchored protein that belongs to the immunoglobulin superfamily (Peles and Salzer, 2000). CNTN1 is necessary to recruit Caspr1 to the plasma membrane and for its targeting to the paranodes: in fact, Caspr1 alone can't be delivered to the plasma membrane in CHO cells if not co-expressed with CNTN1 (Faivre-Sarraillh et al., 2000). This complex interacts then with the 155 kDa isoform of Neurofascin (NF155), another immunoglobulin like protein expressed by glial cells and

concentrated in the paranodal loops. The Caspr1/CNTN1/NF155 complex is necessary for the formation of the septa (Charles et al., 2002).

The juxtaparanodal region is just under the compact myelin sheath, beyond the paranodal junction and is characterized by the presence of heteromultimers of shaker like K⁺ channel (Kv1.1 and Kv1.2, with their Kvβ2 subunit). Two main functions have been proposed for these channels: they can act as a direct current stabilizer to maintain resting potential (Chiu et al., 1999) and/or they can act as an active dumper for reentrant excitation (Vabnick et al., 1999). Although for the first a spatial localization is not necessary, for the second the enrichment at the juxtaparanodal locations can play an important role (Poliak et al., 2003). Caspr2 accumulates at juxtaparanodes, as suggested by the fact that it colocalizes with Kv1.2, and next to Caspr1. It also forms a complex with the K⁺ channels, probably through a PDZ domain containing protein bound to the intracytoplasmic C-terminal portion: in fact, an agarose-bound Caspr2 can pull down Kv1.2 from rat brain lysates, but not when it lacks the last aminoacid. On the contrary, the C-terminal peptide of Kv1.2, but not a peptide lacking the last valine, which is known to be essential for the association of this channel with PDZ domain containing proteins, pulls down Caspr2 from the same lysates (Poliak et al., 1999). Moreover, a Caspr2 null mouse shows a decreased accumulation of Kv1.2 at the juxtaparanodes, without affecting the total internodal concentration of the channel (Poliak et al., 2003). Together these data suggest a role for Caspr2 in the accumulation of K⁺ channels at the juxtaparanodes, through a non-direct interaction.

As stated before, the formation of the domains on a myelinated axon is driven by glia. Even though at the paranodes this involves the Caspr1/CNTN1/NF155 complex, a different group of proteins should be involved at juxtaparanodes. The idea that emerged is that this complex is formed by Caspr2, that indirectly recruits also the K⁺ channel, and TAG-1 (Transient Axonal Glycoprotein-1, also known as CNTN2), a protein closely related to CNTN1, with which it shares 50% of homology and the fact that they are both GPI-anchored proteins. This idea emerged by the evidence that TAG-1 is enriched at the juxtaparanodes (Traka et al., 2002) and by the similarities of these two putative complexes (Caspr1/CNTN1, Caspr2/CNTN2). TAG-1 is expressed in many neuronal subpopulations as a GPI-anchored protein as well as a released form and is implicated in neurogenesis, neurite extension and fasciculation (Furley et al., 1990). In adult mice it is expressed by both myelinating glial cells and neurons, in

both at the juxtaparanodal level (Traka et al., 2002). TAG-1 null mice display behavioral deficits, such as impaired memory and learning as well as sensorimotor gating and gait defects (Savvaki et al., 2008). Its amino acid sequence reveals six immunoglobulin C2 domains and four domains homologous to the type III repeats of fibronectin (FNIII) (Furley et al., 1990, see also fig. 1.6).

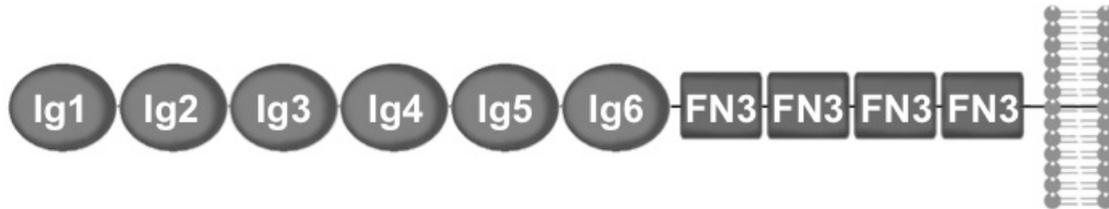


Figure 1.6: Domain organization of TAG-1. Ig: Immunoglobulin domain; FN3: fibronectin type III domain.

In 2003, two studies published on the same issue of the *Journal of Cell Biology* (Poliak et al., 2003; Traka et al., 2003) tried to prove the binding between Caspr2 and TAG-1. They reported that Caspr2 null mice lose the enrichment in K^+ channels and TAG-1 at juxtaparanodes both in CNS and in PNS, but the total content of the proteins between the nodes remains unaffected; Na^+ channels at the nodes and Caspr1 at the paranodes are instead unaffected. On the other hand, TAG-1 null mice show a similar displacement of K^+ channels and Caspr2. Interestingly no other morphological or functional deficit was reported in the myelinated fibers from these mice; also the nerve conduction seemed unaffected. Moreover, co-IP experiments on mice brain showed that an antibody (Ab) specific for Caspr2 can precipitate both TAG-1 and Kv1.2 and an Ab for TAG-1 can precipitate Caspr2 and Kv1.2 (fig. 1.7A-B and 1.8A-B). Co-IP on transfected cells (HEK or COS) showed similar results, but only when the two proteins were co-transfected: in fact no association between Caspr2 and TAG-1 can be detected mixing two independently transfected cell lysates (fig. 1.7C and 1.8C-D). Finally, a soluble TAG-1-Fc fusion protein binds to cells expressing TAG-1 or both the proteins, but not to cells expressing only Caspr2. The authors interpreted all these data suggesting the model in fig. 1.9: K^+ channels bind Caspr2 through some PDZ domain-containing proteins; TAG-1 interacts in trans with a molecule of TAG-1 expressed by glial cells and in cis with Caspr2, thus providing the structural support to localize the complex (Poliak et al., 2003; Traka et al., 2003). To this date, however, the Caspr2-TAG-1 complex hasn't been yet characterized.

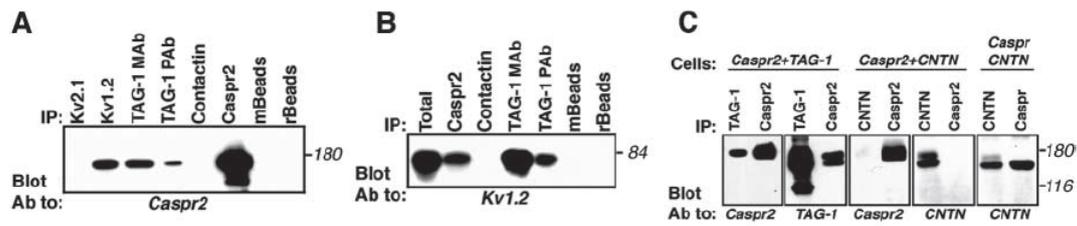


Figure 1.7: WB describing the association of Caspr2 to TAG-1 as published by Poliak et al. 2003. A and B: WB analysis on proteins immunoprecipitated from rat brain lysate; C: WB analysis on proteins immunoprecipitated from HEK cells co-transfected with the specified proteins. Caspr2 and TAG-1 seem to interact with Kv1.2 (A and B) and directly bind to each other (C), but not interact with the other tested proteins.

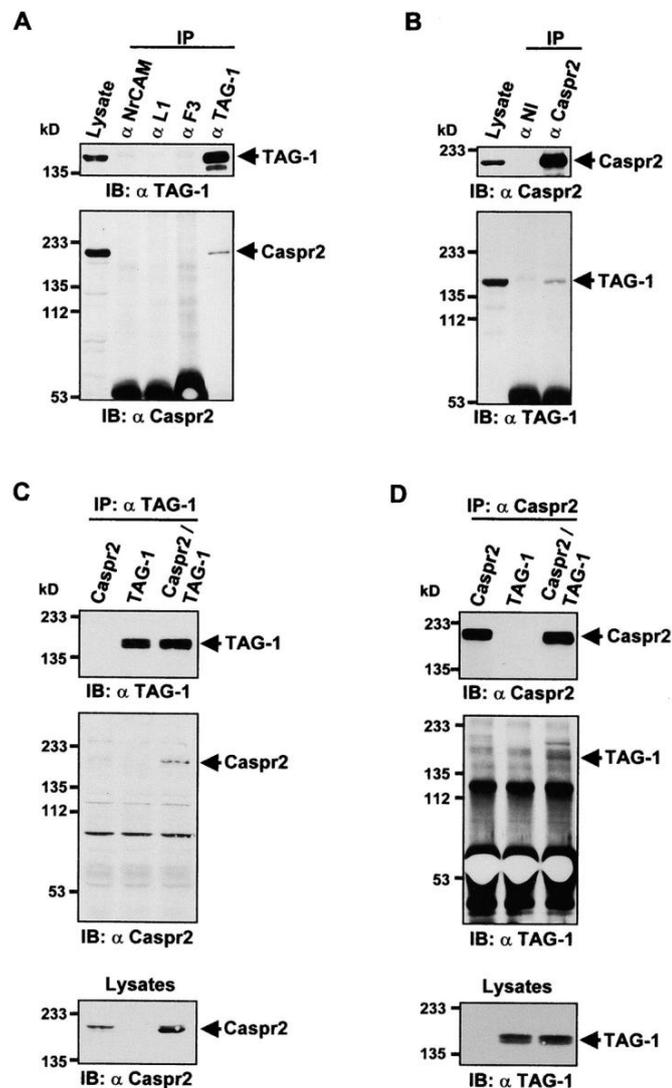


Figure 1.8: WB describing the association of Caspr2 to TAG-1 as published by Traka et al., 2003. A and B: WB analysis on proteins immunoprecipitated from rat brain lysate; C and D: WB analysis on proteins immunoprecipitated from HEK cells transfected with Caspr2, TAG-1 or both. As before Caspr2 and TAG-1 seem to directly bind to each other.

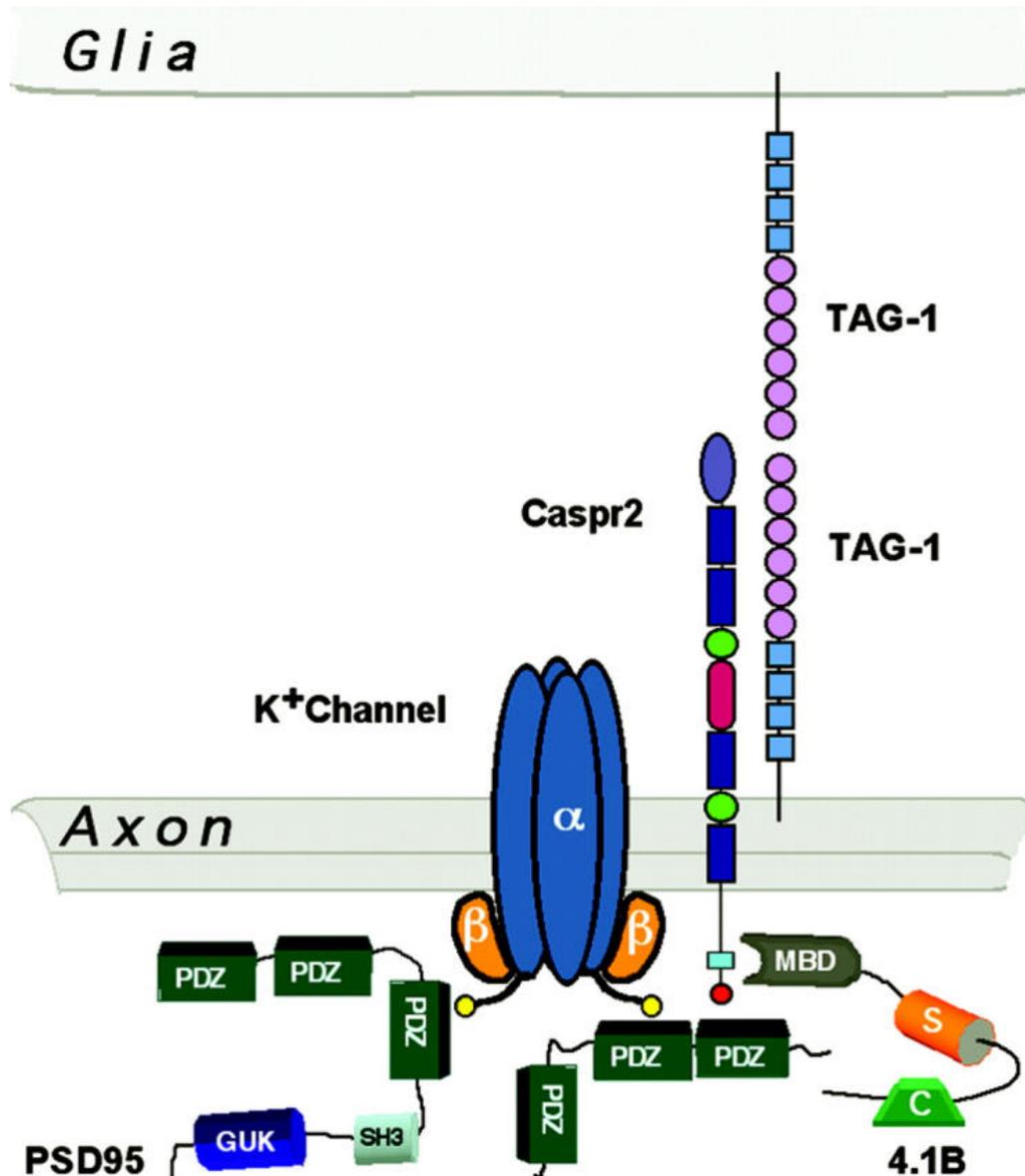


Figure 1.9: Schematic representation of the model proposed for K⁺ channel/Caspr2/Tag-1 complex at the juxtaparanode (Poliak et al. 2003).

Another important intracellular ligand of Caspr2 is protein 4.1B. The latter is enriched all along the axon, with the exception of the Node of Ranvier and can bind the intracellular domains of both Caspr1 and Caspr2. A recent study showed that protein 4.1B knockout mice present in the sciatic nerve a decreased and altered distribution of Caspr1 to the paranode and an almost complete loss of Caspr2, TAG-1, Kv1.1 and Kv1.2 at the juxtaparanodes. Once again, the total amount of the protein was unaffected, reflecting a lack of their enrichment at the right site and not a protein destabilization. Moreover, the paranode appeared non homogeneous and with less

clearly defined limits than in WT mice, that lead to the mislocalization of the juxtaparanodal K^+ channels to the paranode. The protein 4.1 family members are known to bind spectrin, a component of the intermediate filaments, so it can form a bridge between Caspr1 or Caspr2 and the cytoskeleton, revealing the importance of this protein in the stabilization of the paranodal and juxtaparanodal regions (Cifuentes-Diaz C et al. 2011).

1.2.5 Other Caspr2 localizations

Although the presence of Caspr2 at the juxtaparanodes has been well established, little is currently known on CASPR2 subcellular localization in the CNS (Poliak et al., 1999; Inda et al., 2006; Bakkaloglu et al., 2008; Ogawa et al., 2008; Bel et al., 2009; Zweier et al., 2009).

It has been reported that Caspr2 is enriched at the axon initial segment (AIS) of pyramidal cells (Inda et al., 2006; Ogawa et al., 2008). This region is innervated by chandelier's terminals and it is therefore devoid of myelin. Chandelier cells are a particular type of GABAergic interneuron and alteration in its connectivity has been associated with syndromes such as epilepsy and schizophrenia. In contrast to the axon, the AIS present only a partial segregation of Na^+ and K^+ channels (Inda et al., 2006). Interestingly, Caspr2 colocalizes with K^+ channels, but is not required for its localization at the AIS (Ogawa et al., 2008).

Some evidence is reported on the expression of Caspr2 at synaptic compartments. A subfractional analysis on rat lysates found Caspr2 and TAG-1 in the fraction containing synaptic plasma membrane (Bakkaloglu et al., 2008). Immunocytochemistry data showed a signal in cells bodies and dendrites in rat hippocampal neurons (Poliak et al., 1999). However, a 2009 study found a polarized expression of Caspr2 in these neurons, showing how Caspr2 is rapidly internalized in the somatodendric compartments, but not in the axon (Bel et al., 2009). Finally, a work in *Drosophila* points to a localization of NrX-IV (the orthologous protein of Caspr2) to the pre-synaptic compartment, a finding consistent with the association to the NRXN family of proteins (Zweier et al., 2009).

1.3 NEUREXINS AND NEUROLIGINS

1.3.1 Neurexins and Neuroligins in ASD

As ASD has been proposed to result from disconnections of brain regions, that disrupt a higher order association in favor of an increased local connectivity (Kana et al., 2006; Geshwind and Levitt, 2007) it is clear that synapse function and plasticity, playing a key role in neuron-neuron contact, have to be important in the development of these diseases. Among the synaptic proteins two classes have been proposed to be responsible for the function of the synapse, its specialization (excitatory or inhibitory) and the asymmetric differentiation of the branches of the two involved cells: these are the neurexins' (NRXN) and neuroligins' (NLGN) families of adhesion proteins. It is not therefore surprising that several studies found an association between ASD and these genes, in particular NRXN1, NLGN3, NLGN4 and NLGN4Y (Südhof, 2008).

Regarding the NRXN family, Feng et al. reported in 2006 two missense structural variants of NRXN1- β in ASD-affected subjects not present in healthy controls. Moreover, missense mutations or splice variants also of the isoform NRXN1- α were found in another case control study in 2008 (Yan et al., 2008). In those years several other point mutations, translocation events and large-scale deletions in the NRXN1 gene were detected in patients with autism (Szatmari et al., 2007; Kim et al., 2008; Zahir et al., 2008; Marshall et al., 2008).

The first evidence linking a mutation in a member of the NLGN family to ASD was reported in 2003 by Jamain et al. Starting from the evidence of de novo rearrangements in Xp22.3 they identified the transcripts of NLGN4. Noting that NLGN3, another member of the family, maps to Xq13.1, a region also known to be rearranged in autistic subjects, they screened for mutations of these two genes in 36 affected siblings pairs and 122 trios. They found in each gene a frameshift mutation segregating in two Swedish families. In the following years, several other studies reported mutations in NLGN3, 4 and 4Y (Laumonnier et al., 2004; Talebizadeh et al., 2006; Yan et al., 2005; Lawson-Yuen et al., 2008; Zhang et al., 2009), strengthening the idea that this class of proteins can be involved in the pathogenesis of ASD.

Other insights linking these proteins to ASD come from experiments on animal models. In some studies, in fact, autism-like phenotypes were observed in the KO animals for either NLGNs or NRXNs. Mice lacking NLGN1 showed impaired spatial memory, reduced Long Term Potentiation (LTP) in hippocampus and spent

more time in grooming than their wild type counterparts: the latter is likely to be the result of increased repetitive behaviors, that is one of the core symptoms of ASD (Blundell et al., 2010). On the other hand, mice lacking NRXN1 showed the same increased grooming phenotype seen in NLGN1 knock out mice (Etherton et al., 2009). When NLGN2 was deleted instead, the mice showed an increased anxiety-like behavior and decreased pain sensitivity (Blundell et al., 2009). Deletion of these genes leads also to impairments in sensory processing in honeybees and *C. elegans* and in long-term memory in *Aplysia* (Knight et al., 2011). Finally, SHANK3, another intracellular scaffolding protein that can bind to NLGNs through PSD95 and GKAP, has been found associated to ASD, further strengthening the evidence of the association of the NLGNs' family to this disease (Südhof, 2008).

Another aspect that makes the association of these proteins with ASD important is the localization of NLGN3, NLGN4 and NLGN4Y at the sexual chromosomes. This can, at least in part, explain why ASD has a higher incidence in male rather than in female subjects.

Like Caspr2, these genes also have been found associated to other diseases, like Tourette's syndrome (Lawson-Yuen et al., 2008), mental retardation (Laumonnier et al., 2004) and schizophrenia (Kirov et al., 2008). This is not surprising because of the important role that they exert in the synaptic junction: at this level a similar molecular alteration can produce different changes depending on the circuit involved, leading then to different neurological symptoms, which can be classified as distinct cognitive diseases. Consistent with this, two family members carrying the same mutation can present different cognitive disorders (Südhof, 2008).

1.3.2 Neurexins and Neuroligins: two synaptic adhesion proteins

Neurexins (NRXNs) are a class of type I membrane proteins, firstly identified as a receptor for α -latrotoxin, a component of the venom of the black-widow spider (Ushkaryov et al., 1992). They can be classified in two types: α - and β -NRXNs. The N-terminal extracellular portion of α -NRXNs presents six LNS domains (a repeat of Laminin, NRXN, Sex-hormone-binding globulin domains), intercalated by three EGF (epidermal growth factor) like domains, whereas β -NRXNs have only a short specific sequence and the last LNS domain. The two types of proteins have identical stalk domains, rich in O-glycosylated residues, transmembrane and the intracellular C-terminal domains, that present a PDZ binding motif (fig. 1.10) (Südhof, 2008).

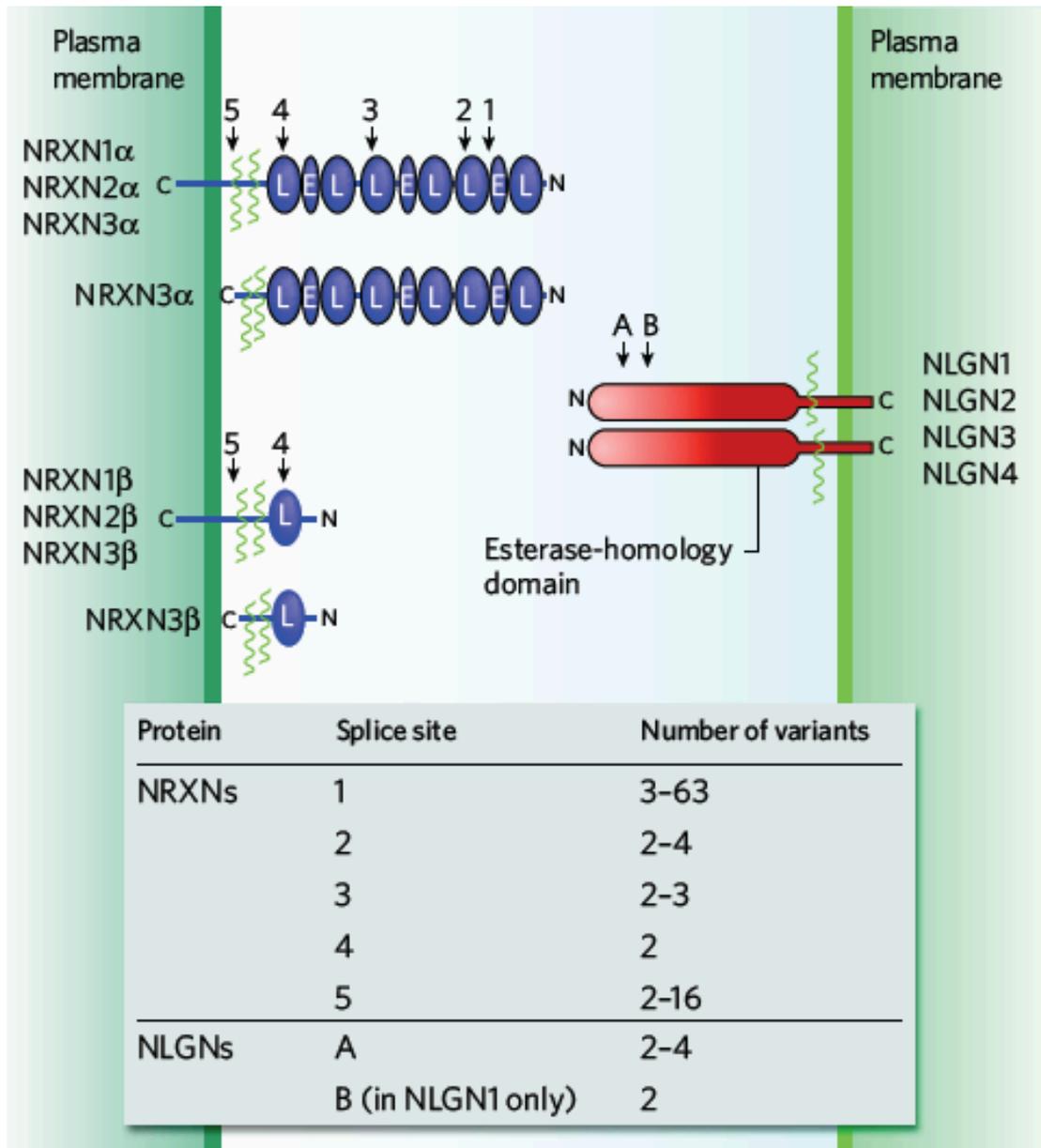


Figure 1.10: Domain organization and splicing sites of NRXNs and NLGNs. In the table are also reported the number of different variants available for each splicing site. L: LNS domains; E: EGF like domains (Südhof, 2008).

The human genome contains three NRXN genes, mapping at 2p16.3, 11q13 and 14q31 respectively. Neurexins are highly conserved in mammals, with more than 95% sequence identity between species (Missler and Südhof, 1998). Each gene codes an α - and a β -NRXN, starting from different promoters and all three are also extensively spliced during transcription: the α -NRXN genes each contain five splice sites (SS1 to SS5), but only the last two are shared with β -NRXNs (SS4 and SS5). These splice sites are used independently of each other and each one has several

alternative variants (fig. 1.10), generating virtually thousands of possible different proteins. Moreover, both the expression of the six NRXNs and the alternative splicing vary between different brain regions, raising the combinations available (Ullrich et al., 1995).

Immunofluorescence data show that NRXNs are located on pre-synaptic terminals, but it's not clear if they're confined there, because deletion of genes encoding α -NRXN has also post-synaptic effects (Südhof, 2008). Moreover, NRXN3 has been found outside the brain (Occhi et al., 2002) and one of its splice variants lacks the transmembrane domain and can be thus released (Missler and Südhof, 1998).

Neuroligins (NLGNs) are also a class of type I membrane proteins. They were identified due to their ability to bind NRXNs (Ichtchenko et al., 1995). The extracellular N-terminal part of NLGNs shows a single domain, that is homologous with acetylcholinesterases, but it lacks residues crucial for the active site, resulting thus in a disabled activity. A stalk domain, rich in O-glycosylated residues, a transmembrane and an intracellular C-terminal domain follow (fig. 1.10). Similar to NRXNs, NLGNs also have a type I PDZ binding motif in their intracellular domain (Südhof, 2008).

There are five genes encoding for NLGNs in the human genome, mapping at 3q26.31, 17p13.1, Xq13.1, Xp22.3 and Yq11.221 respectively: the latter, reported either as NLGN4Y or as NLGN5, complements NLGN4 on the Y chromosome, sharing with it a very high homology. NLGNs, as NRXNs, are highly conserved: for the same member of the family the sequence identity among mammals is 98%. NLGNs have only one splicing site (SSA) plus a second one present only in NLGN1 (SSB): both have multiple alternative variants, even if in a less amount than NRXNs' (fig. 1.10). Sequence comparison indicates that NLGN1, NLGN3 and NLGN4/4Y are more similar to each other than to NLGN2 (Südhof, 2008).

NLGN1 is exclusively localized in excitatory post-synaptic terminals, as it can colocalize with glutamate, but not GABA receptors (Song et al., 1999); the opposite result is observed with NLGN2, showing that it localizes at the inhibitory post-synaptic terminals (Varoqueaux et al., 2004). NLGN3 on the other hand colocalizes with NLGN1 and glutamatergic markers as it does with NLGN2 and GABAergic markers, showing thus localization both in excitatory and in inhibitory post-synaptic terminals (Budreck and Scheiffele, 2007). NLGN3 has been found also in many

classes of glial cells, particularly olfactory ensheathing glia, suggesting the possibility that it mediates glia-glia and glia-neuron interactions (Gilbert et al., 2001). NLGN4 accounts for 3% of NLGN proteins in mouse brain, thus its exact localization is yet unknown (Varoqueaux et al., 2006). A northern blot study detected NLGN4 also outside the brain, but its role there is still unknown (Bolliger et al., 2001).

1.3.3 The Neurexin-Neuroigin complex

The synapse is the asymmetric site of neuron-neuron contact that permits the transmission of a signal from one cell to the other. On the presynaptic side it is formed by an active zone for neurotransmitter release, a network of scaffolding proteins and a cluster of neurotransmitter containing vesicles; on the postsynaptic site, instead, there are scaffolding proteins and a cluster of neurotransmitter receptors placed on the membrane opposed to the active zone of release. Between the two membranes there is a 20 nm wide extracellular space that separates the 2 neurons, where several proteins are located. In this space the binding of Neurexins and Neuroigin occurs, forming an interaction layer, separated from the two plasma membranes by the glycosylated linker regions of these proteins.

NLGNs bind both α - and β -NRXNs with nanomolar affinities (Comoletti et al., 2006). Two molecules of NLGN dimerize through a four helix bundle, formed by two helices from each acetylcholinesterase-homolog domain and the connecting sequence between the globular lobes of the dimer and the cell membrane is elongated, projecting away from the dimer interface (Comoletti et al., 2007). Two molecules of NRXN bind, instead, to the opposite faces of the central dimerization site. The binding surface is then formed from the sixth LNS domain of α -NRXNs, that corresponds to the only LNS of the β -NRXNs, and the esterase homology domain of NLGN in a position opposite to the crippled acetylcholinesterase site (fig. 1.11). In each of the two binding interfaces two Ca^{2+} binding sites coordinated by residues from both proteins can be found, explaining thereby the calcium dependence of this association (Araç et al, 2007).

The splice site SSB of NLGN1 is included in the binding interface and SSA is close to SS4 of NRXN1. This observation can provide a structural reason for the effect of alternative splicing on the binding affinity of the complex. In fact, the affinities between NRXNs and NLGNs are variable and are controlled from the splice

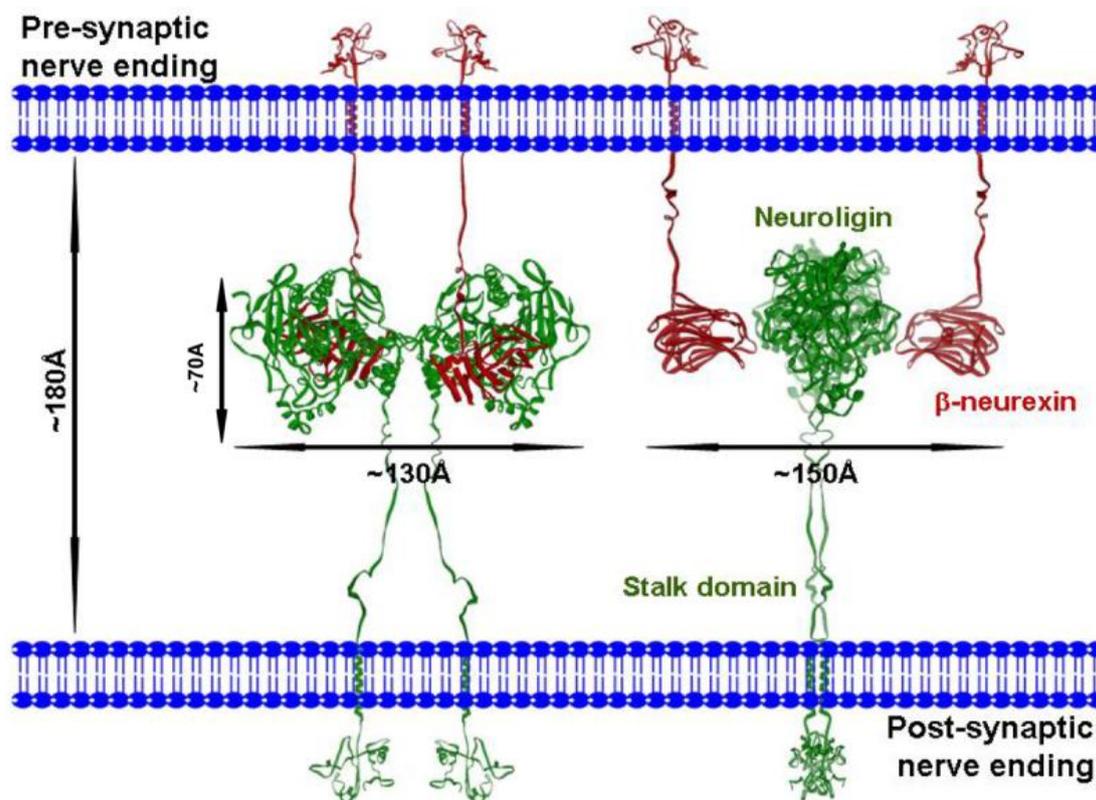


Figure 1.11: The NLGN1- β -NRXN1 complex seen in two orientations. Left: en-face, to illustrate the NLGN1 dimer. Right: 90° rotation to illustrate the binding (Comoletti et al., 2007).

variants that are included in the molecules: in this way, through the post-transcriptional event of splicing, a recognition code that can modulate synapse plasticity and development is generated (table 1.3). For example, the presence of SSB in NLGN1 restricts the binding of this protein only to β -NRXNs that lack SS4, but the lack of those eight residues leads to the binding of both α - and β -NRXNs, regardless of SS4 (Boucard et al., 2005). The lack of SSA in NLGN1, instead, leads to a decreased affinity for β -NRXN1 lacking SS4, but to a smaller extent. Moreover, the presence of SS4 of β -NRXN1 seems to alter the rank order of selectivity for different NLGNs: unlike NLGN1, in fact, NLGN2 and NLGN4 display a slight preference for β -NRXN1 that has SS4, rather than the one lacking it (Comoletti et al., 2006). Finally, of interest is the recent finding that the alternative splicing of NRXNs can be driven by synaptic activity: under depolarizing condition the mRNAs of NRXNs lacking SS4 increase, while the ones carrying that sequence are decreased, in a process that involves calcium/calmodulin-dependent kinase IV signaling and the protein SAM68.

In this way the activity of the neuron can shift the NRXNs' ligand preference from NLGN2 to NLGN1, thus affecting the synapse properties (Iijima et al., 2011).

	α -NRXN + SS4	α -NRXN Δ SS4	β -NRXN + SS4	β -NRXN Δ SS4
NLGN1A	–	+	+	+
NLGN1B	–	–	–	+
NLGN1AB	–	–	–	+
NLGN1(–)	–	+	+	+
NLGN2	+	+	+	+
NLGN3	–	+	+	+
NLGN4	–	+	+	+
Cbln	+	–	+	–
LRRTM	–	+	–	+

Table 1.3: Summary of the code of binding between different splice variants of NRXNs and NLGNs (Bang and Owczarek, 2013).

Recently, other proteins have been found to bind NRXN or NLGN. For example, leucine-rich repeat transmembrane (LRRTM) neuronal proteins bind β -NRXN1 lacking SS4 with a Kd of 20nM. Moreover, both NLGN and NRXN bind receptor like protein tyrosine phosphatases (PTPRT). As a result, other proteins can regulate synapse formation.

1.3.4 Biological role of the complex

The first evidence for the role of NLGN and NRXN in synapse came from the work of Scheiffele and colleagues in 2000. They showed that overexpressing either NLGN1 or NLGN2 proteins in non-neuronal cells was possible to trigger the formation of pre-synaptic structures in contacting neurons. This effect was specific, because an excess of soluble β -NRXN could oppose it. Moreover, using NLGN-coated beads instead of cells or artificially clustering NRXNs by specific antibodies the same results can be obtained, suggesting that this effect is due to neuroligins' potential of clustering neurexins (Dean et al., 2003). These findings were then complemented by the observation that non-neuronal cells expressing NRXNs proteins could induce the formation of both GABAergic and glutamatergic post-synaptic specialization in neurons. As for NLGN, NRXN-coated beads can exert the same effect (Graf et al., 2004). Finally, RNA interference-mediated down-regulation of

NLGN1, NLGN2 and NLGN3 in cultured neurons reduces the number of excitatory and inhibitory synapses (Chih et al., 2005); the same effect can be achieved using soluble β -NRXN to block the NLGN-NRXN association (Levinson et al., 2005).

Together these results led to the conclusion that the binding between NLGN and NRXN may induce synapse formation, but analysis of gene knock-out mice proved that this is not the case. In fact, mice lacking NLGN1, NLGN2 and NLGN3 die at birth, but they all have a normal synapse number. The respiratory failure that leads to death is instead a consequence of reduced GABAergic and glutamatergic synaptic function that leads to an abnormal network activity in brainstem centers that control respiration (Varoqueaux et al., 2006). The deletion of all three α -NRXNs leads to a reduction in the density of synapses and to the impairment of the machinery that regulates the release of synaptic vesicles (Missler et al., 2003). It is then reasonable to think that NLGN-NRXN binding is necessary for synapse maturation, instead of synapse formation, and that the results reported before reflect more the stabilization of otherwise transient synaptic contacts. The extracellular NLGN-NRXN binding triggers a set of signals both at the pre-synaptic and the post-synaptic level, activating functions specific for the different localizations and the different kind of synapse, that can also be modulated by the neuronal activity (Südhof, 2008).

These signals generated by the extracellular binding are bidirectional and pass through the intracellular domains of NLGNs and NRXNs, that can recruit intracellular proteins, like those containing PDZ regions (fig. 1.12). NLGNs and NRXNs bind different proteins, exerting different functions: in this way, their binding can trigger the formation of the asymmetric junction.

NRXNs bind protein 4.1 and the class II PDZ domain of CASK. The latter is a MAGUK (membrane-associated guanylate kinase) protein composed by two regions: the C-terminal contains PDZ, SH3 and guanylate-kinase domains, while the N-terminal has a Ca²⁺/calmodulin-dependent protein kinase (CAMK) that shows a Mg²⁺ independent activity. CASK can phosphorylate NRXN and bind to protein 4.1, VELI and MINT proteins, nucleating the assembly of actin on the cytoplasmic sequence of NRXN, that can in turn trap and stabilize clusters of synaptic vesicles (fig. 1.12). Moreover, CASK and MINT both contain multiple PDZ domains and can therefore recruit a larger scaffold of protein, with additional NRXNs, to form a large adhesive contact (Butz et al., 1998; Dean and Dresbach, 2006; Südhof, 2008). In this

way, CASK can translate the extracellular signal of NLGN-NRXN binding to an intracellular response, like the clustering of NRXNs and the modulation of actin.

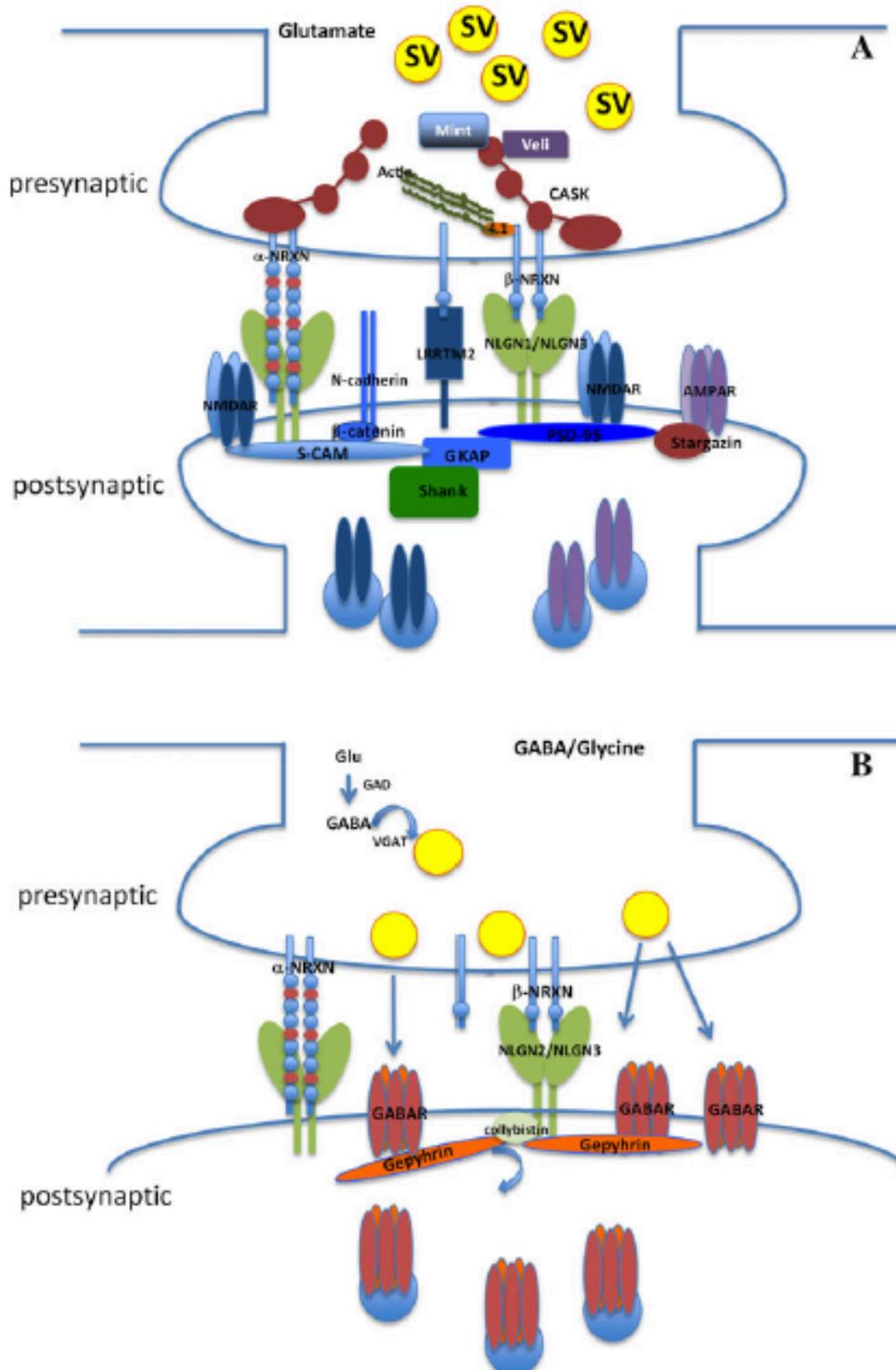


Figure 1.12: Schematic representation the asymmetric contact site generated by the NRXN:NLGN binding at the excitatory (A) or inhibitory (B) synapses (Bang and Owczarek, 2013).

In glutamatergic synapses, NLGN1 binds to class I PDZ domains such as those contained in PSD-95, a post-synaptic MAGUK protein. This protein contains different PDZ domains and each one is specialized for a distinct function, enabling PSD-95 to bind both NLGN1 and NMDA channels (Irie et al., 1997). Moreover, PSD-95 can bind to other proteins, especially to GKAP, which, in turn, binds to SHANK, bridging the recruitment of scaffolding proteins at the post-synaptic site, including the cytoskeleton (Sheng and Hoogenraad, 2007). PSD-95 has been proposed to have a role on the balance of excitatory and inhibitory synapses: its overexpression can in fact lead to a reduction of inhibitory synapses and an increase of the excitatory ones, while its suppression has the opposite effect (Prange et al., 2004; Dean and Dresbach, 2006). The maturation of glutamatergic synapses passes through a two-step process. The initial contact between axon and dendrite leads to the accumulation of NLGN1 and NMDA receptors on the post-synaptic side. As a second maturation step, the NLGN1-PSD-95 complex recruits AMPA receptors to the post-synaptic membrane (fig. 1.12A). This process requires synaptic activity: specifically the recruitment of GluR1-containing AMPA receptor requires synapse activity, whereas GluR2-containing AMPA receptors accumulate also in resting conditions. The importance of the synapse activity is also exhibited by the potential of NMDA co-agonist in rescuing the abnormal grooming behavior showed by NLGN1 null mice (Blundell et al., 2010). Finally, while GluR1 seems to be inserted on the synaptic membrane at the sites of NRXN-NLGN binding, GluR2 is trapped at that site through a mechanism dependent on NLGN1-PSD-95 interaction. The overexpression of NLGN3 in neurons, instead, enhances AMPA, but not NMDA receptor-mediated currents, suggesting that NLGN3 can also play a role in the unsilencing of the excitatory synapses (Bang and Owczarek, 2013).

In GABAergic synapses, NLGN2 is involved in the recruitment of gephyrin and collybistin that are two key components of the inhibitory postsynaptic scaffold: the first is critical for GABA receptors' localization, while the second, when activated, tethers gephyrin to the postsynaptic membrane (fig. 1.12B). Thus, NLGN2 can drive the clustering of GABA receptors (Bang and Owczarek, 2013). It has been shown that also the refinement of developing GABAergic synapses depends on activation of NMDA receptors and that this induces a down-regulation of the receptor itself (Gillespie et al., 2005).

NRXN isoforms can play a role in the specification of the synapses' properties: α -NRXNs seem to be primarily expressed by GABAergic synapses, whereas β -NRXNs localize also to the glutamatergic ones. Moreover, α -NRXNs and β -NRXNs with SS4 induce specifically the formation of GABAergic synapses (Chih et al. 2006). Fu and Huang (2010) proposed a model in which α -NRXNs, that can diffuse along the axon membrane, can function as an exploring mechanism, searching for post-synaptic partners. The first α -NRXN-NLGN weak binding can recruit CASK and the presynaptic release machineries. Synapse-specific β -NRXNs are then delivered to the presynaptic membrane, engage in high affinity binding with the corresponding NLGN and are further stabilized by the release of the correct neurotransmitter.

1.4 OTHER PROTEINS IDENTIFIED IN THE WORK

1.4.1 Secreted Frizzled Related Protein 1

Secreted Frizzled Related Proteins (SFRPs) is a family of five secreted proteins, widely expressed in human tissues. The protein structure (fig. 1.13) involves an N-terminal Cysteine-Rich Domain (CRD) and a C-terminal Netrin-Like domain. The CRD domain, usually called the Frizzled domain, is homologous to the extracellular portion of Frizzled (Fz) receptors, sharing with them about 50% of identity, including the ten key cysteines that are highly conserved residues. The Netrin-like domain, instead, is also present in other proteins, where it appears to bind heparin and function in the inhibitory activity of these molecules against extracellular matrix metalloproteinases (Jones and Jomary, 2002).



Figure 1.13: Protein structure of Secreted Frizzled Related Protein 1. Red line: leader peptide; Fri: Frizzled domain; C345C: Netrin domain, found also in complement proteins C3, C4, and C5.

The known mechanism of action of SFRPs is due to their CRD domain. The homology with Fz receptors can let them bind the ligands without provoking a signal cascade. Moreover, the CRD of SFRPs seems able to dimerize with the corresponding domain both in other SFRPs and Fz receptors. Therefore, because these receptors are involved in the Wnt signaling, SFRPs are involved in its modulation, having then a role in regulation of cell growth and differentiation in specific cell types. The proposed models of action are: 1) sequestering Wnts; 2) acting in a negative-dominant fashion through the formation of inactive complexes with Fz receptors that prevent signal activation; 3) favoring Wnt-Fz interactions by simultaneously binding to both molecules and promoting signal activation. Depending on their mechanism of action different SFRPs can either favor or oppose Wnt signaling, exerting opposite functions. For example, while SFRP1 blocks the Wnt mediated stabilization of β -catenin and promotes cell-death, SFRP2 promotes β -catenin accumulation and increases cell resistance to apoptosis (Bovolenta et al., 2008).

Another role exerted by SFRP1 that does not depend on its ability to bind the Wnt, but only the Fz receptor, is axon guidance. SFRP1 can directly modify and reorient the growth of retinal ganglion cell growth cones: it requires pertussis-toxin-sensitive activation of G α protein, it involves protein synthesis and degradation, and it is modulated by different levels of cAMP and cGMP (Rodriguez et al., 2005).

Other than its role performed by the binding to Wnt and Fz receptors, SFRPs appear to be promiscuous proteins that can interact with molecules that are unrelated to components of the Wnt signaling cascades and among which there is no apparent relationship. These include fibronectin, Unc5H3, receptor activator for nuclear factor kappa B (NF- κ B) ligand (RANKL) and bone morphogenetic protein (BMP)/Tolloid (Bovolenta et al., 2008).

Wnt proteins were first identified as mammary oncoproteins. Indeed, aberrant activation of canonical Wnt signaling occurs in a large proportion of tumors, and is associated with the loss of controlled growth and the impairment of cell differentiation. In line with a tumor suppressor function of SFRPs, loss or significant down-regulation of SFRP1 expression has been observed in a large proportion of invasive carcinomas (Bovolenta et al., 2008). Interestingly, deregulation of the Wnt pathway has also been posited as central to the pathogenesis of Alzheimer's disease (AD), potentially at the level of β -catenin or GSK-3 β activity through altered

interactions with the presenilin proteins (which are mutated in familial AD) (Jones and Jomary, 2002).

1.4.2 Clusterin

Clusterin (CLU, also known as Apolipoprotein J or APOJ) is a secreted protein expressed in almost all mammalian tissues but with higher expression level in brain than in many other tissues. The structure of CLU presents two large domains (fig. 1.14) that are cleaved during the maturation of the protein and remain bound by five disulfide bridges. The protein is highly glycosylated, with carbohydrate accounting for approximately 20-25% of the total mass of the mature molecule (Stewart et al., 2007). In general, both the structural and functional characteristics strongly indicate that the main function of CLU is to act as a chaperone molecule. CLU contains three amphipathic and two coiled-coil α -helices, which are typical structures for molecular chaperones. Interestingly, the secondary structure of CLU also contains three large intrinsic disordered regions, i.e. molten globule domains. Molten globule regions are flexible, protein–protein interacting domains present in many hub proteins (Nuutinen et al., 2009).



Figure 1.14: Protein structure of Clusterin. Red line: leader peptide; CLa: Clusterin Alpha chain; CLb: Clusterin Beta chain.

CLU is an ATP-independent type of chaperone, which with its molten globule domains confers broad substrate specificity. CLU does not refold the proteins, but stabilizes stressed proteins into a state that is suitable for refolding. Fig. 1.15 provides a summary of the known roles of CLU in Alzheimer's disease, in which it acts as a versatile pro-survival factor which has few, if any, harmful characteristics. In fact, in brain CLU is involved in clearance of Amyloid- β ($A\beta$) protein. It can either prevent $A\beta$ oligomerization or enhance the formation of fibrillar structures, depending on the ratio between clusterin and $A\beta$ peptides. Moreover, by binding LRP-2/megalin (low-

density lipoprotein receptor-related protein-2) it is involved in the clearance of the A β peptides through the blood–brain barrier (BBB). Finally, Clusterin can also inhibit the activation of the complement system. This is an important property since several studies have demonstrated that protein aggregates can activate the complement system and provoke inflammatory responses in AD (Nuutinen et al., 2009).

Other than its role in AD, CLU, as other apolipoproteins, is included in lipid particles that transport lipids, e.g. cholesterol and phospholipids, in body fluids. In this role its main binding partner is apolipoprotein A-I in HDL. The major difference between CLU and other apolipoproteins is that the former is present in lipid-poor particles (Nuutinen et al., 2009).

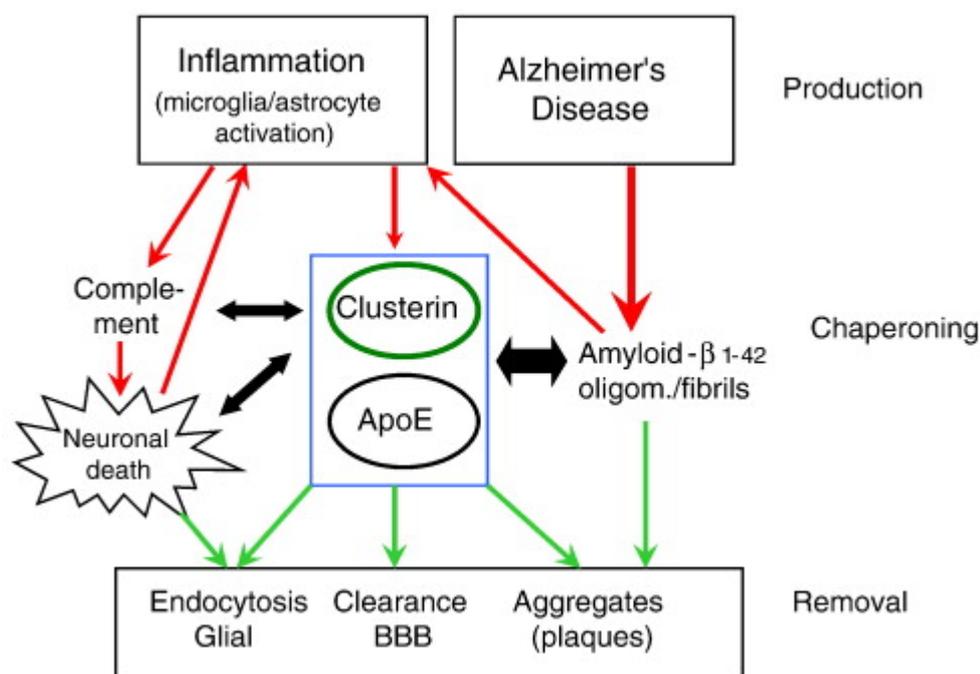


Figure 1.15: The functions of Clusterin and Apolipoprotein E chaperones in Alzheimer's disease.

1.4.3 Apolipoprotein E

Apolipoprotein E (APOE) is a secreted protein produced predominantly from the liver, where the bulk of plasma APOE originates, and from the brain, which has a different pool of this protein, because it cannot pass the blood–brain barrier. In the lipid-free state, the structure of APOE is similar to the one of CLU, with two large independently folded domains; the difference is that APOE is not cleaved and the two

halves are connected by a hinge region, rather than disulfide bonds. The N-terminal domain is an elongated four-helix bundle, while the C-terminal domain is highly α -helical, but its structure is unknown. The α -helices that form this protein are highly amphipathic, making the protein stable both in lipid-bound and lipid-free conformation. Moreover, full-length APOE forms multimeric complexes in aqueous solution. Thr194 is the only glycosylated residue and its glycosylation is not necessary for APOE expression (Hauser et al., 2011).

Functionally, APOE can associate with lipids: in this case the conformation of the protein changes from the lipid-free form, a transition that is necessary for the high affinity binding of target proteins like the LDL receptor, the VLDL receptor, the LDL receptor related protein 1 (LRP1) and the APOE receptor. By binding these receptors, APOE-containing lipoproteins are cleared from the plasma, thereby regulating plasma cholesterol levels. These processes can be very important in brain, as a lack of cholesterol delivery causes synaptic and dendritic spine degeneration and results in failed neurotransmission and decreased synaptic plasticity. Moreover, APOE is involved in the clearance of cholesterol and lipid debris resulting from neuronal cell damage, cell death, traumatic brain injury or terminal differentiation. Finally, APOE is involved in the clearance of Amyloid- β protein in Alzheimer's disease (Hauser et al., 2011, see also fig. 1.15).

ApoE is polymorphic and this influences its functional and structural properties. The three common allelic isoforms, apoE2, apoE3 and apoE4, differ at positions 112 and 158. ApoE3, the most common isoform, contains cysteine and arginine, respectively, whereas apoE2 has two cysteines and apoE4 has two arginines at these positions: these substitutions affect the position of some internal salt-bridges. ApoE3 and apoE4 bind to LDL receptors with similarly high affinity, but the binding of apoE2 is 50- to 100- times weaker. ApoE4, instead, is a major risk factor for Alzheimer's disease, accounting for 40–60% of the genetic variability observed in the disease. Two key properties of apoE4 – domain interaction and reduced stability relative to apoE2 and apoE3 – have been suggested to underlie the association of apoE4 with this disease (Hatters et al., 2006).

1.4.4 Tenascin R

Tenascin R (TNR) is a member of the Tenascin family, a group of four extracellular matrix glycoproteins, and is expressed exclusively in the nervous

system. TNR is highly conserved across the species, with 93% homology between the rat and human version. The protein structure (fig. 1.16) involves 4.5 epidermal growth factor-like (EGF) domains, 8 or 9 fibronectin type III repeats (FN III repeats), and a carboxyl-terminal region globular domain resembling the β and γ chains of fibrinogen (FG) (Carnemolla et al, 1996). TNR has two major molecular forms, TNR 160 (160 kDa) and TNR 180 (180 kDa), probably generated by an alternative splicing on the exon coding for FN6. Isoforms possibly have a role in the formation of dimers and trimers. The relative amounts of the 180- and 160-kDa TNR isoforms change during CNS development. TNR undergoes extensive post-translational modification with sulfated oligosaccharides, both N-linked and O-linked, accounting for 10–20% of weight. The glycosylation pattern of TNR is developmentally regulated and affects its adhesion, motility and migration (Woodworth et al., 2004).



Figure 1.16: Protein structure of Tenascin R. Red line: leader peptide; Green line: coiled coil region; EGF: EGF-like domain; FN3: Fibronectin type 3 domain; FBG: Fibrinogen-related domain.

TNR interacts with proteins on the surface of both glial and neuronal cells, as well as with extracellular matrix proteins. Among the cell membrane associated ligands of TNR there are CNTN1 and voltage-gated Na^+ channels; in the extracellular matrix proteins members of the lectican family of chondroitin sulfate proteoglycans (CSPGs) can be found, including aggrecan, versican, neurocan and, with highest affinity, brevican (Anlar and Gunel-Ozcan, 2012).

TNR has multiple, sometimes opposite roles: its versatility includes adhesive or antiadhesive, growth-promoting or inhibiting effects, depending on the type of targeted cells, receptors and signaling pathways, their location in the CNS, the composition of the surrounding ECM, and the time of interactions. EGF, FG, FN1–2 and FN3–5 domains serve for anti-adhesive properties while FN6–8 domains mediate adhesive properties. TNR interaction with CNTN1 has a repulsive effect. Depending on the developmental context, instead, lecticans can collaborate with TNR or inhibit its action. In general, TNR is growth-promoting when constituting a uniform

substrate, and inhibiting when forming a sharp boundary of the substrate. TNR deficiency in vitro results in reduced oligodendrocyte differentiation and delayed myelin basic protein expression. On the other hand, overexpression results in reduced migration in embryonic stem cells, which travel shorter distances and differentiate more. TNR also induces the generation of GABAergic neurons (Anlar and Gunel-Ozcan, 2012).

2. AIM OF THE RESEARCH

Autism Spectrum Disorder (ASD) is a common and heterogeneous childhood neurodevelopmental disorder characterized by widespread abnormalities of social interactions and communication and a markedly restricted repertoire of repetitive self-stimulatory behaviors. Learning disability (in about 70% of cases) and epilepsy (in about 30% of cases) are also frequently observed. Genetic studies on affected twins proved that, among cognitive diseases, ASDs are the most heritable (about 80%).

Although several subtle but widespread abnormalities throughout the cortex have been found in autistic patients, recent findings suggest a potential unifying model in which higher-order association areas of the brain that normally connect to the frontal lobe are partially disconnected during development (Geschwind and Levitt 2007). A key feature of ASDs is that they typically develop before two or three years of age, a time characterized by pruning of existing synapses and formation and maturation of new ones, suggesting that these can be some of the disrupted pathways that can lead to the brain disconnection.

Several mutations associated to ASD have been found in *CNTNAP2*, whose encoded protein is Caspr2, a member of the NRXNs' superfamily, as suggested by the primary sequence. This protein is primarily located at the juxtaparanodes of the myelinated neurons, where it favors the clustering of potassium channels Kv1 (Poliak et al., 1999). Two papers published in 2003 reported the direct binding of Caspr2 with CNTN2 (also known as TAG-1), a GPI-anchored protein expressed both by neurons and glial cells in the juxtaparanodal region (Poliak et al., 2003; Traka et al., 2003): the authors proposed that this association would help the accumulation of the Caspr2/Kv1 complex in the correct localization. So far, little is known on the structure of the

Caspr2/TAG-1 complex. Moreover, Caspr2 has been proposed to have other localizations in the brain, such as the axon initial segment (AIS) of pyramidal cells, a portion that is devoid of myelin (Inda et al., 2006; Ogawa et al., 2008), and the synaptic terminals, even if it's not clear if at the pre- or post- synaptic side (Bakkaloglu et al., 2008; Poliak et al., 1999; Zweier et al., 2009).

A consistent observation emerging from recent studies is the discovery of mutations in the genes encoding NRXN1, NLGN3 and NLGN4 (Südhof, 2008). NLGNs and NRXNs families of proteins are synaptic cell-adhesion molecules that connect presynaptic and postsynaptic neurons, also modulating the maturation of the synapse and specifying its function. NRXNs are type I membrane proteins localized at the pre-synaptic button: there are three genes encoding for NRXNs in the human genome, each one encoding a longer α - and a shorter β -NRXN. Also NLGNs are type I membrane proteins: there are four genes encoding for NLGNs in the genome. Although all localized at the post-synaptic side, different NLGNs are present in different kinds of synapse: NLGN1 on the excitatory ones, NLGN2 on the inhibitory and NLGN3 and 4 on both. Sequence comparison indicates that NLGN1, NLGN3 and NLGN4 are more similar to each other than to NLGN2 (Südhof, 2008). Moreover, the binding affinity of NLGN2 for NRXNs is significantly lower than the other NLGNs (Comoletti et al., 2006).

This study is aimed at the research of novel binding partners for Caspr2 and NLGN2. Both these proteins have known binding partners, but it's reasonable to think that they might have other ligands in the brain. In fact, Caspr2 has been found in the synapse, where there's no clustering of K⁺ channels, possibly exerting an uncharacterized function. On the other hand, NLGN2 shows several differences from the other family members. The lower affinity for NRXNs and the lower sequence similarity to the other members of the family can reflect the presence of additional specialized functions for this protein in the inhibitory synapse that could be exerted through the association with an unidentified protein. The identification of new binding partners would give new insight in the molecular biology of these proteins and of the ASD pathology. For this purpose, fishing experiments were performed: these consisted of an affinity chromatography on the protein extract of a brain, using either Caspr2 or NLGN2 bound to the stationary phase and a mobile

phase, coupled with a mass spectrometry-based identification of the retained proteins. The identified proteins were then ranked on the probability to bind the target and the most promising were tested by means of co-immunoprecipitation (co-IP) and Isothermal Titration Calorimetry (ITC) experiments.

Another goal of this work is to better characterize the binding of Caspr2 with TAG-1. Since 2003, the year in which the direct association of these proteins was proposed, there hasn't been any further in-depth analysis on this aspect; some preliminary experiments could open the way for further analysis on the structure of the complex. For this purpose, ITC experiments to measure the binding stoichiometry and affinity of the two proteins and co-IP experiments, using fractions of Caspr2 to map the region of binding, were performed.

3. MATERIALS AND METHODS

3.1 MOLECULAR AND CELL BIOLOGY

3.1.1 Eukaryotic Cell Culture

HEK-293 (Human Embryonic Kidney) cells were obtained from American Type Culture Collection (ATCC) and routinely grown as monolayers in Dulbecco's modified Eagle's medium (DMEM) with high glucose and 4mM L-Glutamine (Fisher Scientific) and supplemented with 10% FBS (Fisherbrand™ Research Grade Fetal Bovine Serum, Fisher Scientific). Cells were maintained in 10 cm dishes (Becton-Dickinson) in a humidified incubator (Forma™ Steri-cycle™, Fisher Scientific), at a 5% CO₂ atmosphere.

To create stable selected lines, HEK-293 cells lacking the enzymatic activity of N-acetylglucosaminyltransferase-1 (Gnt1⁻) were used. For the clonal selection (see also par. 3.1.9), 800 µg/ml G418 (Geneticin, Sigma) was added to the HEK-293 media.

To produce high amounts of proteins, cells were grown in 3-layer flasks (Nunc™ TripleFlasks, Thermo Scientific) using DMEM supplemented with different amount of FBS (ranging from 10% to 0%) and with a 100X penicillin and streptomycin mix (HyClone™, Fisher Scientific).

3.1.2 Bacterial culture

For the amplification of the plasmids (see also par. 3.1.3) *E. coli* DH5α strain [genotype: F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK⁻ mK⁺), λ⁻] was used.

The bacteria were grown in Luria-Bertani medium (10 g/l tryptone, 5 g/l Yeast Extract, 10 g/l NaCl) to which the appropriate antibiotic was added as needed. The final concentrations used are:

- Ampicillin: 100 µg / ml;
- Kanamycin: 40 µg / ml;
- Chloramphenicol: 20 µg / ml.

To plate bacteria on solid media, 1.5 % Agar was added to the LB. The media were treated at 120°C for 20 minutes and then cooled down before adding the antibiotic.

3.1.3 Plasmids used

The cDNAs encoding most of the proteins used in this work are cloned into a modified version of the pCMV6-XL4 vector. This was customized in Dr. Comoletti's lab and has the following features (Fig. 3.1):

- **Optimized Kozac sequence**, bearing the sequence gccRccATGG (where ATG is the first Met), recognized by the ribosome as the translational start site, for an enhanced translation efficiency;
- A sequence encoding the **Prolactin Leader Peptide (PLP)**, that is one of the most efficient translocation signals to direct the nascent protein through the Endoplasmic Reticulum (ER) and the Golgi to be either delivered to the plasma membrane or secreted (Hegde and Bernstein, 2006);
- A sequence encoding the **FLAG sequence** DYKDDDDK placed between the PLP sequence and the Multiple Cloning Site (MCS);
- Sequences encoding for a **3C protease cleavage site** (LEVLFQ/GP) and for the **Fc portion of the human IgG** placed downstream the MCS. The IgG coding sequence contains introns that help stabilize the mRNA before translation.

After cloning a cDNA in this vector, the produced protein would be secreted and would contain a FLAG tag at the N-terminal and a removable IgG tag at the C-terminal. In some cases a former version of this plasmid that doesn't have a FLAG sequence was used, producing a protein carrying only the IgG tag.

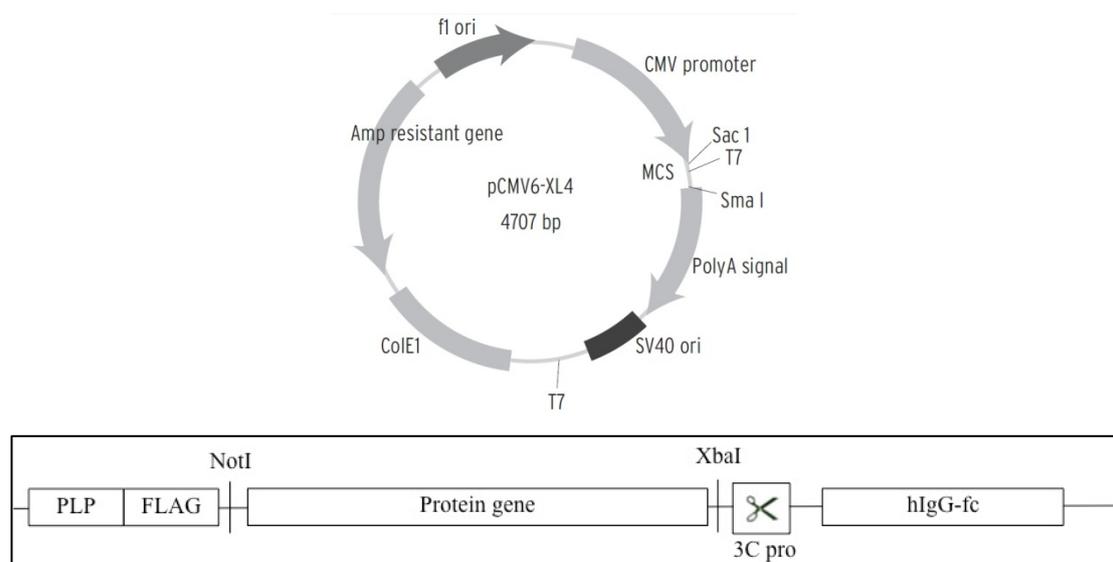


Figure 3.1: Scheme of the original pCMV-XL4 vector (top) and of the modified portion, after the cloning of a hypothetical cDNA between the sites NotI and XbaI (bottom).

Other than for the cloning of the cDNAs reported in par. 3.1.4 and 3.1.5, these vectors were used to express the following proteins:

- Caspr2-1261*- IgG;
- Caspr2-183*-IgG;
- TAG-1-1004*- IgG;
- FLAG- TAG-1-28-901- IgG;
- FLAG- TAG-1-28-461- IgG;
- FLAG- TAG-1-28-597- IgG;
- FLAG- TAG-1-600-1027- IgG;
- α -NRXN1-1309*-IgG;
- FLAG-NLGN2-616*;
- FLAG-NLGN1-639*- IgG;
- FLAG-NLGN1-639*-G500A;
- β NRXN1-300*- IgG;

As a control, the vector encoding only the IgG was used: this has been called IgC that stands for “IgG Control”.

For the expression of the full-length membrane proteins instead the pCDNA3.1 vector was used. Those proteins were:

- HA-Caspr2 wt;
- FLAG-TAG-1 wt;
- FLAG-NLGN1 wt
- FLAG-NLGN2 wt

As a control, an empty pCDNA3.1 vector was used.

3.1.4 cDNAs

The cDNAs encoding for CNTN1, SFRP1, CLU, APOE and TNR were purchased from Open Biosystems (Thermo Scientific). Their sequences can be found in the GenBank database with the following accession numbers:

- SFRP1: BC036503.1
- CLU: BC019588.2
- APOE: BC003557.1
- CNTN1: BC066864.1
- TNR: BC129830.1

The plasmids containing these cDNAs were extracted from the bacteria provided as described in par. 3.1.7.

3.1.5 Cloning

Portions of the coding sequences of the cDNAs reported in par. 3.1.4 were cloned in the pCMV6-XL4-FLAG-IgG or pCMV6-XL4-IgG (par. 3.1.3), as reported in the results section. To achieve this goal, a PCR was used to insert a NotI site at the 5' end of the region of interest and an XbaI site at the 3' end. The primers used were:

SFRP1 → FOR: AGTAAAGCGGCCGCAAGCGAGTACGACTACGT
REV: CGGCCTCGTCTAGACCCTTAAACTCGGACTGAAAGG

CLU → FOR: AGTAAAGCGGCCGCAAGCGAGTACGACTACGT
REV: CGGCCTCGTCTAGACCCTCCTCCCGGTGCTTT

APOE → FOR: AGTAAAGCGGCCGCAAGGTGGAGCAAGCGGT
REV: CGGCCTCGTCTAGATGATTGTCGCTGGGCACA

CNTN1 → FOR: AGTAAAGCGGCCGCAGACTTTACCTGGCACAG
 REV: CGGCCTCGTCTAGACCGAATTCCGAGTAGACAAGA

TNR → FOR: AGTAAAGCGGCCGCAGGCCAGACCTCAGAC
 REV: CGGCCTCGTCTAGACCGAACTGTAAGGACTGCC

The PCR reactions were conducted in a final volume of 50 µl using the above mentioned primers and the phusion high-fidelity DNA polymerase (New England Biolabs):

- 10 µl of reaction buffer (5x);
- 1.2 µl of forward primers (5 µM);
- 1.2 µl of reverse primers (5 µM);
- 1.5 µl of dNTPs (10 mM each);
- 1 µl of template DNA (5 ng/µl);
- 34.6 of µl H₂O;
- 0.5 µl of High fidelity TAQ polymerase;

Cloning PCR cycle			
Step	Temperature (°C)	Incubation time	Cycles
Initial denaturation	98	2 min	1
Denaturation	98	20 sec	2
Annealing	50	30 sec	
Elongation	72	30 sec per Kbp	
Denaturation	98	20 sec	25
Annealing	65	30 sec	
Elongation	72	30 sec per Kbp	
Final elongation	72	10 min	1

Table 3.1: Thermo cycle used for the cloning PCR

The amplified DNA and the vector were digested at 37°C for 1 hour, using NotI and XbaI restriction enzyme (New England Biolabs) in a final volume of 20 µl.

For each reaction:

- DNA → 1 µl for the vector, 10 µl for the PCR-amplified DNA (insert);
- 2 µl of NEbuffer 3.1 (10x);
- 1 µl of NotI enzyme;
- 1 µl of XbaI enzyme;
- H₂O up to 20 µl.

The digested DNAs were purified from a 1% agarose gel, using a NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer protocol. The purified DNAs were quantified on a 1% agarose gel and ligated using a T4 DNA ligase (New England Biolabs) in a final volume of 20 µl, incubating 30 minutes at room temperature and over-night at 4°C:

- DNA → 3:1 molar ratio insert:vector;
- 2 µl of T4 buffer (10x);
- 1 µl of T4 DNA ligase;
- H₂O up to 20 µl.

Chemically competent DH5α bacteria were transformed with 2 µl of ligation product (protocol described in par. 3.1.6).

3.1.6 Bacterial transformation

Chemically competent cells were prepared in the Comoletti laboratory and conserved at -80°C. Bacterial transformation was performed through a heat shock at 42°C for 45 seconds. The protocols used were described by Seidman et al. in “Current Protocols in Molecular Biology” (1997).

3.1.7 Plasmid DNA purification

The purification of plasmidic DNA was carried out using NucleoSpin® Plasmid Mini-prep kit (Macherey-Nagel) or with NucleoBond® Xtra Maxi-prep kit (Macherey-Nagel), according to the manufacturer protocol. The concentration of the purified DNA was tested using a Nano drop 2000c (Thermo Scientific).

3.1.8 HEK cells' transfection

70% confluent HEK cells were transfected with the calcium phosphate method as described by Kingston et al. in "Current Protocols in Molecular Biology" (1996), using 10 µg of each plasmidic DNA.

3.1.9 Selection of stable transfected cells

In order to achieve protein purification, HEK293 Gnt1⁻ cells were transfected with a pCMV6-XL4 vector encoding the selected protein (see also par. 3.1.3, 3.1.4, 3.1.5 and 3.1.8). The cells were also transfected with 3 µg of a pCDNA3.1 carrying the geneticin resistance gene (neo: aminoglycoside 3'-phosphotransferase). The cells were then selected by growth in 800 µg/ml G418-added media (see also par. 3.1.1). After ~20 days the cells usually form colony-like spots each representing a single clone. The clones were then selected and grown in a 24-well plate. After ~3-4 days the medium of each clone was tested by western blot (see also par. 3.2.1) to verify the level of expression of the protein. The clones with the higher level of expression were expanded on 10 cm plates.

3.1.10 Protein extraction from HEK cells

To collect the protein content of HEK cells the following protocol was used. The cells were harvested and centrifuged at 1200 g for 5 minutes. The pellet was resuspended in 1 ml of lysis buffer (150 mM NaCl + 50 mM Tris-HCl pH 8 + 1% triton X100) to which leupeptin 1µM (Sigma) was added. After 20 minutes of incubation on ice, the samples were centrifuged at max speed for 10 minutes and the supernatant was collected.

3.1.11 Protein purification

In this work several proteins were purified. These are naturally secreted proteins, GPI anchored proteins or single-pass trans-membrane proteins. In the last two cases the attention has been focused on the extracellular portion and, for this reason, the gene encoding the protein was cloned till few residues before the anchored residues or the trans-membrane domain. Therefore, the proteins to be purified were all released by the cells in which they were produced.

Stably transfected HEK 293 Gnt1⁻ cells (see also par. 3.1.9) were cultured in a 3-layer flask (see also par. 3.1.1). In the DMEM media used, a 100X pen-strep mix and different concentrations of FBS were added. The concentration of FBS used were:

- 10% to plate the cells and for the first 4 days;
- 5% for the following 2 days;
- 2% till when the desired amount of media was collected, replacing it every day.

The 2% FBS concentration was chosen to allow cell survival while reducing the amount of unwanted serum proteins that can lower the protein yield in the purification steps.

The media collected from the flasks was spun down to eliminate any detached cell and then a protease inhibitor cocktail (Benzamidine 4%, Bacitracine 1%, Na Azide 2%, EGTA 100 mM) was added for storage till the purification step.

The buffers used for the purification were:

TNE:	50 mM Tris-HCl pH 8.0	+
	150 mM NaCl	+
	1 mM EDTA	
TNED:	TNE	+
	1 mM DTT	
Washing buffer:	TNE	+
	450 mM NaCl	
HBS:	20 mM HEPES	+
	150 mM NaCl	

The collected media was passed through a chromatography column by gravity flow to purify the protein of interest. In the column a protein-A-coupled resin (Captiv-A PriMab affinity resin by RepliGen) was placed to allow the binding of the IgG-fused protein. Once all the media was passed through the column, the resin was washed 3 times with washing buffer, to remove any residual media, equilibrated with

TNE and then TNED. The reaction with 10 µg/ml of 3Cpro-GST fusion enzyme was carried-on over-night at 4°C in 1 resin volume of TNED added of leupeptin 1 µM. After the elution of the protein with 6 volumes of HBS, the resin was regenerated incubating it in a solution containing 50mM Gly pH2.0 and 500mM NaCl for 10 minutes and then washing 3 times with about 20 ml of washing buffer. The flow-through of the first round of purification was loaded again to enhance the yield of the purification. The eluted mix was incubated for 1 hour at 4°C on a lab roller with 0.1 ml of protein-A resin and 50 µl of a Glutathione (GSH) Sepharose resin (GE healthcare) and then spun to remove possible uncleaved proteins and the 3C pro enzyme. The purified protein, contained in the supernatant, was concentrated in an appropriate sized “vivaspin 6” column (Sartorius stedim biotech), quantified using a Nano drop 2000c (Thermo Scientific) and checked for purity, aggregation and degradation by size exclusion chromatography (SEC) (see also par. 3.1.12) and SDS-PAGE stained with Coomassie (see also par. 3.2.1).

3.1.12 Size exclusion chromatography

Size exclusion chromatography (SEC) experiments were used to assess purity, aggregation and degradation of purified protein, as well as for checking the binding properties of two injected proteins. The experiments were performed using an AKTA instrument (GE Healthcare) with a constant flow speed of 0.5 ml/min, using a buffer containing Hepes 10mM pH 7.4 and NaCl 150 mM. 200 µl of protein were loaded for each protein and the chromatogram was recorded.

3.2 BIOCHEMISTRY

3.2.1 SDS-PAGE and Western Blot

Coomassie stained SDS-PAGE and Western blot analysis were used for several purposes during this study, like the evaluation of purified proteins (par. 3.1.11), the estimation of the level of expression of stable transfected cells (par. 3.1.9) and the detection of protein pulled down through co-immunoprecipitation (par. 3.2.2).

The protein samples were added with 4x SDS-PAGE loading buffer and boiled for 5 minutes at 100°C. They were then loaded on a 10% SDS polyacrylamide

precast gel (Bio-Rad) and run in a SDS-PAGE apparatus (Bio-Rad) at 200 V for 35 minutes.

At the end, the proteins were either stained with Coomassie Brilliant Blue R-250 (Bio-Rad) for 1 hour or transferred on a nitrocellulose membrane in a transfer apparatus (Bio-Rad) at 100 V for 45 minutes for immuno-blotting.

The membrane was blotted with the following protocol:

- 1 hour incubation in 5% milk;
- 1 hour incubation in a solution of 1% milk in which a specific antibody was added;
- 3 washes in PBS;
- 1 hour incubation in a solution of 1% milk in which a secondary antibody conjugated with horseradish peroxidase (HRP) was added;
- 3 washes in PBS;
- addition of the HRP substrate (HyGLO™ Chemiluminescent HRP Antibody Detection Reagent, Denville Scientific)
- exposure of an autoradiography film (HyBlot CL, Denville scientific Inc.) and development in a KODAK MIN-R film processor.

The following antibodies (Ab) were used, at the final concentration suggested by the manufacturer:

- Peroxidase-conjugated AffiniPure Goat Anti-Human IgG (Jackson Immuno Research, USA);
- Monoclonal mouse Anti-FLAG (Sigma);
- Monoclonal mouse Anti-HA (Fisher Scientific);
- Goat anti-mouse IgG-HRP (Santa Cruz Biotechnology);
- Goat anti-rabbit IgG-HRP (Santa Cruz Biotechnology).

Note: when using the anti-human IgG to blot proteins fused with the human-fc, no primary Ab is needed.

3.2.2 Co-immunoprecipitation experiments

In order to evaluate protein binding, co-immunoprecipitation (co-IP) experiments were performed. HEK cells were transiently transfected, as described in

par. 3.1.8. Depending on the goal of the experiment, either one plate of cells was co-transfected with the plasmids encoding the proteins to be tested or two different plates were independently transfected with one of them and then the cell lysates (or the culture media) were mixed.

- **Co-IP on cell lysate:** the protein content of the cells was collected as in par. 3.1.10. The samples were then incubated with 40 μ l of a protein-A-coupled resin (Captiv-A PriMab affinity resin by RepliGen) or an anti-FLAG resin (Sigma) and placed on a lab roller for 1 hour. At the end they were spun 30 seconds on a microcentrifuge and the pellet was washed 3 times with lysis buffer. 20 μ l of 2x SDS-PAGE loading buffer was added and the samples were boiled for 5 minutes at 100°C and analyzed through western blotting (par. 3.2.1).
- **Co-IP on culture media:** the culture medium of the cells was collected and centrifuged at 1200 g for 5 minutes to remove residual cells. The samples were then incubated with 40 μ l of a protein-A-coupled resin (Captiv-A PriMab affinity resin by RepliGen) or an anti-FLAG resin (Sigma) and placed on a lab roller for 1 hour. At the end they were spun 30 seconds on a microcentrifuge and the pellet was washed 3 times with PBS (Fisher). 20 μ l of 2x SDS-PAGE loading buffer was added and the samples were boiled for 5 minutes at 100°C and analyzed through western blotting (par. 3.2.1).

3.3 PROTEOMICS

3.3.1 Fishing experiments

To identify new candidate binding partners for Caspr2 and NLGN2, fishing experiments were performed on three different brain samples.

- Human adult brain (2 grams of frontal cortex bought from a tissue bank);
- Human embryonic brain (2 grams of frontal cortex bought from a tissue bank);
- Rat adult brain.

These experiments were performed as represented in fig. 4.9 (pag. 70). For each brain sample, the extracted proteins were pre-cleared with an IgC-bound resin to filter out all the proteins that would non-specifically bind to the resin or to the human

IgG portion that is used to link the resin to the bait molecule (i.e. Caspr2, NLGN2, etc.). The unbound portion was then incubated with a Caspr2-bound resin and then split in 3 parts and incubated with α -NRXN1-, NLGN1- or NLGN2-bound resin. This order was chosen to both give the priority to the identification of Caspr2 ligands and avoid the progressive depletion of the protein pool after each step.

The buffers used were:

- Buffer A: 20 mM Hepes pH 7.4 +
150 mM NaCl +
1,2% triton X100

- Buffer B: Buffer A +
4 mM CaCl₂ +
4 mM MgCl₂

The resins used were prepared incubating 750 μ l of a protein-A-coupled resin (Captiv-A PriMab affinity resin by RepliGen) with a saturating amount of the following purified proteins (see also par. 3.1.3 and 3.1.11) for 1 hour at 4°C on a lab roller:

- IgC, intended as the control IgG, corresponding to only the fc portion fused with the other proteins;
- Caspr2-IgG;
- α -NRXN1-IgG;
- NLGN1-IgG;
- NLGN2-IgG.

Each brain sample was homogenized in about 70 ml of buffer A. The volume was then adjusted to 210 ml (105 ml buffer A + 105 ml buffer B). The lysates were incubated for 2 hours at 4°C on a lab roller and then for 30 minutes at 37°C. To precipitate the cellular debris and collect the proteins, the lysate was centrifuged 3 times, collecting at each step the supernatant, as follows:

- 10000 g, 20 min, 4°C;
- 16000 g, 20 min, 4°C;

- 16000 g, 20 min, 4°C;

The clarified lysate was incubated with 750 μ l of the IgC-bound resin and incubated for 1 hour at 4°C on a lab roller. The mix was then poured in a column (Borosilicate Glass Economy Chromatography Columns, Kimble Chase) and the flow-through was collected. The resin was collected in 750 μ l of buffer B for the washing and elution step (see below).

The flow-through was incubated with 750 μ l of the Caspr2-bound resin for 1 hour at 4°C on a lab roller. The flow-through and the resins were collected as before.

This flow-through was split in 3 equal amounts and each was incubated with 250 μ l of the α -NRXN1-, NLGN1- or NLGN2-bound resin for 1 hour at 4°C on a lab roller. The resin was collected as before, while the flow-through was discarded.

All the collected resins were washed in 2 volumes of buffer B and eluted by boiling at 100°C for 5 minutes. At this point the proteins were precipitated by incubating for 10 minutes in 4 volumes of Trichloroacetic acid (TCA) and centrifuging 4 minutes at max speed. The protein pellet was then washed two times in cold acetone and dried at 95°C for 10 minutes. The proteins collected were then analyzed by Mass Spectrometry at the Biomolecular/Proteomics Mass Spectrometry Facility located in the University of California San Diego (UCSD).

3.4 BIOPHYSICS

3.4.1 Isothermal Titration Calorimetry

To study the binding properties of two purified proteins, ITC experiments using Microcal iTC200 (GE Healthcare) were performed, using HEPES buffer with Ca^{2+} . The scheme presented in fig. 3.2 shows the principle of this technique. A protein is titrated into a solution of its binding partner. The instrument measures the binding enthalpy change due to the interaction by detecting the heat exchange in the chamber: this is calculated by the instrument through the power that it applies to the sample cell to keep the temperature difference with a reference cell constant. Several injections of the titrant are performed and the heat exchange is recorded after each one: in case of association the signal diminishes as the system reaches saturation, until the only heat exchange is due to the dilution of the protein and background noise are

observed. A graph recording the $\mu\text{cal}/\text{sec}$ exchanged over time is reported as results; a binding curve is then obtained from a plot of the integral of each heat exchange peak from each injection against the ratio of ligand and binding partner in the cell. The binding curve is analyzed with the appropriate binding model to determine K_d , stoichiometry (N), ΔH and ΔS (see also fig. 4.1 for some examples).

The samples used for ITC experiments need to be in matching buffers and the two proteins have to be pure to avoid non-specific signals. As a control, a separate experiment injecting the titrant protein in a chamber containing only buffer is performed and the values are subtracted from the previous one, to correct the signal due to the dilution of the protein.

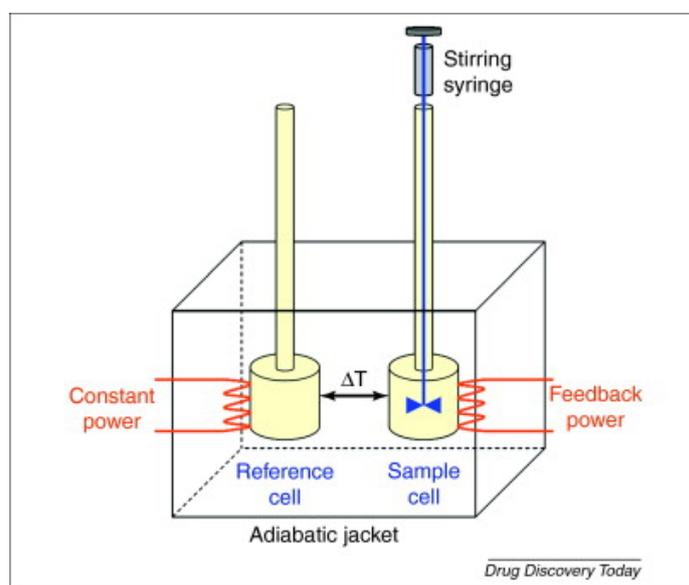


Figure 3.2: Basic configuration of an isothermal titration calorimeter (Núñez et al., 2012).

4. RESULTS

4.1 CHARACTERIZATION OF THE BINDING BETWEEN CASPR2 AND TAG-1

Since when the association of Caspr2 and TAG-1 was proposed in 2003, research has been focused on its physiological role in myelinated neurons, trying to explain the association of this complex with several diseases, such as ASD, Epilepsy, Gilles de la Tourette syndrome, etc. Little attention has been given instead to the characterization of the binding proprieties of the complex. All the available data come from pull-down experiments performed on rat brain and on cell lysates from HEK co-transfected with the two exogenous proteins. Multiple questions remain thus unsolved, as, for example, what the affinity and the stoichiometry of the complex are and which domains and which residues it involves.

4.1.1 Biophysical analysis of the complex

For a first characterization of the binding properties of Caspr2 and TAG-1, an Isothermal Titration Calorimetry (ITC) experiment was performed. This technique uses purified soluble proteins to calculate the binding affinity of a complex. It has the advantage that the association can happen in solution, removing potential artifacts, such as binding site unavailability in immobilization-based techniques. Since Caspr2 has a trans-membrane hydrophobic domain, it is not suitable for an experiment performed in an aqueous buffer. In any case, what emerges from the available information on the two molecules is that, with TAG-1 being a GPI-anchored protein, the binding should take place between the extracellular portions of the proteins. It is therefore possible to focus on these soluble parts.

The ITC system uses the concept that the binding process involves a certain amount of energy exchange with the environment; by measuring that heat exchange we can have an indication of the association potential. Two interacting molecules will show decreasing heat exchange values after each injection of one of them, used as a titrant, in a solution containing its binding partner, reflecting the progressive occupation of all the protein binding sites, resulting then in a curve sigmoidal and asymptotic to the zero line; non interacting proteins, instead, won't show any heat exchange, if not the one due to the dilution of the titrant, resulting in a flat curve centered on the zero line. By measuring the extent of the energy exchanged and the concentration of the proteins, the parameters of the binding, such as K_d and stoichiometry, can also be calculated.

In fig. 4.1A the results of the ITC experiments are reported. The purified proteins used are Caspr2-1261* and TAG-1-1004*, corresponding to their entire extracellular domains. Surprisingly, no binding was detected. As a control, an ITC experiment was performed using the proteins NLGN1-639* and β NRXN1-300*. As for Caspr2 these two proteins have a trans-membrane hydrophobic domain, so only the extracellular soluble domains were purified and used for this binding experiment. Moreover, these two proteins are well known to interact: for the NLGN1-639* / β NRXN1-300* association, Comoletti and colleagues (2006) reported 30 nM by SPR; Araç and colleagues (2007), instead, reported 97 nM, and Chen and colleagues (2008) also reported 93 nM, both by ITC. This couple of proteins represents thus a good positive control. As shown in fig. 4.1B, the results of this experiment reported an excellent binding curve with a K_d of 116 nM and a stoichiometry very close to 1, proving the reliability of this system.

Because these results were not consistent with the generally accepted idea that Caspr2 and TAG-1 directly bind *in vivo*, it was necessary to perform other experiments using a different approach, in order to gain more insights into this controversial aspect. For this reason, a size exclusion chromatography experiment was performed. This technique gives the possibility to separate proteins on the basis of their molecular weight and shape. The peak profile of the two proteins injected after one hour of co-incubation would reflect their actual binding abilities: in the case that a tight complex is formed, it would be eluted faster than the two proteins injected alone and an additional peak, shifted to a lower elution volume, would appear; instead, in the case that no association resulted from the incubation, no peak shift

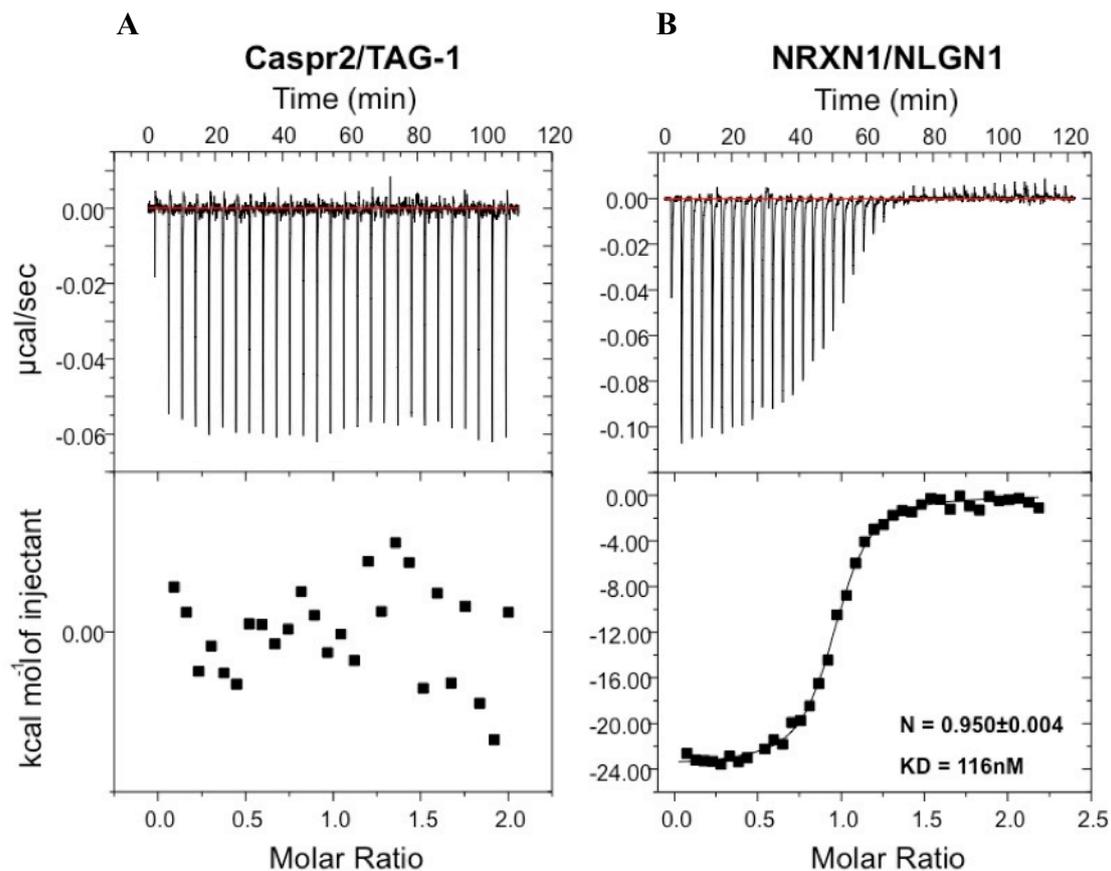


Figure 4.1: Results for the ITC analysis using the proteins Caspr2-1261*, used as titrant at a concentration of 61 μM , and TAG-1-1004*, used at the concentration of 6 μM (A), or $\beta\text{NRXN1-300}^*$, used as titrant at the concentration of 100 μM , and NLGN1-639*, used at the concentration of 10 μM (B). In case of association the calculated values for K_d and stoichiometric ratio (N) of the interaction are also reported.

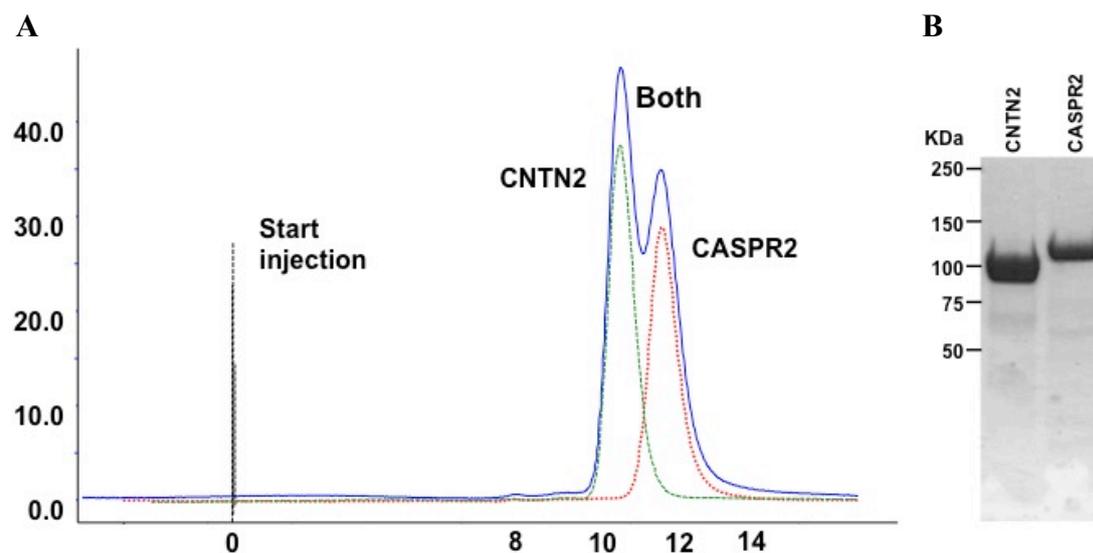


Figure 4.2: (A) Size exclusion chromatography results of Caspr2 (red line) and TAG-1 (green line) injected alone and of the two proteins injected together (blue line). (B) SDS-PAGE analysis of the purified Caspr2 and TAG-1 proteins.

would be observed. As seen in fig. 4.2A, no differences can be noted in the elution peak profiles when the proteins are injected together (blue line) in comparison to the independent injection of Caspr2 (red line) and TAG-1 (green line), while the relative height of the peaks is also conserved. Moreover, the analysis of the two purified proteins by means of Coomassie blue staining, reported in fig. 4.2B, showed two single bands at the expected molecular weight for Caspr2 and TAG-1: together with the symmetry of their corresponding peaks in the chromatogram, this proves that the proteins used in these experiments were pure, monodispersed, and not degraded. All together, these results are consistent with the ITC analysis previously shown, providing strong evidence that when these two proteins are put together in an aqueous medium *in vitro*, they don't interact.

4.1.2 Co-immunoprecipitations using full-length and truncated proteins

The ITC and chromatography experiments, performed using the purified extracellular portion of Caspr2 and TAG-1, failed to confirm the association between these two proteins. Even though they provided strong data proposing the idea that the two molecules do not actually interact, these experiments had the limitation of being performed *in vitro*. It was, therefore, interesting to replicate the binding data published in 2003 (Poliak et al., 2003, Traka et al., 2003, see also fig. 1.7 and 1.8), whose authors instead used an *ex-vivo* approach. For this purpose, the same co-immunoprecipitation (co-IP) experiments proposed in those articles were performed. These consisted of the use of two different protocols:

1. **Co-transfection:** HEK-293 cells were transfected with both the plasmid encoding HA-Caspr2 wt and the one encoding FLAG-TAG-1 wt and then the cell lysate was used for immunoprecipitation, using anti-FLAG antibodies;
2. **Independent transfection:** two different plates of HEK-293 cells were transfected with either of the two plasmids independently. The two cell lysates were then mixed together and used for immunoprecipitation, using anti-FLAG antibodies.

In both cases an experiment using Caspr2 and an empty vector was performed as a control.

Both Poliak and Traka reported that the interaction happens only when the proteins are co-expressed, proposing that this was due to the need for the two proteins to associate in “cis” on the same cell surface membrane. Consistent with those results,

the western blot analysis shown in the top panel of fig. 4.3, reveals that, using the co-transfection protocol, the precipitation of TAG-1 pulls-down also Caspr2, while the mock experiment result was negative. On the contrary, when transfecting the two plasmids independently, the Caspr2 band was not visible after the FLAG immunoprecipitation (fig. 4.3, bottom panel). Thus, our experiments replicated the data available in literature.

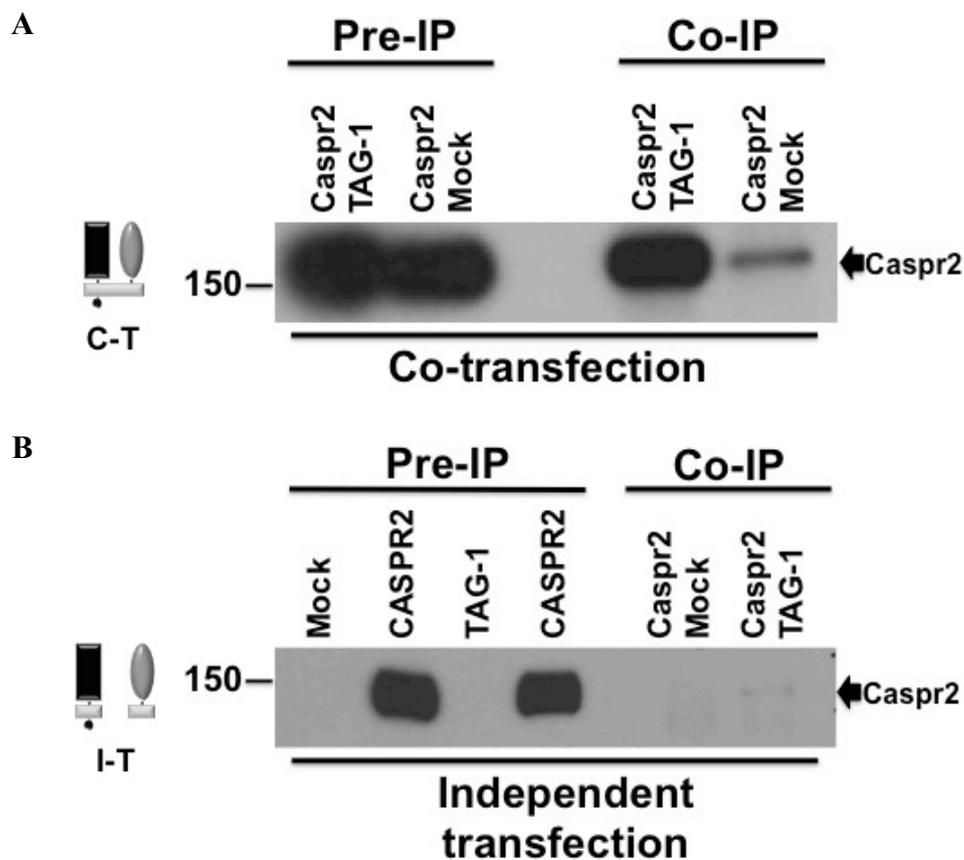


Figure 4.3: HEK-293 cells were co-transfected (**A**) or independently transfected (**B**) with plasmids encoding HA-Caspr2 wt or FLAG-TAG-1 wt. The final lysates (that in the case of independent transfection result from the mix of two different lysates) were incubated with anti-FLAG antibodies for the immunoprecipitation of TAG-1. As negative control, an empty pCDNA3.1 vector, coding only for the FLAG peptide (Mock), was used in place of TAG-1. The precipitated proteins were evaluated by western blotting using anti-HA antibodies.

The need of the Caspr2-TAG-1 complex to be formed during protein maturation and delivery to the cell membrane would explain the negative results obtained using purified proteins *in vitro*, but if this is the case more evidence should

be provided to test this hypothesis. For example, the same model has been proposed for the interaction of the proteins Caspr1 and CNTN1: the binding happens in the ER and the complex is delivered already formed to the cell membrane. However, the transport of Caspr1 to the surface is dependent on CNTN1, since when expressed alone in CHO cells, Caspr1 is retained in the ER (Faivre-Sarrailh et al., 2000). This, instead, doesn't happen to Caspr2 that is trafficked to the cell membrane even when expressed independently of TAG-1 (Falivelli et al., 2012).

To prove if effectively Caspr2 needs to bind TAG-1 in the ER in order to form a complex, the extracellular domains alone can be used. By co-expressing HA-Caspr2-1261* and FLAG-TAG-1-1004*, the complex would in theory be secreted and it could be pulled-down from both the culture media and the cell lysate. In fig. 4.4 the results are reported: as expected, after the immunoprecipitation of TAG-1 in the lysate, the band corresponding to Caspr2 is visible in western blotting. However, the same is not true when TAG-1 is pulled down in the media, indicating that the complex of the truncated forms is not secreted. As a control, TAG-1 was correctly secreted and pulled down, both in lysates and in media.

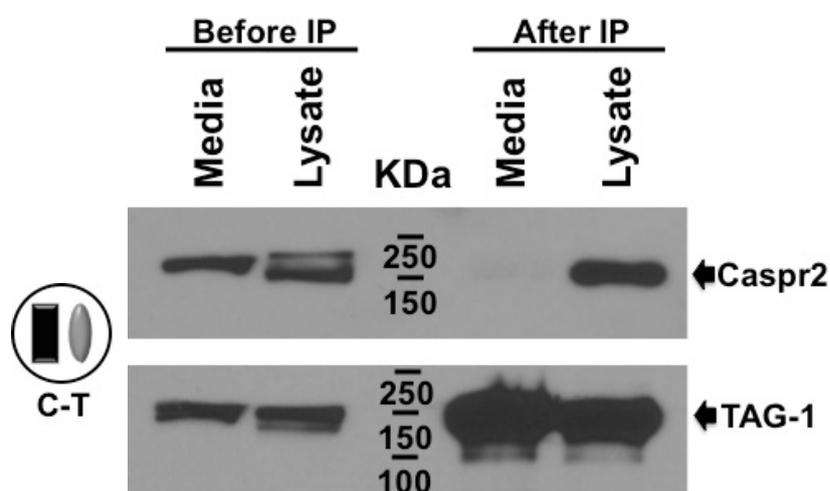


Figure 4.4: HEK-293 cells were co-transfected with plasmids encoding HA-Caspr2-1261* or FLAG-TAG-1-901*. Both cell lysates and culture medium were incubated with anti-FLAG antibodies for the immunoprecipitation of TAG-1. The precipitated proteins were evaluated by western blotting using anti-HA (to detect Caspr2) or anti-FLAG (to detect TAG-1) antibodies.

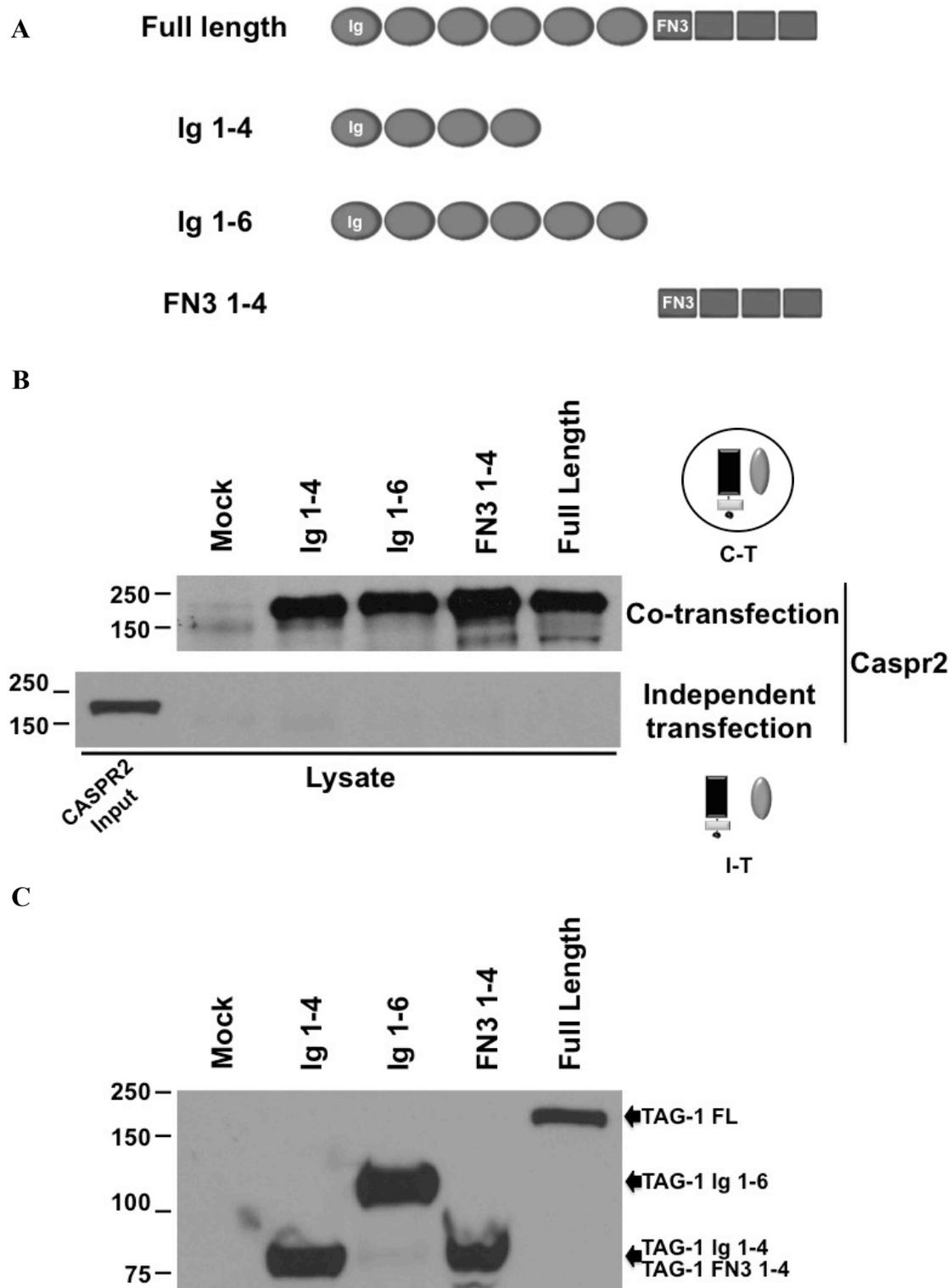


Figure 4.5: (A) Schematic representation of the truncated fragments of TAG-1 used in co-IP experiments. Ig: Immunoglobulin domains; FN3: Fibronectin like domains. (B-C) Results of the co-IP experiments using truncated versions of TAG-1. HEK-cells were co-transfected or independently transfected with plasmids encoding HA-Caspr2 wt or the FLAG-tagged fragments of TAG-1 above reported. The final lysates (that in the case of independent transfection result from the mix of two different lysates) were incubated with anti-FLAG antibodies for the immunoprecipitation of TAG-1. The precipitated proteins were evaluated by western blotting using anti-HA antibodies to detect Caspr2 (B) or anti-FLAG antibodies to detect TAG-1 for the independent transfection (C).

Moreover, in order to get further insight, the characterization of the putative binding site conveyed interesting information. Co-IP experiments were performed using truncated versions of TAG-1. The fragments tested are reported in fig. 4.5A and consist of:

- Immunoglobulin (Ig) domains 1-4 (aa 28-461),
- Ig domains 1-6 (aa 28-597);
- Fibronectin (FN) like domains 1-4 (aa 600-1027).
- Full-length TAG-1.

The experiments were carried out with both the co-transfection and the independent-transfection protocols, using also HA-Caspr2 wt. If in the neuron binding between Caspr2 and TAG-1 and between the latter and another TAG-1 molecule expressed from the glial cells occurs, it is expected that these two events would require different surfaces: therefore only one of the fragment used would give a positive result. Unexpectedly, the results, reported in fig. 4.5B, showed that, using the co-transfection protocol, a positive band is visible for each of the truncated forms of TAG-1 tested, even when they don't share any overlapping regions. Moreover, the intensities of the bands are similar between them and to the full-length protein. As expected, using the independent-transfection protocol, no band is visible, even if all the truncated TAG-1 proteins were correctly expressed and pulled-down (fig. 4.5C).

Altogether, the information provided by these results seems conflicting: on one hand, the complex can be pulled down only in cell lysates and only after the co-expression of the two proteins; on the other hand, the extracellular domains fail to bind when tested by means of both ITC and co-IP. Moreover, non-overlapping portions of TAG-1 can pull down Caspr2. Even if it cannot be excluded that the binding region of the two proteins involves different domains, at least a reduced affinity would be expected when using only a portion of TAG-1: an artifact in the set-up of the co-IP experiment has thus to be taken in account.

4.1.3 Validation of the co-IP protocol

Due to the conflicts among the data presented above, the reliability of the used co-IP protocol had to be tested. For this purpose, both positive and negative controls were needed.

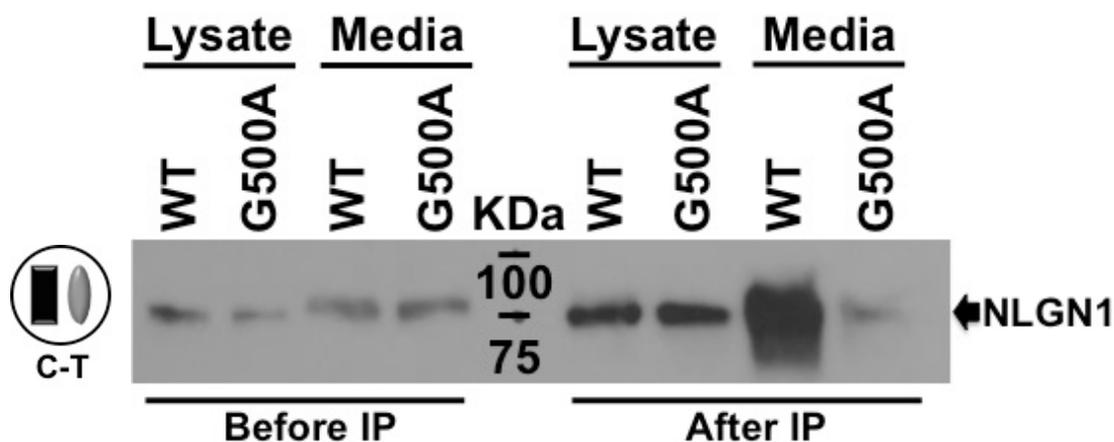


Figure 4.6: HEK-293 cells were co-transfected with plasmids encoding either the couple of proteins FLAG-NLGN1-639* and β NRXN1-300*-IgG or the couple of proteins FLAG-NLGN1-639*-G500A and β NRXN1-300*-IgG. Both cell lysates and culture medium of each plate were incubated with a protein-A coupled resin for the immunoprecipitation of β NRXN1. The precipitated proteins were evaluated by western blotting using anti-FLAG antibodies.

Because no protein, other than TAG-1, has been proposed to directly bind Caspr2, the proteins FLAG-NLGN1-639* and β NRXN1-300*-IgG were used as a positive control. These proteins were already used for the same purpose in the ITC experiments, due to their high affinity interaction. Moreover, in 2010 Leone and colleagues proved that the mutation of the glycine in position 500 to an alanine could abolish the binding propriety of NLGN1 to its counterpart: therefore, the G500A mutant was used as a good negative control for these experiments. All these proteins are truncated forms, corresponding to the extracellular portion that is known to be responsible for the NLGN-NRXN binding (see par. 1.3.3): thus the complex is expected to be precipitated from the cell lysate as well as the media. The co-IP was performed with the same protocols used for the secreted Caspr2 and TAG-1 (fig. 4.4), pulling down β NRXN1 through a protein-A coupled resin that binds its IgG tag. The results are presented in fig. 4.6. As expected NLGN1 wt was co-precipitated both in the media and in the cell lysate. The interesting results come instead from the G500A mutant: as expected, whereas in the media almost no protein was detected, in the cell lysate a band similar in intensity to the wt could be found. Altogether this proves that it is possible to pull-down a secreted complex from the culture media, after the co-expression of the proteins from the same cell, proving how the failure in the co-precipitation of the Caspr2-TAG-1 complex was effectively due to a lack of binding or a low affinity of the putative complex and not to an experimental issue. At the

same time, the co-expression of two glycoproteins leads to the generation of artifacts when the binding ability is evaluated in the cell lysates.

To better analyze this potential to generate artifacts of the proposed co-IP protocol, several proteins that are not supposed to bind Caspr2 were tested. These are:

1. **Caspr2 domain 1** (aa. 1-183): Caspr2 does not dimerize (Comoletti unpublished data). Moreover, this represents only one domain of the extracellular portion of the protein (Discoidin domain, fig 1.3);
2. **CNTN1** (aa 21-986): this protein that belongs to the same family of TAG-1 is supposed to bind Caspr1, but not Caspr2 (Poliak et al., 2003). Also this protein is GPI anchored and for these experiments it was truncated to be released;
3. **Clusterin** (CLU), also known as APOJ (full length): no association has ever been reported between these two proteins. Moreover, fishing data (see par. 4.2.1) does not predict any binding. CLU is a naturally secreted protein, cleaved in two halves that remain associated through disulfide bonds (De Silva et al., 1990).

All the proteins were expressed fused to both a FLAG at the N-terminal and an IgG tag at the C-terminal. The plasmids encoding these proteins and Caspr2-1261* were either co-transfected or independently transfected in HEK cells as already done for the extracellular domains of Caspr2 and TAG-1. The pull-down of the IgG fused protein, using a protein-A coupled resin, resulted in the co-precipitation of Caspr2 in all the cases (fig. 4.7A, top panel). On the contrary, when the proteins were independently expressed and the cell lysates mixed together no Caspr2 band could be detected (fig. 4.7A, bottom panel), even if all the proteins were correctly expressed and pulled-down (fig. 4.7B).

The results just presented are consistent with those emerging from the use of NLGN1 wt and NLGN1 G500A. Anyhow, when comparing these experiments to those reported in fig. 4.3 for the couple Caspr2-TAG-1, it has to be taken in account that even if the protocols used are the same and the results should be comparable, these last two experiments were performed using different tags for the precipitation. In fact, while TAG-1 was fused to a FLAG peptide, these proteins carry a human IgG-fc tag. It is possible that this additional peptide could bridge a non-specific association when it co-localizes with Caspr2 in the ER. To verify that this is not the case, a co-IP

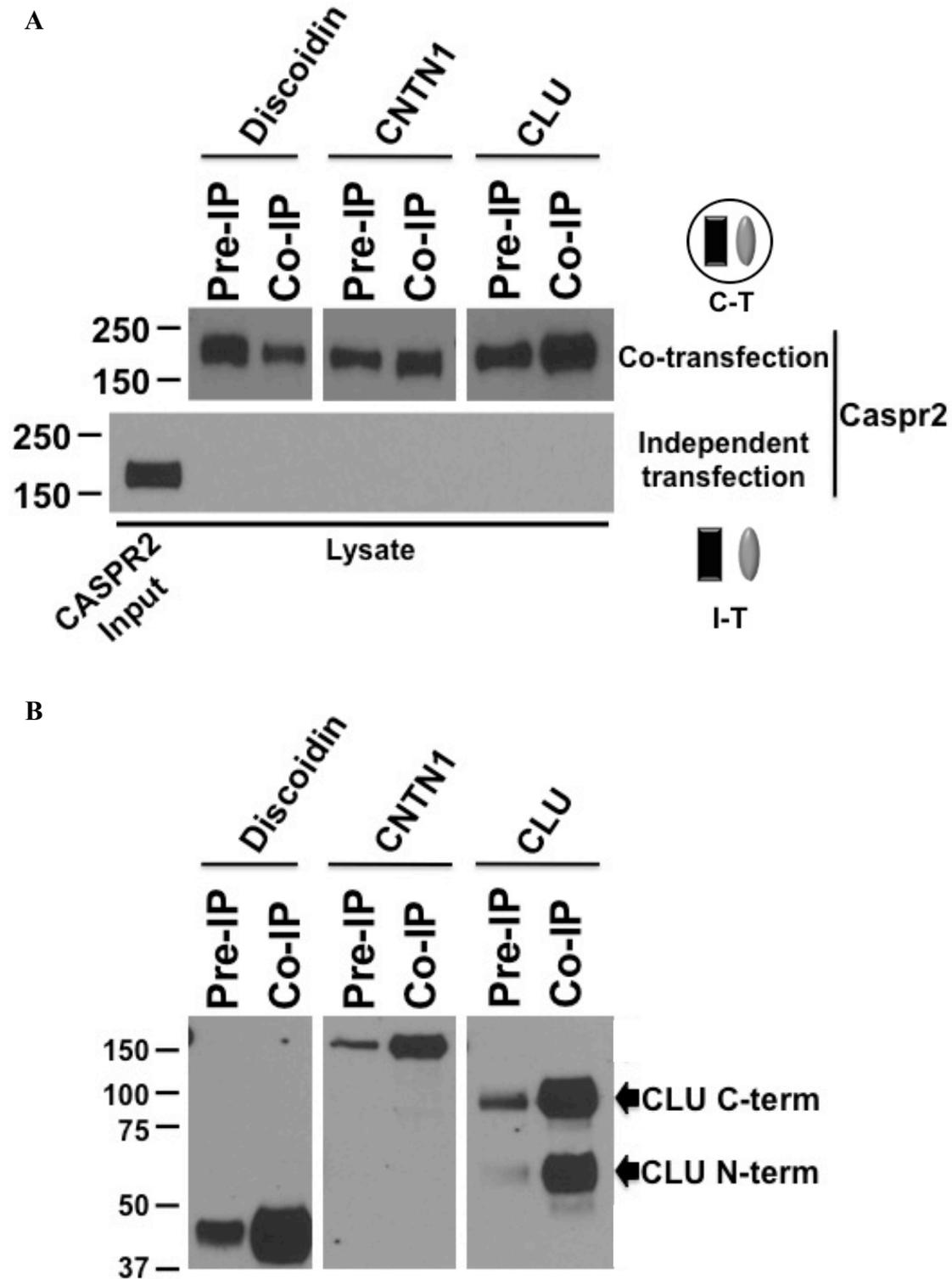


Figure 4.7: HEK-cells were co-transfected or independently transfected with plasmids encoding HA-Caspr2 wt or one protein among FLAG-Caspr2-183*-IgG, FLAG-CNTN1-986*-IgG and FLAG-CLU-wt-IgG. The final lysates (that in the case of independent transfection result from the mix of two different lysates) were incubated with a protein-A coupled resin for the immunoprecipitation of the tested protein. The precipitated proteins were evaluated by western blotting using anti-HA antibodies (A) to detect Caspr2 or anti-human IgG antibodies (B) to prove the correct precipitation of the pulled-down proteins for the independent transfection.

experiment was performed using HA-Caspr2 co-expressed with the IgG peptide alone (IgC) or a TAG-1-IgG fusion protein. Moreover, 1 ml of FBS was added during the precipitation phase of the mock experiment to check if one of the proteins present in the serum, that is always present in the culture of the cells, can interfere with the experiments. As shown in fig. 4.8, only TAG-1 can pull-down Caspr2, proving that these variables are not responsible of the positive bands previously shown.

Altogether the results presented show that the co-transfection protocol for co-IP presented above can't be used when the association of the two glycoproteins is evaluated in the cell lysates, because false-positive artifacts can be generated: in fact, every tested protein can be precipitated using this protocol, regardless of the actual binding affinity. The most likely and simple explanation of this phenomenon is that a cellular protein bridges the contact between the two proteins during their maturation, but is no longer bound when they are delivered to the cell membrane. Unpublished data from Comoletti lab proved that one of the proteins that exerts this function is Calnexin, a resident chaperone of the Endoplasmic Reticulum (ER). This flaw could have affected the original experiments conducted in 2003, in which over-expressing the two proteins in the same cell, the authors observed a direct association between Caspr2 and TAG-1 (Poliak et al., 2003; Traka et al., 2003; see also fig. 1.9 for the proposed model). The results reported here also show no direct binding for this couple of proteins. However, no information is provided on whether or not the two proteins belong to the same macromolecular complex bridged by a third protein. The experiments proving this kind of association were conducted on rat brain's lysates by the same authors: false positive artifacts due to the post-translational binding to ER resident chaperones would affect more likely experiments performed with proteins that were over-expressed rather than those performed on brain lysate's sample. Moreover, strong evidence comes from co-localization data published on the same papers. It is therefore likely that Caspr2 and TAG-1 associate indirectly, but further experiments are needed. This evidence, together with the fact that TAG-1 was the only proposed ligand of the extracellular portion of Caspr2, makes very interesting the characterization of new potential partners for this protein.

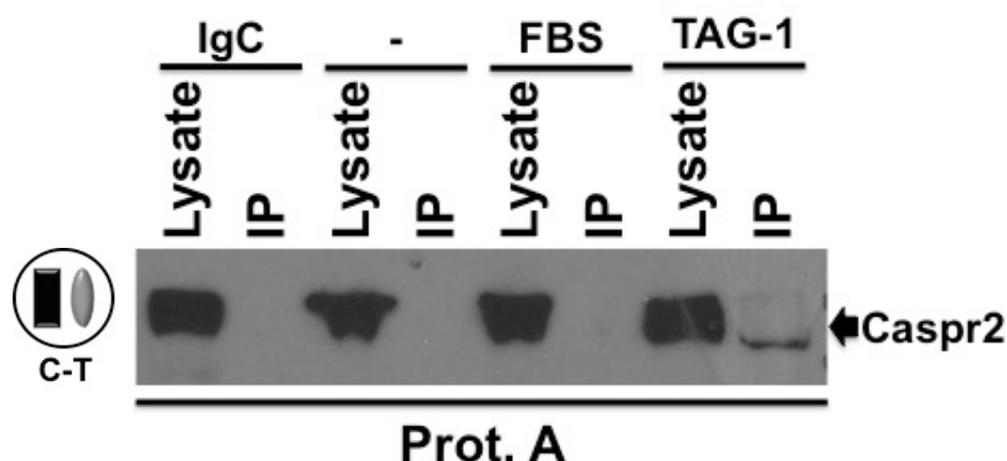


Figure 4.8: HEK-cells were transfected with a plasmid encoding HA-Caspr2 wt. To address if the IgG protein can cause non-specific interaction with Caspr2, the cells were also transfected with a pGL6-XL4 plasmid expressing only the human IgG fc portion. The co-expression of TAG-1 was used as a positive control, as this co-IP was found always positive in this condition. The lysates were incubated with a protein-A coupled resin. To check also if proteins contained in FBS can non-specifically precipitate Caspr2, 1 ml of FBS was added to the lysate of cells expressing only Caspr2. The precipitated proteins were evaluated by western blotting using anti-HA antibodies to detect Caspr2.

4.2 IDENTIFICATION OF NOVEL BINDING PARTNERS FOR CASPR2 AND NLGNs' FAMILY MEMBERS

4.2.1 Fishing experiments

As already mentioned, it is reasonable to think that Caspr2 and NLGN2 have binding partners not yet identified. In fact, Caspr2 has been found in other brain regions, like the synapse, where there's no K^+ channel clustering (Bakkaloglu et al., 2008; Poliak et al., 1999; Zweier et al., 2009). In this case, it would likely exert other non-characterized functions and bind different proteins. Moreover, it has been reported in the first part of this thesis that the association between TAG-1 and Caspr2 is not direct. In any case, it is likely that in-vivo they belong to the same molecular complex: if this is the case, their association has to be mediated by another unknown extracellular protein. Moreover, as long as TAG-1 is the only proposed molecular partner of the extracellular portion of Caspr2, the latter wouldn't have a known binding partner. NLGN2, instead, shows several differences from the other family members both in the aminoacidic sequence and in the affinity for NRXNs (Südhof, 2008; Comoletti et al., 2006), leading to the hypothesis that it has not yet identified associated proteins.

In order to identify uncharacterized binding partners of these two proteins, a fishing experiment was performed. This consisted of an affinity chromatography analysis of brain samples coupled with a mass spectrometry-based identification of the retained proteins. Briefly, the protein content of a small sample of adult and embryonic human prefrontal cortex (provided by a tissue bank) and of a rat brain were passed on a system of affinity chromatography columns, with different proteins bound to the stationary phase. These were:

- IgC (IgG of Control), corresponding to the fc portion of a human IgG. It has been used as a control for the other IgG fusion protein used;
- Caspr2-1261*-IgG;
- NLGN2-616*-IgG;
- NLGN1-639*-IgG;
- α NRXN1-1309*-IgG.

In fig. 4.9 the scheme of the experiment is reported: the brain lysate was incubated with an IgC bound resin and loaded in a chromatography column to separate the beads from the unbound lysate. This step not only provides a control for the other IgG fusion proteins, but also allows to reduce the complexity of the sample, removing the proteins that would bind non-specifically either the resin or the IgG alone. The flow-through was incubated with a Caspr2-IgG-bound resin and the beads were then collected. Finally, the unbound fraction was split in three parts and each was incubated with a different resin (α NRXN1-IgG, NLGN2-IgG and NLGN1-IgG). In all the cases the beads were collected. The bound proteins were finally eluted and identified through mass spectrometry analysis. The scheme was designed in this way in order to balance three different aspects: to give the priority to the identification of Caspr2 ligands, to avoid the progressive depletion of the protein pool after each step and to maximize the use of precious material. NLGN1 was used as a control for NLGN2, being that they are both members of the same family, while α NRXN1 was used as a control for Caspr2, due to the high sequence homology between these two proteins.

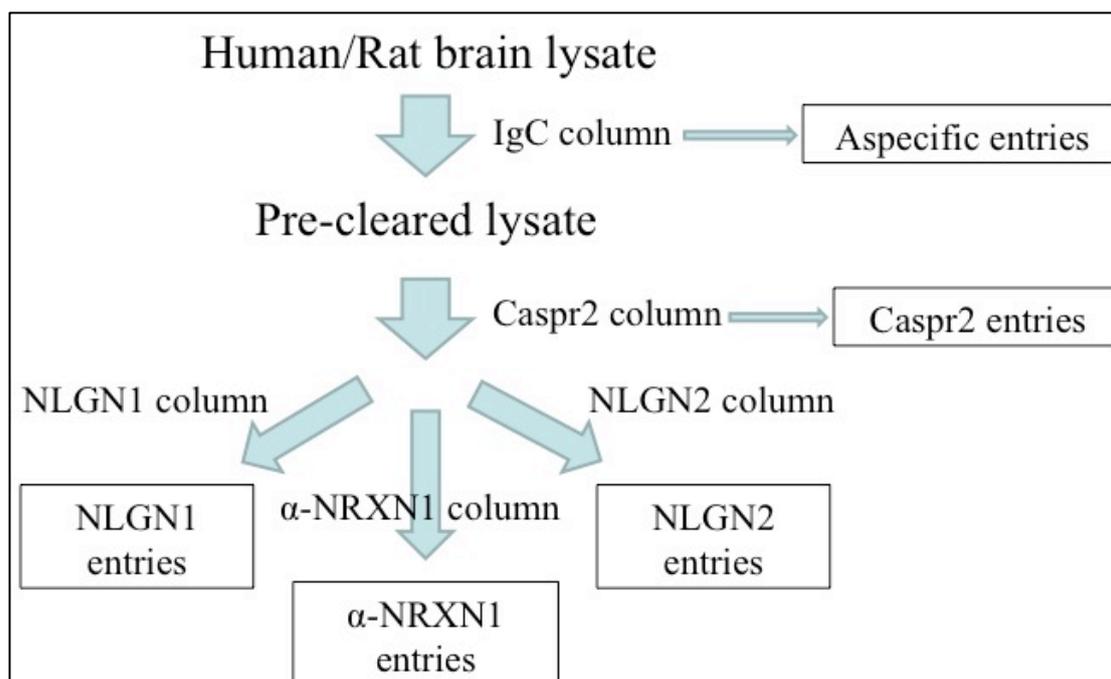


Figure 4.9: Scheme of the fishing experiment.

About three thousands proteins were identified and each one was reported as an entry with an assigned value for each resin analyzed (tab. 4.1 for an example): this number reflects how many times a certain peptide was identified in that resin. Obviously, the higher the number is, more likely it is that the association observed is specific. The proteins were then selected and ranked using the following criteria:

1. A value for the IgC column null or very close to 0;
2. A high value for the Caspr2 or NLGN2 column, coupled with a very low value for the α NRXN1 or NLGN1 column respectively. Priority was given to the proteins identified with this criterion in more than one tissue, because it would reflect a supposedly more specific binding;
3. A localization at least partially extracellular: both Caspr2 and NLGN2 are type I trans-membrane proteins with a big extracellular domain accounting for the functional bindings of the proteins. This study is therefore more interested in finding extracellular binding partners, than intracellular ones.

Due to the variable conditions intrinsic in this kind of experiments it is not possible to set objective criteria for this ranking, so these have to be subjective and interpreted case by case, based also on the available information for the examined

proteins. The values for the NLGN1 protein found in the α NRXN1 column can also be considered a positive control and give some hints while evaluating the values of the other proteins: however, when considering this kind of control, it has to be taken into account that the NRXN1-NLGN1 binding has a high affinity and it's not predictable how the values can change for looser binding. Moreover, it has to be considered that the pull-down of a protein in this kind of fishing experiment can reflect also a non-direct association, but, if this is the case, the bridging protein is expected to be present in the list of candidates. For all these reasons each possible interaction has to be confirmed. The list of the candidate proteins identified is reported in tables 4.1, 4.2 and 4.3.

Human fetus - Identified Proteins						
Symbol	Name	IgC	Caspr2	α NRXN1	NLGN1	NLGN2
α NRXN1	α Neurexin 1 [Homo sapiens]	1	0	1037	320	7
SFRP1	Secreted frizzled-related protein 1 [Homo sapiens]	0	30	5	2	0
CPNE1	Copine-1 [Homo sapiens]	0	19	3	0	5
CLU	Clusterin (Apolipoprotein J) [Homo sapiens]	1	1	0	2	18
GPC1	Glypican-1 [Homo sapiens]	8	3	6	79	3
GPC2	Glypican-2 precursor [Homo sapiens]	0	5	0	68	0
SDC3	Syndecan-3 [Homo sapiens]	2	0	0	62	1
TNC	Tenascin C [Homo sapiens]	8	2	0	47	0
AGRN	Agrin [Homo sapiens]	6	1	1	46	0
NCAN	Neurocan [Homo sapiens]	4	3	5	43	11
VCAN	Versican [Homo sapiens]	3	2	2	43	4
TNR	Tenascin R (restrictin, janusin) [Homo sapiens]	0	1	1	33	0
SDC2	Syndecan-2 [Homo sapiens]	0	0	0	29	0
TMEFF1	Isoform 1 of Tomoregulin-1 [Homo sapiens]	0	0	0	23	0
SCUBE1	Signal peptide, CUB and EGF-like domain-containing protein 1 [Homo sapiens]	0	5	0	16	0
SPON1	Spondin-1 [Homo sapiens]	0	0	0	13	0
BCAN	Brevican [Homo sapiens]	0	0	0	13	0
SDC1	Syndecan-1 [Homo sapiens]	0	2	0	12	0
SDC4	Syndecan-4 [Homo sapiens]	0	0	0	5	0

Table 4.1: List of candidates identified by fishing in the Human Fetus brain sample. For each candidate, the values for the five columns are reported. The candidates are arranged into groups according to the protein they are predicted to bind to. White: the values for α NRXN1 are reported as a positive control; Green: candidates for the binding to Caspr2; Yellow: candidates for the binding to NLGN2; Cyan: candidates for the binding to NLGN1. The proteins are ordered according to their values for the predicted partner.

Human adult - Identified Proteins						
Symbol	Name	IgC	Caspr2	α NRXN1	NLGN1	NLGN2
CLU	Clusterin (Apolipoprotein J) [Homo sapiens]	0	5	0	12	157
APOE	Apolipoprotein E [Homo sapiens]	0	0	0	0	12
REN1	Renin Receptor [Homo sapiens]	0	0	0	0	11
TNR	Tenascin R (restrictin, janusin) [Homo sapiens]	1	0	0	76	0
VCAN	Versican [Homo sapiens]	0	0	0	39	0
BCAN	Brevican [Homo sapiens]	0	0	0	39	0
NCAN	Neurocan [Homo sapiens]	0	0	0	37	0
PTPRZ1	Receptor-type tyrosine-protein phosphatase zeta [human]	0	0	0	24	0
Caspr1	Contactin associated protein 1 [Homo sapiens]	1	0	0	21	0
CNTN1	Contactin 1 [Homo sapiens]	1	0	0	11	0

Table 4.2: List of candidates identified by fishing in the Human Fetus brain sample. For each candidate, the values for the five columns are reported. The candidates are arranged into groups according to the protein they are predicted to bind to. Yellow: candidates for the binding to NLGN2; Cyan: candidates for the binding to NLGN1. The proteins are ordered according to their values for the predicted partner.

Rat - Identified Proteins						
Symbol	Name	IgC	Caspr2	α NRXN1	NLGN1	NLGN2
SFRP1	Secreted frizzled-related protein 1 [rat]	3	32	0	0	0
NPTXR	Neuronal pentraxin receptor [rat]	2	11	3	0	0
TNR	Tenascin R (restrictin, janusin) [rat]	7	12	8	351	6
PTPRZ1	Receptor-type tyrosine-protein phosphatase zeta [rat]	0	2	5	138	0
BCAN	Brevican [rat]	0	2	4	63	3
CNTN1	Contactin-1 [rat]	4	15	7	59	5
VCAN	Versican [rat]	0	2	0	15	0

Table 4.3: List of candidates identified by fishing in the Human Fetus brain sample. For each candidate, the values for the five columns are reported. The candidates are arranged into groups according to the protein they are predicted to bind to. Green: candidates for the binding to Caspr2; Cyan: candidates for the binding to NLGN1. The proteins are ordered according to their values for the predicted partner.

What emerges from these data, is that some candidates for the binding of Caspr2 and NLGN2 were identified, but the highest number of proteins was predicted to bind NLGN1:

- Caspr2: SFRP1 (see par. 1.4.1 for a brief introduction) was found in two different tissues. Even if the values are not high, it can reflect a specific interaction. Other two candidate proteins have been identified.

- NLGN2: two different apolipoproteins were found (see par. 1.4.2 and 1.4.3 for a brief introduction). One of the two, Clusterin, was found in two different tissues and with relatively high values.
- NLGN1: although this protein has been included as a control for NLGN2, the experiments can be read vice versa with NLGN2 being used as a control for NLGN1. A protein that has high fishing values for this protein, but not for its control can be a specific partner. Even if this work is specifically aimed at the identification of new ligands for Caspr2 and NLGN2, in general the goal is to characterize molecular partners of proteins involved in ASD. Therefore, the high number of candidate proteins found in the NLGN1 column is very interesting. It is also noteworthy that most of the proteins identified are present in the results of all three experiments. If NLGN1 would be confirmed to associate with one of them, new information on its role could be provided. Moreover, some of the proteins found here are known to interact with each other (tab. 4.4). So it is possible that one of them binds directly NLGN1 and the others have been pulled-down as a larger complex. Of course, among these proteins, the candidate that most likely would present an actual interaction is TNR (see par. 1.4.4 for a brief introduction), because it is characterized by the highest values. Also interesting is the presence in the list of Caspr1 and its ligand CNTN1: even if this complex is mainly located at the paranodes (see also par. 1.2.4), recently Caspr1 has been found at the synapse (Santos et al., 2012), thus a similar interaction could be possible. Moreover, CNTN1 has been reported to bind also TNR and PTPRZ1, also identified in this fishing experiment.

Protein	Known ligand	References
TNR	TNC	Anlar and Gunel-Ozcan, 2012
	CNTN1	Rathjen and Wolff, 1991; Zacharias et al. 2002
	Neurocan	Milev et al.,1998
	Versican	Aspberg et al.,1995
	Brevican	Aspberg et al.,1997
	Syndecan	Salmivirta et al.,1991
CNTN1	Caspr1	Faivre-Sarraillh et al., 2000 See also par. 1.2.4
	TNR	Rathjen and Wolff, 1991; Zacharias et al. 2002
	PTPRZ1	Bouyain and Watkins, 2010

Table 4.4: List of TNR and CNTN1 known ligands.

All the predicted ligands identified had to be tested to verify their binding potential. Therefore, experiments using only the two hypothetical ligands had to be performed, in order to prove that their association is direct and specific. The high amount of candidates makes this step impossible to be carried out in a short time without an efficient high-throughput system. Therefore, for each target few proteins were selected for validation: these are chosen among the proteins that could most likely result positive. They are reported in tab. 4.5 and they are SFRP1, CLU, APOE, CNTN1 and TNR.

Protein	Predicted ligand	Human fetus	Human adult	Rat
SFRP1	Caspr2	30	/	32
CLU	NLGN2	18	157	/
APOE	NLGN2	/	12	/
CNTN1	NLGN1	/	11	59
TNR	NLGN1	33	76	351

Table 4.5: Summary of the five selected candidates. For each of them the predicted partner and the values observed in the corresponding column in the three brain samples are reported.

For the validation step, co-IP experiments were performed. The protocol used is the one described in par. 4.1.3, whose reliability was tested by the co-immunoprecipitation of NLGN1-639* and β NRXN-300* (fig. 4.6): this uses secreted proteins either co-expressed in the same cell or independently expressed. In some cases the proteins were tested also by means of ITC, which is a technique that uses soluble purified proteins. All the candidates that had to be tested are extracellular and all, except CNTN1, are naturally secreted from the cell: therefore, the full-length proteins could be used. CNTN1, instead, is a GPI-anchored protein, so a truncated form, which lacks the anchored residues and is consequently secreted, was studied. Moreover, while for co-IP purposes the fully glycosylated proteins were used, for ITC analysis purposes they were purified after the expression in HEK 293 Gnt1⁻ cells, a strain that lacks the N-acetylglucosaminyltransferase-1 activity and therefore produces proteins with a reduced number of N-glycans: these proteins were used to reduce the complexity of the interactions to be observed, focusing only on the peptide-peptide interactions.

In all the cases, the sequence encoding the tested proteins, lacking the portion corresponding to the leader peptide, was cloned in the multiple cloning site of a modified pCMV6-XL4 vector (fig. 4.10), between the NotI and XbaI sites. A protein expressed using this plasmid has the following features:

- Prolactin Leader Peptide (PLP): this is an efficient signal for the secretion of the protein and is removed when the protein is released;
- FLAG tag at the N-terminal;
- 3Cprotease cleavage sequence before a human IgG fc peptide at the C-terminal: this tag can be used for western blot analysis, pull-down experiments and purification. For the latter purpose the tag can be then removed using a 3Cprotease.

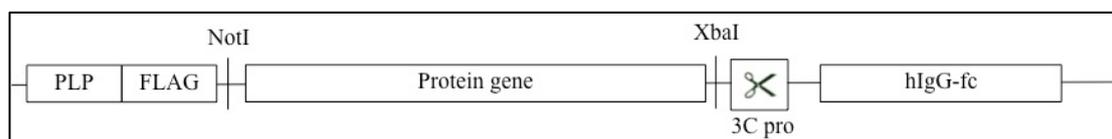


Figure 4.10: Scheme of the modified portion of pCMV-XL4 vector after the cloning of a hypothetical cDNA between the sites NotI and XbaI.

In the following sections the results of the experiments performed for the validation of the predicted associations are reported.

4.2.2 Secreted Frizzled Related Protein 1

In order to perform experiments on the binding ability of Secreted Frizzled Related Protein 1 (SFRP1) to verify its association to Caspr2, the coding sequence of the full-length protein was cloned in the pCMV6-XL4 vector (fig. 4.10). The resulting plasmid was tested, transfecting it in both HEK-293 and HEK-293 Gnt1⁻ cells and evaluating the levels of protein expression and secretion in the culture medium by means of western blotting. As shown in fig. 4.11A, both kinds of cells correctly expressed SFRP1. The protein was also properly secreted in the cell medium.

To test if SFRP1 is able to bind Caspr2, co-IP experiments were performed. The two plasmids encoding FLAG-Caspr2-1261* and FLAG-SFRP1-wt-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture media were collected, mixed together and incubated with a protein-A-bound resin to pull-down SFRP1. The precipitated proteins were then

evaluated by western blotting, using anti-FLAG antibodies. As shown in fig. 4.11B no Caspr2 protein was present after the pull-down, while it was present in the media and, therefore, correctly expressed. Even if the SFRP1 band is not visible in the lane corresponding to the media sample, the enrichment of the protein after the immunoprecipitation proves that it was correctly expressed, secreted and pulled-down. These results show that SFRP1 does not bind to Caspr2, as was instead predicted by the fishing data.

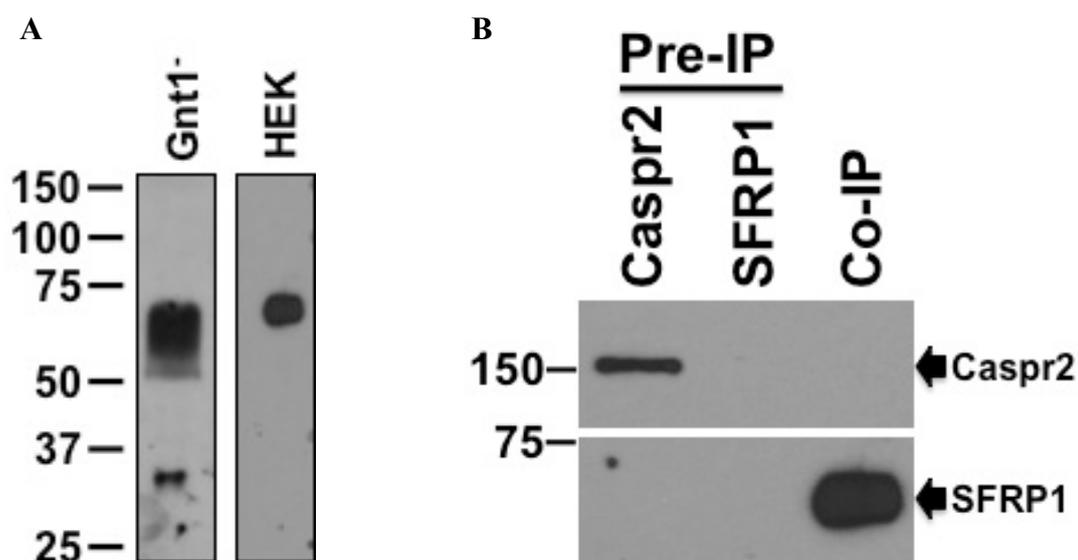


Figure 4.11: **A)** Check of expression and secretion of SFRP1 from HEK-293 and HEK-293 Gnt1⁻ cells. The cells were transfected with the pCMV6-XL4 plasmid encoding SFRP1 and after 24 hours the media were analyzed by western blotting using anti-human IgG antibodies. Expected molecular weight: 35 kDa for SFRP1 + 30 kDa for the IgG tag. The apparent molecular weight for SFRP1 expressed by HEK-293 cells is higher due to glycosylation; **B)** co-IP experiments for the couple of proteins Caspr2 and SFRP1. The two plasmids encoding FLAG-Caspr2-1261* and FLAG-SFRP1-wt-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture media were collected, mixed together and incubated with a protein-A-bound resin to pull-down SFRP1. The precipitated proteins were then evaluated by western blotting, using anti-FLAG antibodies.

To confirm this result, the purification of the protein was attempted in order to perform ITC experiments. For this purpose, a stable transfected HEK-293 Gnt1⁻ cell line expressing the FLAG-SFRP1-wt-IgG protein from the pCMV6-XL4 vector was generated. Even if the protein was correctly expressed and secreted, as shown in fig. 4.11A, no yield was obtained from the purification. This process involves an affinity chromatography, using a protein-A bound resin, followed by the elution of the protein upon digestion with 3Cprotease that separates the protein from the IgG tag. This

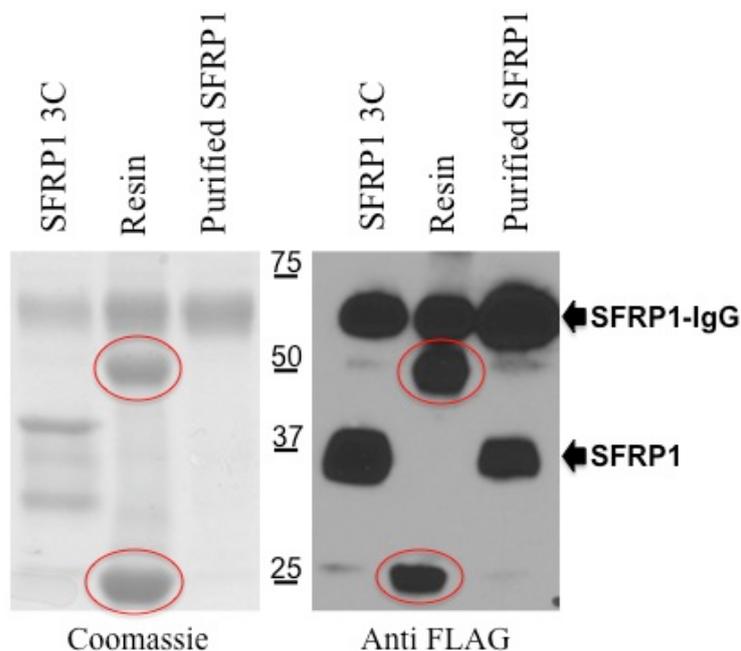


Figure 4.12: 3C protease digestion of SFRP1. Culture medium of HEK-293 Gnt1⁻ cells stably transfected with the pCMV6-XL4 plasmid encoding SFRP1 was collected. The protein was purified by a pull-down performed using an anti-FLAG resin and eluted with an excess of FLAG peptide. The purified protein was then incubated over night with a 3C protease enzyme. The day after, the protein samples of the purification resin, of the pre-cleavage protein and of the digested one were analyzed by SDS-PAGE stained with Coomassie blue and immuno-blotted with anti-FLAG antibodies. Red circles: heavy and light chains of the anti-FLAG antibodies. Expected molecular weight: 35 kDa for SFRP1 + 30 kDa for the IgG tag.

digestion is a very sensible step: a low efficiency would affect the final yield of the whole purification process. To check if this was the case, the protein was purified using FLAG antibodies and then the digestion was carried out. Samples from each step were analyzed by SDS-PAGE stained with coomassie blue and by western blotting. As shown in fig. 4.12, in all the samples a large amount of undigested protein could be found, meaning that the resin retained part of the protein and that the digestion process was not totally efficient, as other digestions could instead be carried out to completion (data not shown). The digested protein, instead, was present after purification both before and after the 3C protease treatment, reflecting probably a partial self-digestion of the protein: in any case, this band is missing in the coomassie stained gel for the untreated sample, probably due to an amount of protein below the detection limit. Altogether, these results prove that the digestion did take place, but with low efficiency, possibly explaining the low yield of purified protein. The IgG fused protein could not be used in ITC experiments. The purification of the protein

could have been carried out using its FLAG tag, but in that case the un-cleaved proteins should have been eliminated, for example by protein-A precipitation: this step would reduce the concentration of the protein in any case. Therefore, to perform ITC experiments the protein should be expressed fused only to a FLAG-tag and then purified.

In conclusion, even if the ITC experiments were not performed, the co-IP data did not confirm the binding with Caspr2 predicted by the fishing results.

4.2.3 Clusterin

Clusterin (CLU) had to be further studied to confirm its association to NLGN2. Therefore, the coding sequence of the full-length protein was cloned in the pCMV6-XL4 (fig. 4.10). To verify the correct expression of the protein, HEK-293 cells and HEK-293 Gnt1⁻ cells were transfected with the generated plasmid and then the medium was collected to evaluate the protein content through western blotting. As shown in fig. 4.13 both cell lines correctly expressed CLU, also delivering it to the cell medium. This protein has a molecular weight of 80 kDa when un-cleaved, but after maturation it is cleaved in two halves of 40 kDa each, connected by disulfide bonds: in this case, the C-terminal half looks bigger because it is fused to a 30 kDa IgG tag. Moreover, the HEK-293 Gnt1⁻ line performs a simplified glycosylation: therefore, since CLU is a heavily glycosylated protein, with carbohydrate accounting for approximately 20-25% of the total mass of the mature molecule (Stewart et al., 2007), it appears smaller when produced in this kind of cells.

To verify the potential binding of CLU to NLGN2, co-IP experiments were performed. The two plasmids encoding FLAG-NLGN2-616* and FLAG-CLU-wt-IgG were co-transfected in HEK-293 cells. After 24 hours the culture medium was collected and incubated with a protein-A-bound resin to pull-down CLU. The proteins were then eluted and analyzed through western blotting, using anti-FLAG antibodies. As shown in fig. 4.14, both the cleaved and un-cleaved form of CLU were correctly pulled-down in the assay and NLGN2 was co-precipitated as well, showing a binding activity of the two proteins. The unexpected result is that the band corresponding to NLGN2 is weaker after the pull-down in comparison to the medium. When a high affinity complex is precipitated, the involved proteins are expected to be more concentrated (as for NLGN1 and β NRXN1 in fig. 4.6). A reduction in the intensity of the band is then unusual and can reflect a low affinity interaction.

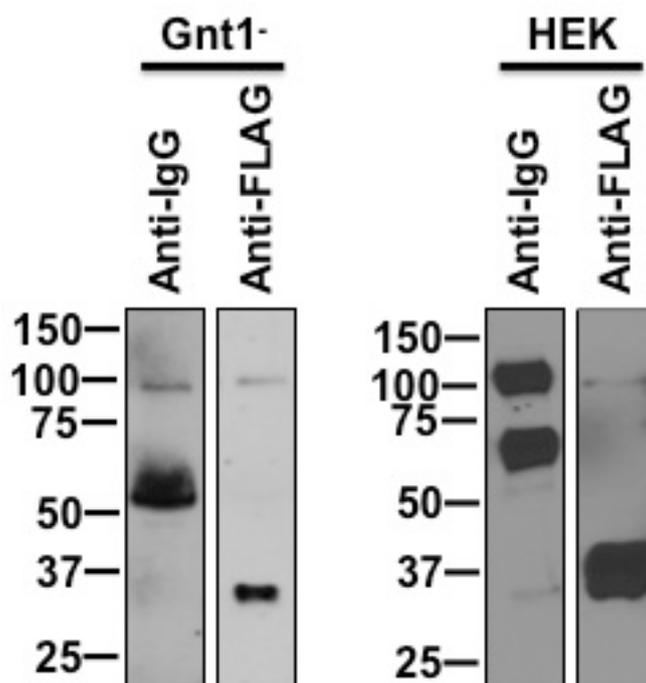


Figure 4.13: Check of expression and secretion of CLU from HEK-293 and HEK-293 Gnt1⁻ cells. The cells were transfected with the pCMV6-XL4 plasmid encoding CLU and after 24 hours the media were analyzed by western blotting using anti-FLAG or anti-human IgG antibodies. Expected molecular weight: 50 kDa for the uncleaved CLU, 25 kDa for each of the two halves, the N-terminal and the C-terminal, 30 kDa for the IgG tag. The apparent molecular weight for CLU expressed by HEK-293 cells is higher due to glycosylation.

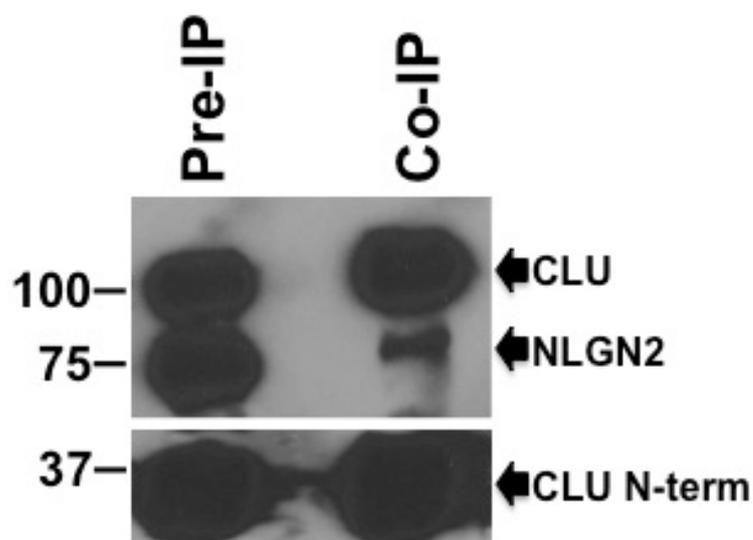


Figure 4.14: Co-IP experiments for the couple of proteins NLGN2 and CLU. The two plasmids encoding FLAG-NLGN2-616* and FLAG-CLU-wt-IgG were co-transfected in HEK-293 cells. After 24 hours the culture medium was collected and incubated with a protein-A-bound resin to pull-down CLU. The proteins were then eluted and analyzed through western blotting, using anti-FLAG antibodies.

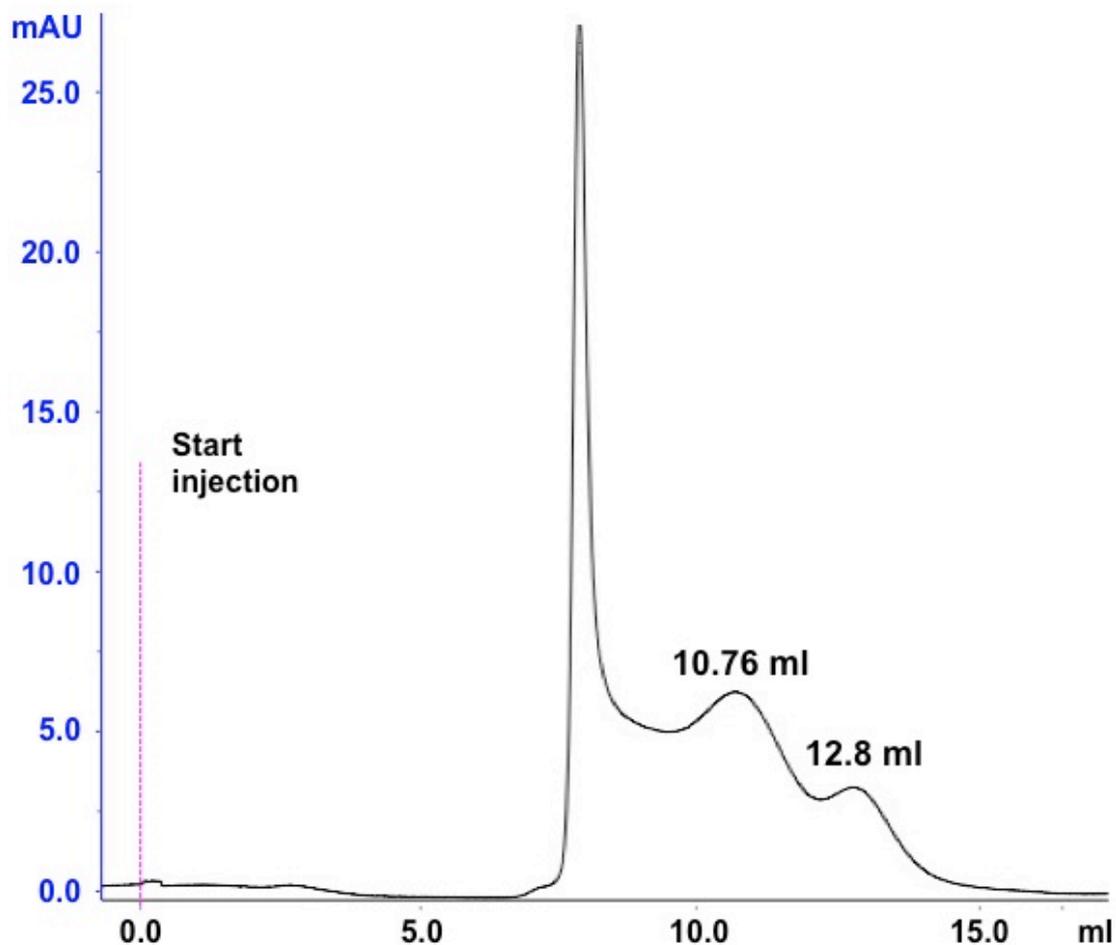


Figure 4.15: Chromatogram of the purified CLU protein. The elution volumes of the peaks are reported, corresponding to about 100 kDa (12.8 ml) and 200 kDa (10.76 ml) proteins. The first peak, instead, results from the aggregation of the concentrated protein.

To confirm the association between these two proteins, CLU and NLGN2 were purified in a form devoid of carbohydrates: even if sugars account for a large portion of the protein, when CLU is de-glycosylated it is still correctly folded and functioning (Stewart et al., 2007) and can therefore be used. For this purpose, a stable transfected HEK-293 Gnt1⁻ cell line expressing the FLAG-CLU-wt-IgG protein from the pCMV6-XL4 vector was generated. The secreted protein was purified from the culture medium using a protein-A affinity chromatography column. The elution step was carried out through the incubation of the bound resin with a GST-tagged 3C-protease: then the pure protein was concentrated. A sample of the purified protein was analyzed by means of FPLC. In fig. 4.15 the resulting chromatogram is reported. Two peaks can be observed at 10.76 and 12.8 ml of elution, probably corresponding to a dimeric and tetrameric form of the protein, due to their small volume of elution that

reflect proteins around 100 kDa and 200 kDa in weight. The first peak, instead, is most likely the result of the aggregation of the concentrated protein. In fact, it is known that CLU, especially in the de-glycosylated form, tends to aggregate in solution.

FLAG-NLGN2-616*, instead was expressed from the pCDNA3.1 in stable transfected HEK-293 Gnt1⁻ and purified from the culture medium using an anti-FLAG resin. A sample of the purified protein was analyzed by FPLC. In fig. 4.16 the chromatogram is reported: a single and symmetric peak was present at 12.08 ml of elution, corresponding to a protein about 120 kDa in mass, consistent with the dimeric form of NLGN2, whose truncated monomer has an expected molecular weight of about 68 kDa.

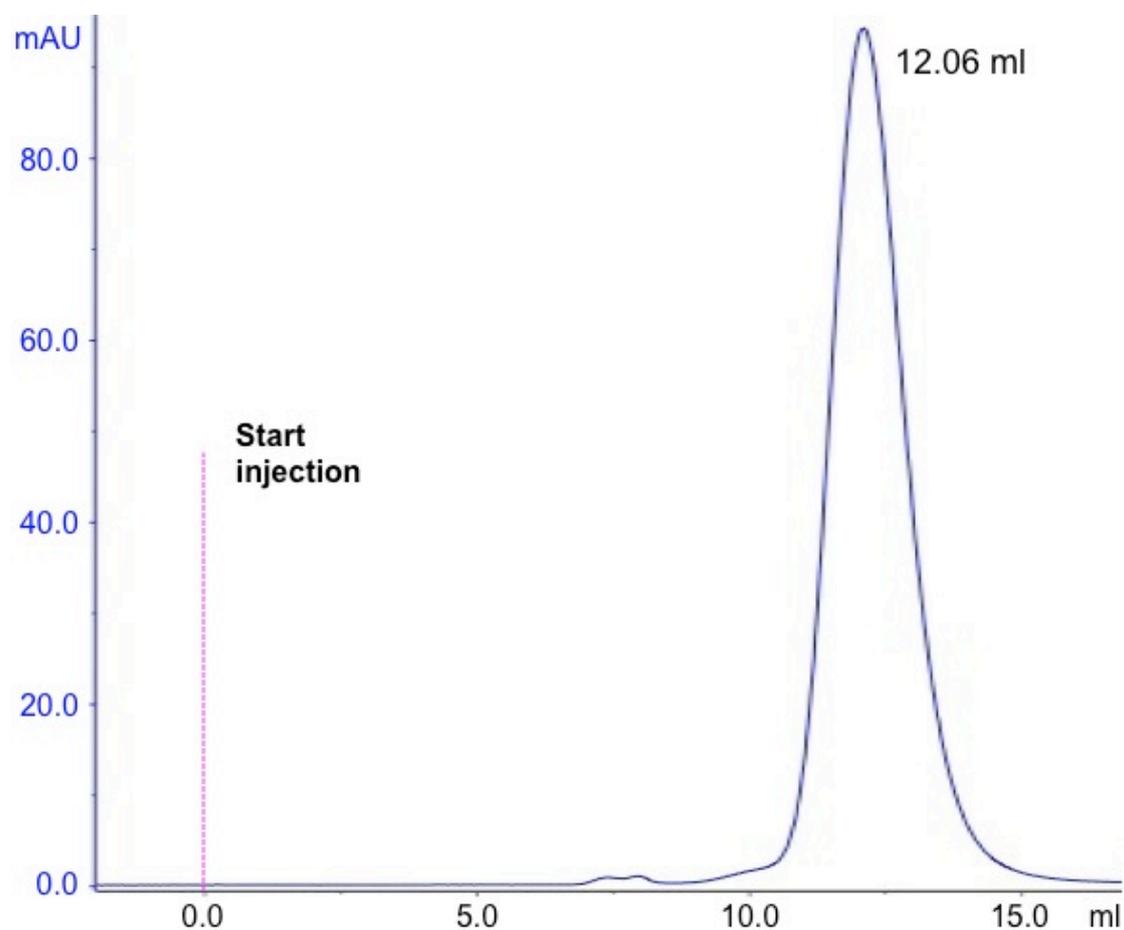


Figure 4.16: Chromatogram of the purified NLGN2 protein. The elution volume of the peak is reported, corresponding to about 122 kDa proteins.

The purified proteins were used to perform ITC experiments. CLU was used as the titrant at a concentration of 200 μM , while NLGN2 was concentrated to 22 μM . In fig. 4.17 the results are reported. The curve shows a binding activity of the two proteins, but with a relatively low affinity ($K_d = 1.2 \mu\text{M}$): this characteristic would be consistent with the co-IP results. Moreover, this experiment revealed an endothermic reaction for the association of these two molecules, instead of the exothermic one usually found for protein-protein interactions.

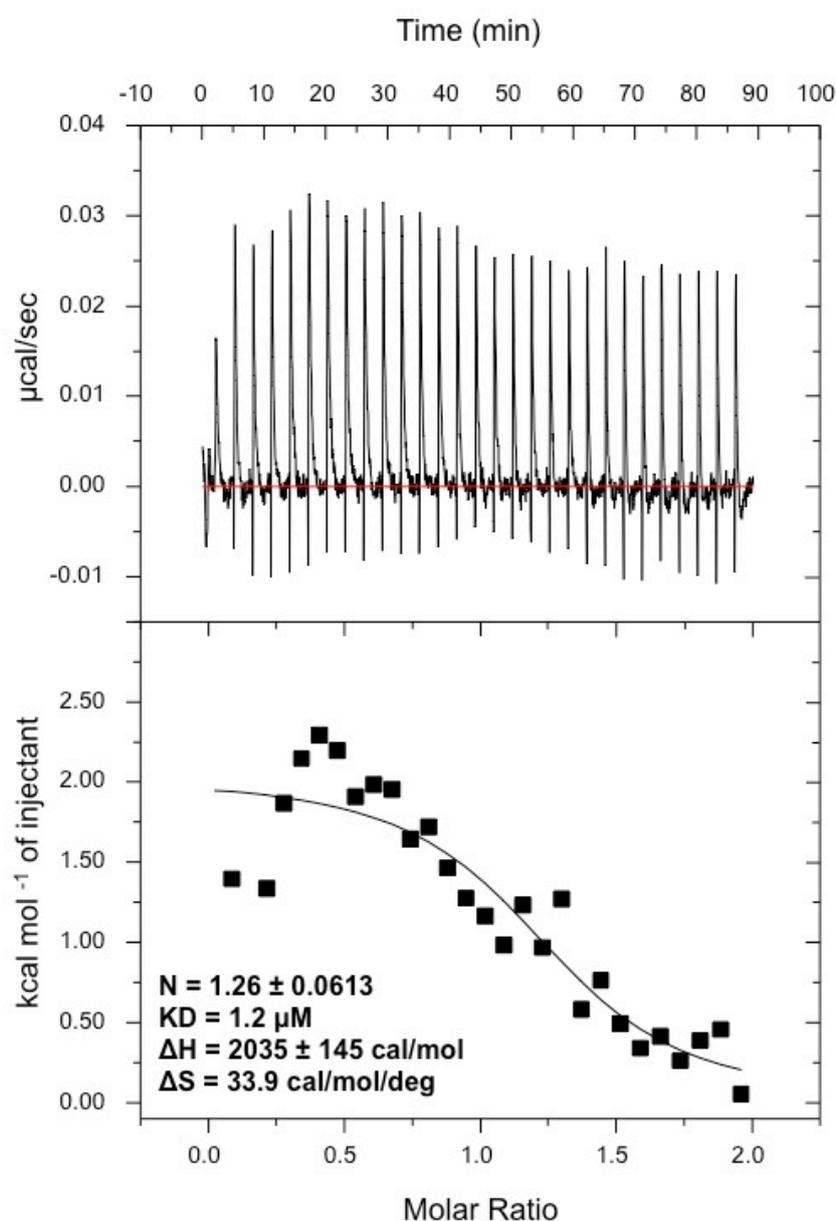


Figure 4.17: Results of the ITC analysis using the proteins CLU, used as titrant at a concentration of 200 μM , and NLGN2, used at a concentration of 22 μM . The curve in the top panel is the result of the difference between the CLU vs NLGN2 and CLU vs Buffer curves. Each point in the bottom panel is the integral of the corresponding spike in the top panel. The calculated values for stoichiometric ratio (N), K_d , ΔH and ΔS of the interaction are reported.

Altogether, these results confirm the binding of CLU and NLGN2, but some uncertainty still remains. For example, being that CLU is a molecular chaperone, could the binding be due to a partially unfolded population of the NLGN2 protein? Even if this was true, it would explain only the ITC results, obtained using purified and concentrated proteins, steps that can affect the folding status of a protein; the co-IP results, instead, would be only partially affected, because the binding happens in an ex-vivo assay conducted in a physiologic environment. In vivo experiments could definitely overcome this issue. For example, the full-length NLGN2 receptor fused to a FLAG-tag can be expressed in HEK cells and the binding with a CLU-IgG protein secreted by other cells, evaluated by a cell-surface immuno-labeling assay.

4.2.4 Apolipoprotein E

Apolipoprotein E (APOE) had to be further studied to confirm its association to NLGN2. Therefore, the coding sequence of the full-length protein was cloned in the pCMV6-XL4 (fig. 4.10). To verify the correct expression of the protein, HEK-293 cells and HEK-293 Gnt1⁻ cells were transfected with the generated plasmid and then the medium was collected to evaluate the protein content through western blotting. As shown in fig. 4.18A, APOE was correctly expressed and secreted by both cell lines.

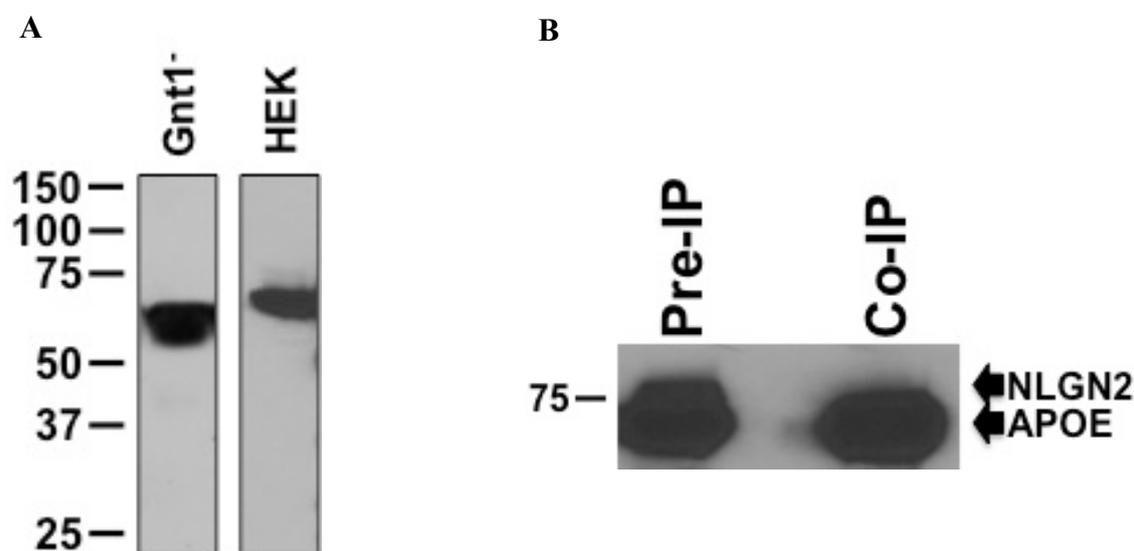


Figure 4.18: **A)** Check of expression and secretion of APOE from HEK-293 and HEK-293 Gnt1⁻ cells. The cells were transfected with the pCMV6-XL4 plasmid encoding APOE and after 24 hours the media were analyzed by western blotting using anti-human IgG antibodies. Expected molecular weight: 36 kDa for APOE + 30 kDa for the IgG tag. The apparent molecular weight for APOE expressed by HEK-293 cells is higher due to glycosylation; **B)** co-IP experiments for the couple of proteins NLGN2 and APOE. The two plasmids encoding FLAG-NLGN2-616* and FLAG-APOE-wt-IgG were co-transfected in HEK-293 cells. After 24 hours the culture medium was collected and incubated with a protein-A-bound resin to pull-down APOE. The proteins were then eluted and analyzed through western blotting, using anti-FLAG antibodies.

To verify the potential binding of APOE to NLGN2, co-IP experiments were performed. The two plasmids encoding FLAG-NLGN2-616* and FLAG-APOE-wt-IgG were co-transfected in HEK-293 cells. After 24 hours the culture medium was collected and incubated with a protein-A-bound resin to pull-down APOE. The proteins were then eluted and analyzed through western blotting, using anti-FLAG antibodies. In fig. 4.18B the results are presented. The FLAG-APOE-IgG protein has a molecular weight very close to the one of NLGN2. Even if, in theory, it would be possible to distinguish them, the western blot analysis of the proteins revealed bands too close to be distinguished: in any case, the doublet that seems to appear before the co-IP is no longer present after it, lacking the upper band, corresponding to NLGN2. To confirm that the co-IP experiment's result is negative, the immunoprecipitated protein should be run again using either a more concentrated gel or a lower amount of protein to increase the separation of the bands; another solution can be to clone the coding sequence of APOE in a vector not encoding the FLAG sequence, so the experiments could be performed using an APOE-IgG protein: this protein would have tags compatible with FLAG-NLGN2.

In order to perform ITC experiments to have more insights on the ability of APOE and NLGN2 to interact, the two proteins were purified. For this purpose a stable transfected HEK-293 Gnt1⁻ cell line expressing the FLAG-CNTN1-639*-IgG protein from the pCMV6-XL4 vector was generated. The secreted protein was purified from the culture medium using a protein-A affinity chromatography column. The elution step was carried out through the incubation of the bound resin with a GST-tagged 3C-protease: then the pure protein was concentrated. A sample of the purified protein was analyzed by means of FPLC. As it can be seen in fig. 4.19, the chromatogram presents two peaks: the first may correspond to aggregated protein, while the second, centered on 10.74 ml of elution volume corresponds to a complex with a total molecular weight of about 250 kDa. Because the expected molecular weight of APOE is 36 kDa, this reflects a multimeric form of APOE. The peak is not perfectly symmetric, suggesting that degraded forms of the protein can be also present.

FLAG-NLGN2-616*, instead was expressed from the pCDNA3.1 in stable transfected HEK-293 Gnt1⁻ and purified from the culture medium using an anti-FLAG resin as reported in par. 4.2.3 (see also fig. 4.16).

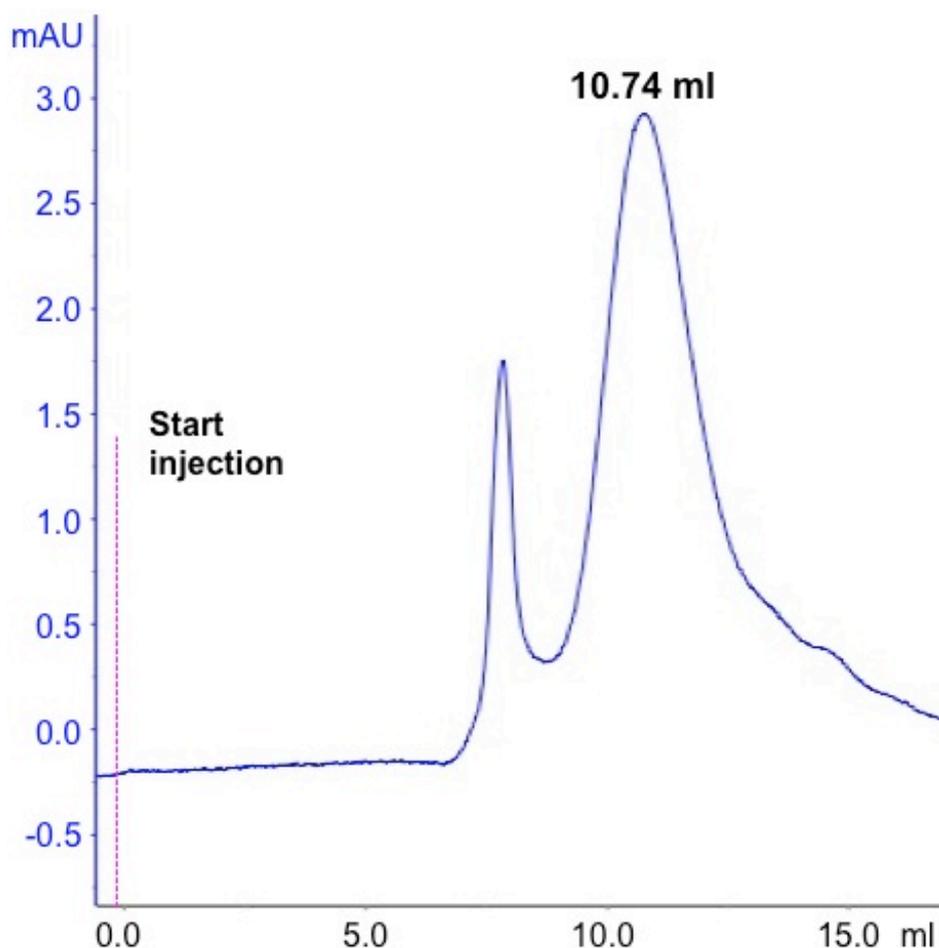


Figure 4.19: Chromatogram of the purified APOE protein. The elution volume of the peak is reported, corresponding to about 250 kDa proteins. The first peak, instead, results from the aggregation of the concentrated protein.

The purified proteins were used to perform ITC experiments. APOE was used as the titrant at the concentration of 120 μM , while NLGN2 was concentrated to 8 μM . In fig. 4.20 the results are reported: the points look randomly scattered, not fitting any statistical model, pointing-out that no binding can be observed using this method. Thus, even if, as already discussed, the co-IP results are not reliable, the ITC experiment failed to confirm an association between APOE and NLGN2.

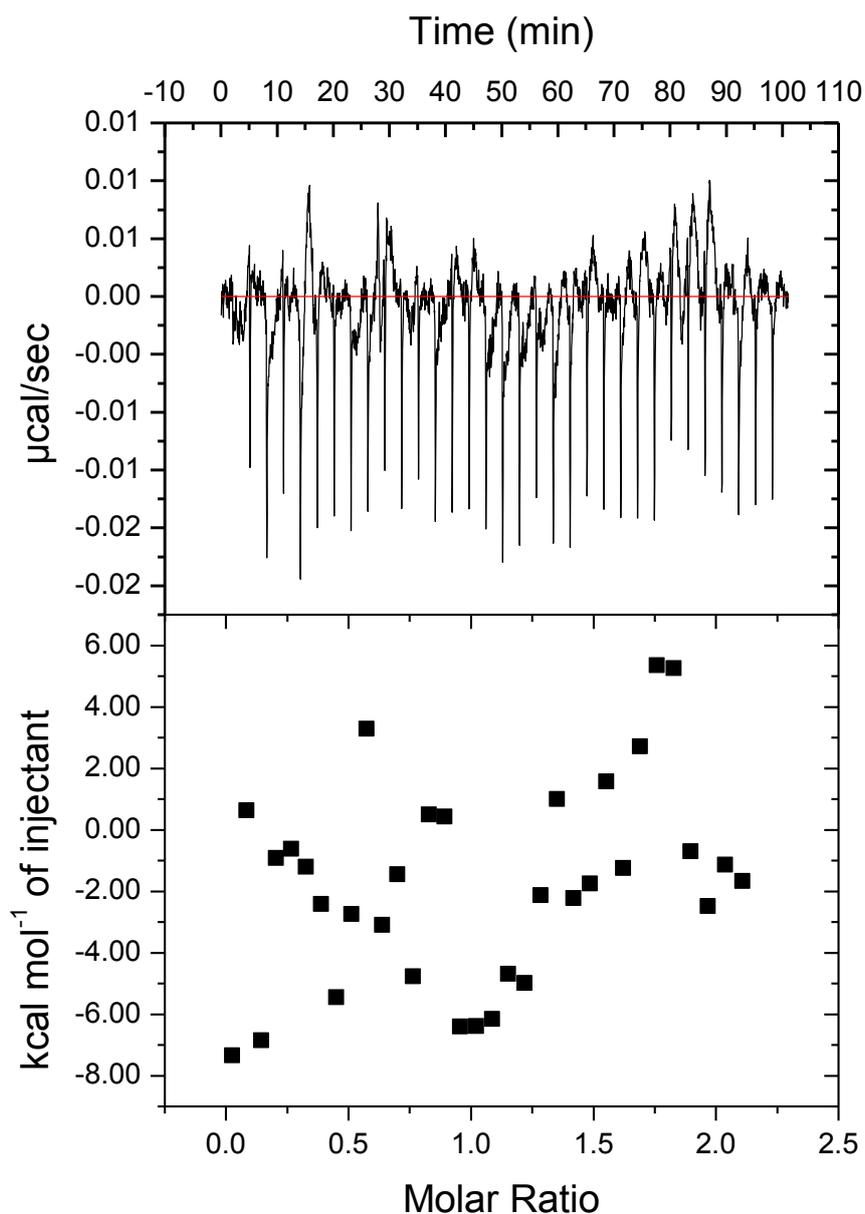


Figure 4.20: Results of the ITC analysis using the proteins APOE, used as titrant at a concentration of 120 μM , and NLGN2, used at the concentration of 8 μM . Each point in the bottom panel is the integral of the correspondent spike in the top panel. No association has been found between the two proteins.

4.2.5 Contactin 1

Due to DNA unavailability, the mouse Contactin 1 (CNTN1) was used to validate its binding to the human NLGN1. Both these proteins have been highly conserved during evolution, sharing a high homology between species: in fact an alignment of the protein sequences performed using the Lalign software provided on the expasy website (http://embnet.vital-it.ch/software/LALIGN_form.html) gives

95.8% of identity for the two CNTN1s and 96.6% of identity for the two NLGN1s. Even if the use of a human Contactin 1 would have been optimal, an association using these proteins was expected to be identified.

The coding sequence of CNTN1-986* protein was cloned in the pCMV6-XL4 vector (fig. 4.10). The resulting protein is truncated on the last residue of the last FN3 domain (the domain organization of CNTN1 is similar to the one of TAG-1/CNTN2; see also fig. 1.6), before the region where the GPI anchor is bound, resulting then in a secreted protein. The generated plasmid was transfected in both HEK-293 and HEK-293 Gnt1⁻ cells and the levels of protein expression and secretion in the culture medium evaluated by means of western blotting. As shown in fig. 4.21, both kinds of cells correctly expressed the protein, but multiple bands corresponding to CNTN1 appeared when the protein was produced and secreted by HEK-293 cells. On the contrary, only one band can be found when evaluating the CNTN1 content of the cell lysate, corresponding to the expected molecular weight of the protein, pointing out that the lower bands can be the product of some degradation event that happens in the media. Moreover, different band patterns could be seen when the western blot was performed using anti-FLAG or anti-human IgG antibodies, maybe due to a degradation happening on both sides of the protein. Nevertheless, these degraded forms of the protein account only for a small portion of the total amount.

To test if CNTN1 is able to bind NLGN1, co-IP experiments were performed. For this purpose the coding sequence of CNTN1-986* protein was cloned in a different version of pCMV6-XL4 vector, bearing only the IgG tag but not also the FLAG. This was necessary because some of the degradation forms of CNTN1, generated when the protein is secreted by HEK-293 cells, have the same molecular weight of NLGN1 and both the proteins have a FLAG tag. The two plasmids encoding FLAG-NLGN1-639* and CNTN1-986*-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture media were collected, mixed together and incubated with a protein-A bound resin to pull-down CNTN1. The precipitated proteins were then evaluated by western blotting, using anti-FLAG or anti-IgG antibodies. As reported in fig. 4.22, no band corresponding to NLGN1 is present after co-IP, even if it can be seen in the culture medium. As a control, CNTN1 was correctly expressed and pulled-down. Therefore, these results showed the lack of association between these proteins.

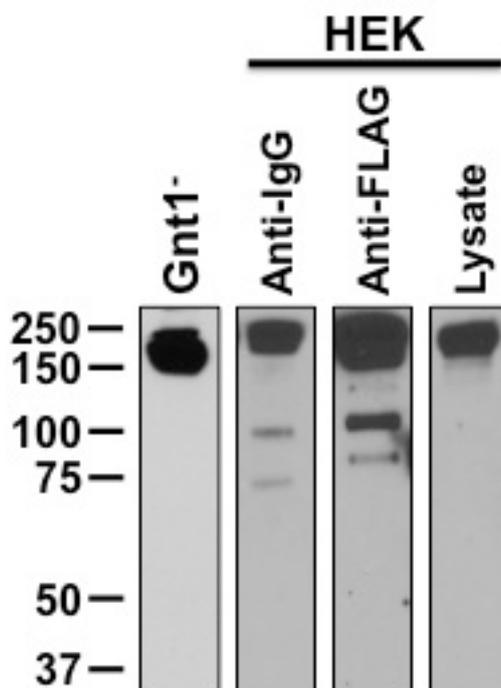


Figure 4.21: Check of expression and secretion of CNTN1 from HEK-293 and HEK-293 Gnt1⁻ cells. The cells were transfected with the pCMV6-XL4 plasmid encoding CNTN1 and after 24 hours the media or the lysate of the HEK-293 cells were analyzed by western blotting using anti-FLAG or anti-human IgG antibodies. Expected molecular weight: 110 kDa for CNTN1 + 30 kDa for the IgG tag. The apparent molecular weight for CNTN1 expressed by HEK-293 cells is higher due to glycosylation.

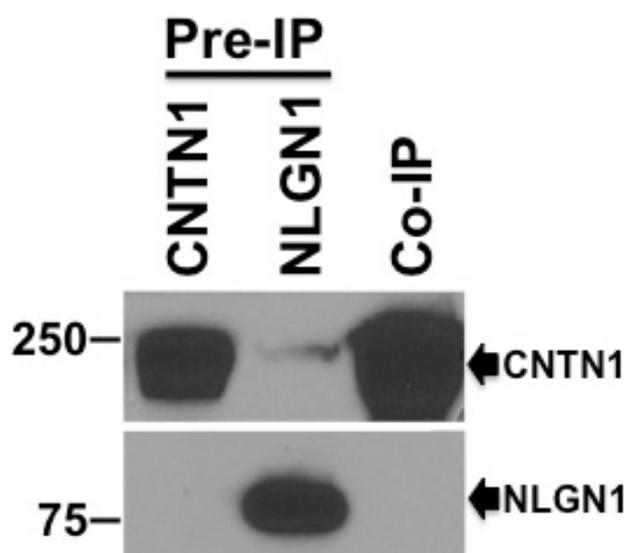


Figure 4.22: Co-IP experiments for the couple of proteins NLGN1 and CNTN1. The two plasmids encoding FLAG-Caspr2-1261* and FLAG-SFRP1-wt-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture media were collected, mixed together and incubated with a protein-A-bound resin to pull-down CNTN1. The precipitated proteins were then evaluated by western blotting, using anti-FLAG (to detect NLGN1) or anti-human IgG (to detect CNTN1) antibodies.

To verify the co-IP data, the CNTN1 and NLGN1 proteins were purified. For this purpose a stable transfected HEK-293 Gnt1⁻ cell line expressing the FLAG-CNTN1-639*-IgG protein from the pCMV6-XL4 vector was generated. The secreted protein was purified from the culture medium using a protein-A affinity chromatography column. The elution step was carried out through the incubation of the bound resin with a GST-tagged 3C-protease: then the pure protein was concentrated. A sample of the purified protein was analyzed by FPLC. In fig. 4.23 the resulting chromatogram is reported. Two peaks are present: the biggest one is centered at 11.2 ml of eluted buffer, corresponding about to 200 kDa, reflecting thus a probable dimeric form; the second is eluted earlier and can be related to a small quantity of aggregated protein. The non-perfect symmetry of the former peak can reflect a small amount of degraded protein.

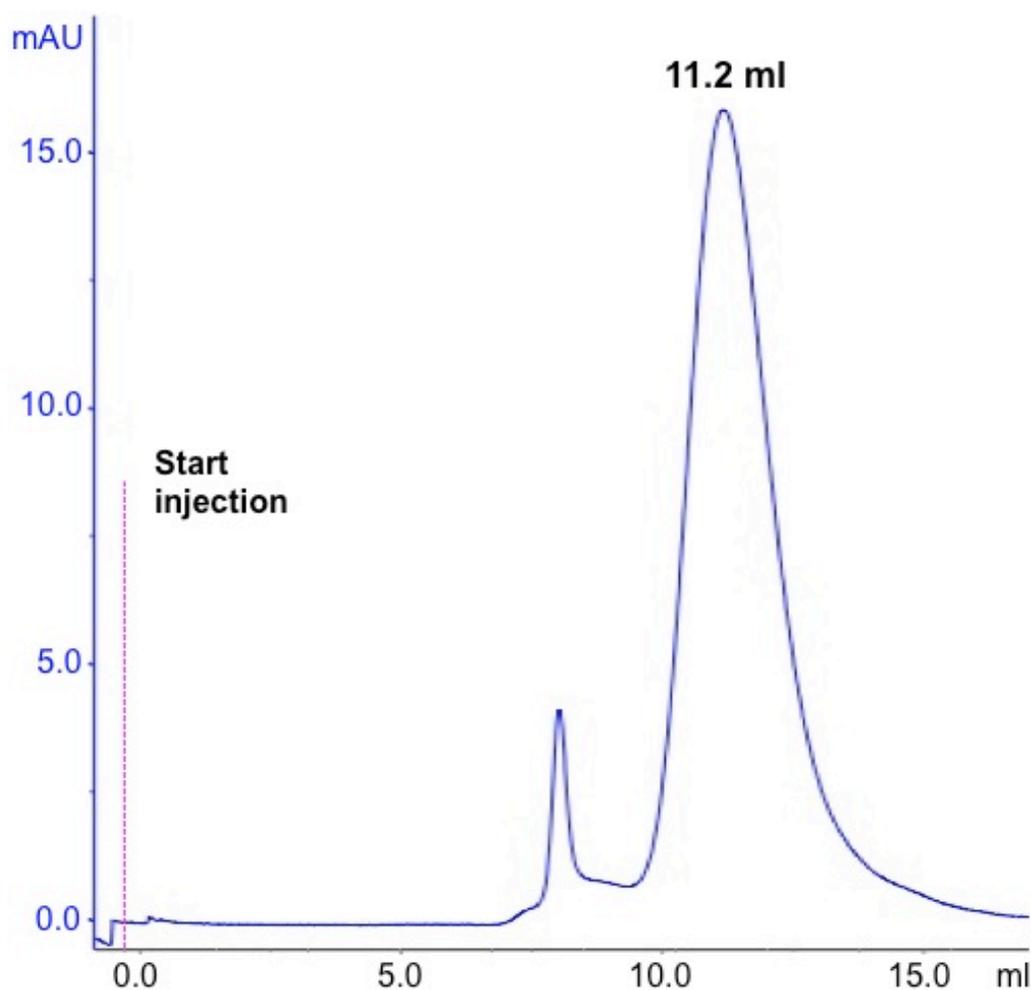


Figure 4.23: Chromatogram of the purified CNTN1 protein. The elution volume of the peak is reported, corresponding to about 200 kDa proteins. The first peak, instead, results from the aggregation of the concentrated protein.

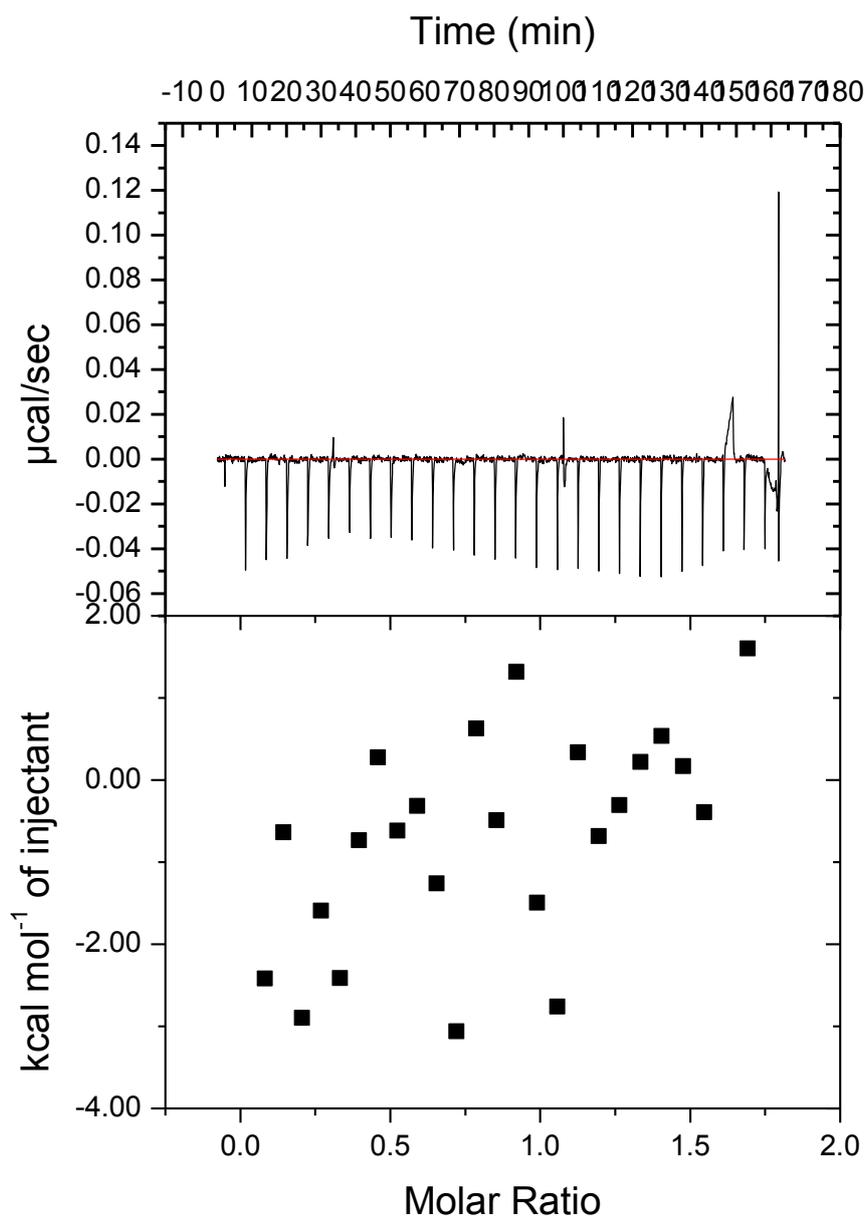


Figure 4.24: Results for the ITC analysis using the proteins CNTN1, used as titrant at a concentration of $82 \mu\text{M}$, and NLGN1, used at the concentration of $9.8 \mu\text{M}$. Each point in the bottom panel is the integral of the corresponding spike in the top panel. No association has been found between the two proteins.

FLAG-NLGN1-639*, instead, was expressed from the pCDNA3.1 in stable transfected HEK-293 Gnt1⁻ and purified from the culture medium using an anti-FLAG resin.

The purified proteins were used to perform ITC experiments. NLGN1 was used as the titrant at the concentration of $82 \mu\text{M}$, while CNTN1 was concentrated to

9.8 μM . In fig. 4.24 the results are reported: the points look randomly scattered, not fitting any statistical model, pointing-out that no binding can be observed using this method. These results, therefore, did not confirm the association between CNTN1 and NLGN1, as already emerged by co-IP analysis.

4.2.6 Tenascin-R

The last protein selected to confirm its predicted association was Tenascin R (TNR): the fishing data strongly suggest the binding of this protein to NLGN1. To test this prediction, the coding sequence of TNR was cloned in the pCMV6-XL4 (fig. 4.10). The expression and secretion of this protein was tested only in HEK-293 cells as its purification wasn't attempted. For this purpose, the cells were transfected with the generated plasmid and then the medium was collected to evaluate the protein content through western blotting. As shown in fig. 4.25A, TNR was correctly expressed and secreted. TNR has two major molecular forms, TNR 160 (160 kDa) and TNR 180 (180 kDa). Moreover, the two isoforms go through extensive process of glycosylation that raise their molecular weight (Anlar and Gunel-Ozcan, 2012, see also par. 1.4.4). Therefore, the doublet found here is consistent with what is reported in literature.

To test if TNR is able to bind NLGN1, co-IP experiments were performed. The two plasmids encoding FLAG-NLGN1-639* and FLAG-TNR-wt-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture media were collected, mixed together and incubated with a protein-A-bound resin to pull-down TNR. The precipitated proteins were then evaluated by western blotting, using anti-FLAG antibodies. As shown in fig. 4.25B, TNR was correctly expressed and both the 180 kDa and the 160 kDa forms were enriched after immunoprecipitation. NLGN1 was present in the culture medium, but the corresponding band is faint, at the limit of detection, while it's missing after the co-IP, apparently suggesting a lack of association. If the binding between the two proteins had high affinity, NLGN1 should be enriched, as TNR is, and the band should be visible; if the association had a low affinity instead, no enrichment would be expected, as happened for CLU and NLGN2. This means that the low amount of NLGN1 can mask some kind of interactions: therefore the experiment should be performed again, using a more concentrated sample. Moreover, the purification of the protein, in order to perform ITC experiments, has not been performed yet: other

interesting data, necessary to confirm or exclude a binding with NLGN1, would emerge from that analysis.

Altogether these data do not confirm, nor exclude the interaction of TNR with NLGN1 and point out that more experiments are needed.

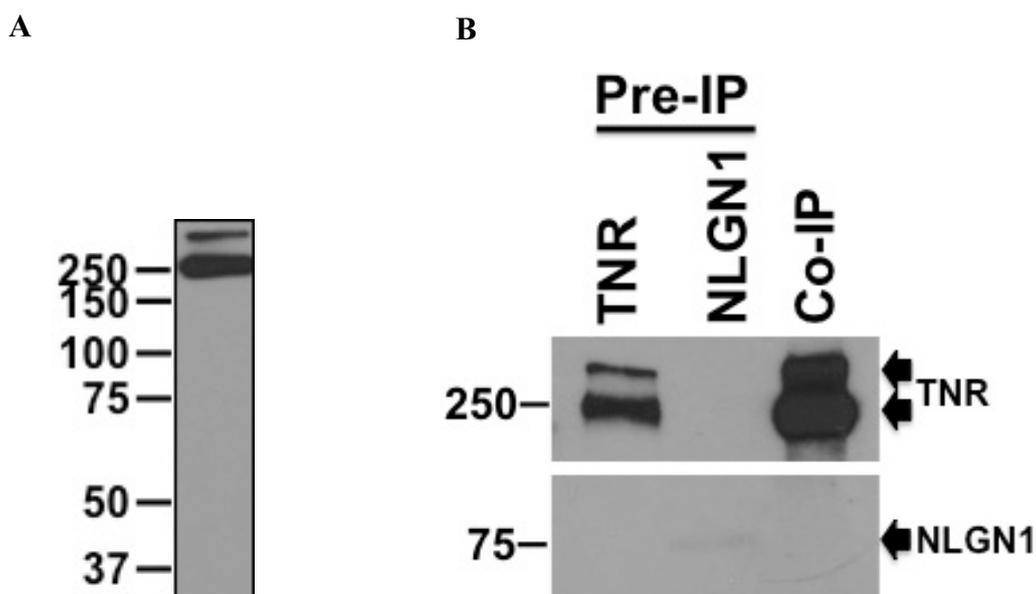


Figure 4.25: A) Check of expression and secretion of TNR from HEK-293 cells. The cells were transfected with the pCMV6-XL4 plasmid encoding SFRP1 and after 24 hours the medium was analyzed by western blotting using anti-human IgG antibodies. Expected molecular weight: 150 kDa for TNR + 30 kDa for the IgG tag. The apparent molecular weight is higher due to glycosylation; B) co-IP experiments for the couple of proteins NLGN1 and TNR. The two plasmids encoding FLAG-NLGN1-639* and FLAG-TNR-wt-IgG were independently transfected in HEK-293 cells cultured in two different plates. After 24 hours the two culture medium were collected, mixed together and incubated with a protein-A-bound resin to pull-down TNR. The precipitated proteins were then evaluated by western blotting, using anti-FLAG antibodies.

5. CONCLUSIONS

The aim of this thesis was to characterize new ligands of proteins involved in Autism Spectrum Disorder (ASD): among all the proteins associated with this disease the attention was focused on Caspr2 and NLGN2. A better understanding on the binding proprieties of these two molecules would shed new light on their function and help to understand their role in ASD.

Caspr2 is localized at the juxtaparanodes, where it promotes the clustering of the shaker like K^+ channel. In 2003, two groups proposed that the juxtaparanodal localization is guaranteed by the direct binding between Caspr2 and TAG-1 and since then several lines of research based their conclusions on this assumption. The evidence supporting this idea came from the fact that TAG-1 deficient mice have mislocalized Caspr2 proteins and vice versa. Moreover, co-immunoprecipitation (co-IP) experiments performed on brain lysates and co-immunoprecipitation experiments carried out using lysates from HEK or CHO cells transfected with plasmids encoding the two proteins seemed to confirm those results (Poliak et al., 2003; Traka et al., 2003; see also fig. 1.7 and 1.8). If the first two approaches only prove the involvement of Caspr2 and TAG-1 in the same macromolecular complex, the latter one pointed out a direct interaction between them. In particular, the two proteins were co-precipitated from the cell lysate only when they were co-expressed in the same cell, but not when they were expressed from different ones and then mixed together. These results were interpreted from the authors as a need for the Caspr2/TAG-1 complex to be formed in “cis”, or on the plasma membrane. Alternatively, the complex may have to be formed co-translationally in the endoplasmic reticulum (ER). A similar model was proposed for the Caspr1/CNTN1 complex: nevertheless, while Caspr1 cannot be translocated to

the plasma membrane and it is retained in the ER if it is not co-expressed with CNTN1 (Faivre-Sarrailh et al., 2000), Caspr2, when expressed alone, is correctly transported to the cell surface and also its truncated version, lacking the trans-membrane domain, can be released in the culture medium (Falivelli et al., 2012).

In this thesis, experiments to characterize the interaction between Caspr2 and TAG-1 were reported. Isothermal Titration Calorimetry approaches were used to calculate the affinity of the binding; moreover, the putative complex was analyzed by size exclusion chromatography. In all cases no association was observed for these proteins. Both these techniques use soluble proteins, purified independently from each other: therefore, the need for the complex to be formed in the ER can, in principle, explain the negative results obtained in these biophysical experiments.

To overcome this issue, pull-down experiments were performed co-expressing truncated proteins in HEK cells: these molecules lack the trans-membrane domains and are therefore released. Using this approach, the extracellular domains of Caspr2 and TAG-1, that are proposed to be responsible for the association, can be processed together in the ER and secreted as a complex in the culture medium. Nevertheless, also using this approach it was not possible to detect any association. Moreover, different, non-overlapping fragments of TAG-1 were used together with the extracellular domain of Caspr2 for pull down-experiments in HEK cells. As expected, the complex could be pulled-down from the cell lysates when the two proteins were co-expressed in the same cell, but not when they were expressed from different ones and then mixed together. All the fragments tested resulted to have the same potential to associate with Caspr2, similar also to the one of the full-length protein. This is an unexpected result: in fact, even if the binding involved all the extracellular domains of TAG-1, the binding is expected to be reduced when some portions are missing.

Altogether these results led to the idea that the protocol used to test the association generates false-positive artifacts. Therefore, pull-down experiments to validate the protocol were performed. The association between β NRXN1 and NLGN1 was used as a positive control, while the NLGN1-G500A mutant, which presents an abolished β NRXNs binding activity, was used as a negative control. Moreover, the associations of Caspr2 with unrelated proteins, such as Clusterin, CNTN1 (that should bind specifically Caspr1 and not Caspr2) and Secreted Frizzled Related Protein 1, were used as additional negative controls. These experiments showed that, when the binding was evaluated in cell lysates after co-expression, all the controls presented a

positive result, independently of their actual binding potential. When evaluated in cell lysates after independent transfections or in the cells' culture media, all the controls showed the expected results.

In addition, other unpublished data of the Comoletti lab proved that ER-resident chaperones, as for example Calnexin, can bridge the association between Caspr2 and TAG-1 when they are over-expressed in HEK cells. Altogether these data pointed out that the potential of artifact generation of the protocols that were used to test the direct association of the two proteins biased those experiments. The fact that, instead, no direct binding was detected using multiple other approaches, strongly indicates that this does not take place. Caspr2 and TAG-1, therefore, belong most likely to the same supramolecular complex, but their association is not likely direct. In fact, the flaws reported in the co-IP protocols using overexpressed proteins, would not affect, most likely, the pull-down experiments performed on brain lysates, nor the Caspr2 delocalization that was found in TAG-1 deficient mice.

TAG-1 was currently the only proposed protein associated to the extracellular portion of Caspr2. Therefore, the finding that they don't directly interact makes the latter a protein with no characterized ligands. Moreover, Caspr2 has been found in other localizations such as the axon initial segment (AIS) of pyramidal cells, a portion that unlike juxtaparanodes is devoid of myelin (Inda et al., 2006; Ogawa et al., 2008), and the synaptic terminals, even if it's not clear if at the pre- or post- synaptic side. In these locations Caspr2 can exert different functions, maybe binding different proteins. Another protein that can have uncharacterized molecular partners is NLGN2: as all the other NLGNs, this molecule is located in the post-synaptic terminals and interacts with the NRXNs family members. Nevertheless, NLGN2 is uniquely located at inhibitory synaptic terminals, has the lowest affinity for NRXNs and the lowest sequence similarity to the other members of its family. For all these reasons, a fishing experiment was performed: this consisted of an affinity chromatography using Caspr2 or NLGN2 as bait proteins, to capture their ligands in a brain lysate. The mass spectrometry analysis of the retained proteins can provide a list of candidates for the binding to these proteins. Among all the bound proteins, about ten likely candidates were found and three of them, chosen among those who had the highest chance to be actual ligands, were selected for further analysis: these were Secreted Frizzled Related Protein 1 (SFRP1), Clusterin (CLU) and Apolipoprotein E (APOE). Even if NLGN1 was included in the experiment only as a control for NLGN2, the results

emerging from the fishing using it as bait were interesting, with about eighteen candidates. As all the members of the NLGN family, also NLGN1 is associated to ASD and also the identification of new ligands for this protein would be interesting for the understanding of this disease. Therefore, other two candidates were selected for further analysis: Contactin 1 (CNTN1) and Tenascin R (TNR).

The second step for the identification of new ligands for Caspr2, NLGN1 and NLGN2 consisted of the confirmation of the proposed interactions of the selected candidates. This goal was achieved performing both ITC analysis and co-IP experiments. The latter were carried out using the protocol that was validated in this work. Out of five candidates, only the interaction of Clusterin with NLGN2 could be confirmed, with a reported affinity of 1.2 μ M and an endothermic binding reaction.

In the future, the binding between CLU and NLGN2 has to be better characterized, and if confirmed, the potential biological meaning will have to be investigated. Moreover, the remaining candidates identified by fishing have to be validated. Till now, only five out of about twenty-five were possible to be tested. The validation step is long because every protein has to pass through several steps, including the cloning of the gene, the stable expression in HEK cells and the purification. A fast and reliable test would improve the amount of data that can be generated in a small amount of time, even more if it is suitable for high throughput screenings. A technique that could be considered uses an immunocytochemistry approach to prove the binding in cell culture. By Expressing in HEK cells one of the proteins to test as a trans-membrane molecule and providing the hypothetical ligand in a soluble form, the binding can be evaluated by co-localization immunofluorescence: this technique would have the advantage of being suitable to be performed using transiently expressed proteins, saving the time of protein purification, that would be performed only in case of positive association to evaluate by means of ITC the binding parameters, such as K_d and molar ratio. Moreover, giving multiple soluble proteins at a time, all fused with the same tag, this technique can become suitable for high throughput screenings: the candidates could be processed in groups and in the case of a positive result the single proteins composing the group can be studied individually.

6. BIBLIOGRAPHY

- **Abrahams BS, Geschwind DH.** Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet.* **2008** May;9(5):341-55. Review. Erratum in: *Nat Rev Genet.* 2008 Jun;9(6):493
- **Abrahams BS, Tentler D, Perederiy JV, Oldham MC, Coppola G, Geschwind DH.** Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A.* **2007** Nov 6;104(45):17849-54
- **Alarcón M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, Nelson SF, Cantor RM, Geschwind DH.** Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet.* **2008** Jan;82(1):150-9
- **Amaral DG, Schumann CM, Nordahl CW.** Neuroanatomy of autism. *Trends Neurosci.* **2008** Mar;31(3):137-45. Epub 2008 Feb 6. Review
- **Anlar B, Gunel-Ozcan A.** Tenascin-R: role in the central nervous system. *Int J Biochem Cell Biol.* **2012** Sep;44(9):1385-9. Review
- **Araç D, Boucard AA, Ozkan E, Strop P, Newell E, Südhof TC, Brunger AT.** Structures of neuroligin-1 and the neuroligin-1/neurexin-1 beta complex reveal specific protein-protein and protein-Ca²⁺ interactions. *Neuron.* **2007** Dec 20;56(6):992-1003
- **Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A.** A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet.* **2008** Jan; 82 (1):160-4
- **Aspberg A, Miura R, Bourdoulous S, Shimonaka M, Heinegård D, Schachner M, Ruoslahti E, Yamaguchi Y.** The C-type lectin domains of lecticans, a family of aggregating chondroitin sulfate proteoglycans, bind tenascin-R by protein-protein interactions independent of carbohydrate moiety. *Proc Natl Acad Sci U S A.* **1997** Sep 16;94(19):10116-2
- **Aspberg A, Binkert C, Ruoslahti E.** The versican C-type lectin domain recognizes the adhesion protein tenascin-R. *Proc Natl Acad Sci U S A.* **1995** Nov 7;92(23):10590-4
- **Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators; Centers for Disease Control and Prevention.** Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *MMWR Surveill Summ.* **2012** Mar 30;61(3):1-19

- **Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, Tanriover G, Abrahams BS, Duvall JA, Robbins EM, Geschwind DH, Biederer T, Gunel M, Lifton RP, State MW.** Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am J Hum Genet.* **2008** Jan;82(1):165-73
- **Bang ML, Owczarek S.** A matter of balance: role of neuroligin and neuroligin at the synapse. *Neurochem Res.* **2013** Jun;38(6):1174-89. Review
- **Bel C, Oguievetskaia K, Pitaval C, Goutebroze L, Faivre-Sarrailh C.** Axonal targeting of Caspr2 in hippocampal neurons via selective somatodendritic endocytosis. *J Cell Sci.* **2009** Sep 15;122(Pt 18):3403-13
- **Blundell J, Blaiss CA, Etherton MR, Espinosa F, Tabuchi K, Walz C, Bolliger MF, Südhof TC, Powell CM.** Neuroligin-1 deletion results in impaired spatial memory and increased repetitive behavior. *J Neurosci.* **2010** Feb 10;30(6):2115-29
- **Blundell J, Tabuchi K, Bolliger MF, Blaiss CA, Brose N, Liu X, Südhof TC, Powell CM.** Increased anxiety-like behavior in mice lacking the inhibitory synapse cell adhesion molecule neuroligin 2. *Genes Brain Behav.* **2009** Feb;8(1):114-26
- **Bolliger MF, Frei K, Winterhalter KH, Gloor SM.** Identification of a novel neuroligin in humans which binds to PSD-95 and has a widespread expression. *Biochem J.* **2001** Jun 1;356(Pt 2):581-8
- **Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, Bailey A, Rutter M.** A case-control family history study of autism. *J Child Psychol Psychiatry.* **1994** Jul;35(5):877-900
- **Boucard AA, Chubykin AA, Comoletti D, Taylor P, Südhof TC.** A splice code for trans-synaptic cell adhesion mediated by binding of neuroligin 1 to alpha- and beta-neurexins. *Neuron.* **2005** Oct 20;48(2):229-36
- **Bouyain S, Watkins DJ.** The protein tyrosine phosphatases PTPRZ and PTPRG bind to distinct members of the contactin family of neural recognition molecules. *Proc Natl Acad Sci U S A.* **2010** Feb 9;107(6):2443-8
- **Bovolenta P, Esteve P, Ruiz JM, Cisneros E, Lopez-Rios J.** Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. *J Cell Sci.* **2008** Mar 15;121(Pt 6):737-46. Review
- **Budreck EC, Scheiffele P.** Neuroligin-3 is a neuronal adhesion protein at GABAergic and glutamatergic synapses. *Eur J Neurosci.* **2007** Oct;26(7):1738-48

- **Butz S, Okamoto M, Südhof TC.** A tripartite protein complex with the potential to couple synaptic vesicle exocytosis to cell adhesion in brain. *Cell*. **1998** Sep 18;94(6):773-82
- **Carnemolla B, Leprini A, Borsi L, Querzé G, Urbini S, Zardi L.** Human tenascin-R. Complete primary structure, pre-mRNA alternative splicing and gene localization on chromosome 1q23-q24. *J Biol Chem*. **1996** Apr 5;271(14):8157-60
- **Charles P, Tait S, Faivre-Sarrailh C, Barbin G, Gunn-Moore F, Denisenko-Nehrbass N, Guennoc AM, Girault JA, Brophy PJ, Lubetzki C.** Neurofascin is a glial receptor for the paranodin/Caspr-contactin axonal complex at the axoglial junction. *Curr Biol*. **2002** Feb 5;12(3):217-20
- **Chen X, Liu H, Shim AH, Focia PJ, He X.** Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions. *Nat Struct Mol Biol*. **2008** Jan;15(1):50-6
- **Chih B, Gollan L, Scheiffele P.** Alternative splicing controls selective trans-synaptic interactions of the neuroligin-neurexin complex. *Neuron*. **2006** Jul 20;51(2):171-8
- **Chih B, Engelman H, Scheiffele P.** Control of excitatory and inhibitory synapse formation by neuroligins. *Science*. **2005** Feb 25;307(5713):1324-8. Erratum in: *Science*. 2005 Jun 3;308(5727):1413
- **Chiu SY, Zhou L, Zhang CL, Messing A.** Analysis of potassium channel functions in mammalian axons by gene knockouts. *J Neurocytol*. **1999** Apr-May;28(4-5):349-64
- **Cifuentes-Diaz C, Chareyre F, Garcia M, Devaux J, Carnaud M, Levasseur G, Niwa-Kawakita M, Harroch S, Girault JA, Giovannini M, Goutebroze L.** Protein 4.1B contributes to the organization of peripheral myelinated axons. *PLoS One*. **2011**;6(9):e25043
- **Clarke RA, Lee S, Eapen V.** Pathogenetic model for Tourette syndrome delineates overlap with related neurodevelopmental disorders including Autism. *Transl Psychiatry*. **2012** Sep 4;2:e158. Review. Erratum in: *Transl Psychiatry*. 2012;2:e163
- **Comoletti D, Grishaev A, Whitten AE, Tsigelny I, Taylor P, Trehella J.** Synaptic arrangement of the neuroligin/beta-neurexin complex revealed by X-ray and neutron scattering. *Structure*. 2007 Jun;15(6):693-705
- **Comoletti D, Flynn RE, Boucard AA, Demeler B, Schirf V, Shi J, Jennings LL, Newlin HR, Südhof TC, Taylor P.** Gene selection, alternative splicing, and post-translational processing regulate neuroligin selectivity for beta-neurexins. *Biochemistry*. **2006** Oct 24;45(42):12816-27

- **Condro MC, White SA.** Distribution of language-related Cntnap2 protein in neural circuits critical for vocal learning. *J Comp Neurol.* **2014** Jan 1;522(1):169-85
- **Courchesne E, Pierce K.** Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Curr Opin Neurobiol.* **2005** Apr;15(2):225-30. Review
- **Dean C, Dresbach T.** Neuroligins and neuroligins: linking cell adhesion, synapse formation and cognitive function. *Trends Neurosci.* **2006** Jan;29(1):21-9
- **Dean C, Scholl FG, Choih J, DeMaria S, Berger J, Isacoff E, Scheiffele P.** Neuroligin mediates the assembly of presynaptic terminals. *Nat Neurosci.* **2003** Jul;6(7):708-16
- **De Silva HV, Harmony JA, Stuart WD, Gil CM, Robbins J.** Apolipoprotein J: structure and tissue distribution. *Biochemistry.* **1990** Jun 5;29(22):5380-9
- **Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN, Barnes KA, Dubis JW, Feczko E, Coalson RS, Pruett JR Jr, Barch DM, Petersen SE, Schlaggar BL.** Prediction of individual brain maturity using fMRI. *Science.* **2010** Sep 10;329(5997):1358-61. Erratum in: *Science.* **2010** Nov 5;330(6005):756
- **Elia J, Gai X, Xie HM, Perin JC, Geiger E, Glessner JT, D'arcy M, deBerardinis R, Frackelton E, Kim C, Lantieri F, Muganga BM, Wang L, Takeda T, Rappaport EF, Grant SF, Berrettini W, Devoto M, Shaikh TH, Hakonarson H, White PS.** Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol Psychiatry.* **2010** Jun;15(6):637-46. Erratum in: *Mol Psychiatry.* **2010** Nov;15(11):1122
- **Etherton MR, Blaiss CA, Powell CM, Südhof TC.** Mouse neuroligin-1alpha deletion causes correlated electrophysiological and behavioral changes consistent with cognitive impairments. *Proc Natl Acad Sci U S A.* **2009** Oct 20;106(42):17998-8003
- **Faivre-Sarrailh C, Gauthier F, Denisenko-Nehrbass N, Le Bivic A, Rougon G, Girault JA.** The glycosylphosphatidylinositol-anchored adhesion molecule F3/contactin is required for surface transport of paranodin/contactin-associated protein (caspr). *J Cell Biol.* **2000** Apr 17;149(2):491-502
- **Falivelli G, De Jaco A, Favaloro FL, Kim H, Wilson J, Dubi N, Ellisman MH, Abrahams BS, Taylor P, Comoletti D.** Inherited genetic variants in autism-related CNTNAP2 show perturbed trafficking and ATF6 activation. *Hum Mol Genet.* **2012** Nov 1;21(21):4761-73

- **Feng J, Schroer R, Yan J, Song W, Yang C, Bockholt A, Cook EH Jr, Skinner C, Schwartz CE, Sommer SS.** High frequency of neurexin 1beta signal peptide structural variants in patients with autism. *Neurosci Lett.* **2006** Nov 27;409(1):10-3
- **Folstein SE, Rosen-Sheidley B.** Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet.* **2001** Dec;2(12):943-55. Review
- **Fombonne E.** Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry.* **2005**;66 Suppl 10:3-8
- **Friedman JI, Vrijenhoek T, Markx S, Janssen IM, van der Vliet WA, Faas BH, Knoers NV, Cahn W, Kahn RS, Edelmann L, Davis KL, Silverman JM, Brunner HG, van Kessel AG, Wijmenga C, Ophoff RA, Veltman JA.** CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol Psychiatry.* **2008** Mar;13(3):261-6. Erratum in: *Mol Psychiatry.* 2010 Nov;15(11):1121
- **Frith C.** Is autism a disconnection disorder? *Lancet Neurol.* **2004** Oct;3(10):577. Review
- **Fu Y, Huang ZJ.** Differential dynamics and activity-dependent regulation of alpha- and beta-neurexins at developing GABAergic synapses. *Proc Natl Acad Sci U S A.* **2010** Dec 28;107(52):22699-704
- **Furley AJ, Morton SB, Manalo D, Karagogeos D, Dodd J, Jessell TM.** The axonal glycoprotein TAG-1 is an immunoglobulin superfamily member with neurite outgrowth-promoting activity. *Cell.* **1990** Apr 6;61(1):157-70
- **Geschwind DH, Levitt P.** Autism spectrum disorders: developmental disconnection syndromes. *Curr Opin Neurobiol.* **2007** Feb;17(1):103-11. Feb 1. Review
- **Gilbert M, Smith J, Roskams AJ, Auld VJ.** Neuroligin 3 is a vertebrate gliotactin expressed in the olfactory ensheathing glia, a growth-promoting class of macroglia. *Glia.* **2001** May;34(3):151-64
- **Gillespie DC, Kim G, Kandler K.** Inhibitory synapses in the developing auditory system are glutamatergic. *Nat Neurosci.* **2005** Mar;8(3):332-8
- **Graf ER, Zhang X, Jin SX, Linhoff MW, Craig AM.** Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell.* **2004** Dec 29;119(7):1013-26
- **Happé F, Ronald A.** The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychol Rev.* **2008** Dec;18(4):287-304. Review

- **Hatters DM, Peters-Libeu CA, Weisgraber KH.** Apolipoprotein E structure: insights into function. *Trends Biochem Sci.* 2006 Aug;31(8):445-54. Epub 2006 Jul 3. Review
- **Hauser PS, Narayanaswami V, Ryan RO.** Apolipoprotein E: from lipid transport to neurobiology. *Prog Lipid Res.* 2011 Jan;50(1):62-74. Review
- **Hegde RS, Bernstein HD.** The surprising complexity of signal sequences. *Trends Biochem Sci.* 2006 Oct;31(10):563-71
- **Howlin P, Goode S, Hutton J, Rutter M.** Adult outcome for children with autism. *J Child Psychol Psychiatry.* 2004 Feb;45(2):212-29
- **Ichtchenko K, Hata Y, Nguyen T, Ullrich B, Missler M, Moomaw C, Südhof TC.** Neuroligin 1: a splice site-specific ligand for beta-neurexins. *Cell.* 1995 May 5;81(3):435-43
- **Iijima T, Wu K, Witte H, Hanno-Iijima Y, Glatter T, Richard S, Scheiffele P.** SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. *Cell.* 2011 Dec 23;147(7):1601-14
- **Inda MC, DeFelipe J, Muñoz A.** Voltage-gated ion channels in the axon initial segment of human cortical pyramidal cells and their relationship with chandelier cells. *Proc Natl Acad Sci U S A.* 2006 Feb 21;103(8):2920-5
- **Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y, Rosahl TW, Südhof TC.** Binding of neuroligins to PSD-95. *Science.* 1997 Sep 5;277(5331):1511-5.
- **Jackman C, Horn ND, Molleston JP, Sokol DK.** Gene associated with seizures, autism, and hepatomegaly in an Amish girl. *Pediatr Neurol.* 2009 Apr;40(4):310-3
- **Jamain S, Quach H, Betancurá C, Råstam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T; Paris Autism Research International Sibpair Study.** Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet.* 2003 May;34(1):27-9
- **Jones SE, Jomary C.** Secreted Frizzled-related proteins: searching for relationships and patterns. *Bioessays.* 2002 Sep;24(9):811-20. Review
- **Just MA, Cherkassky VL, Keller TA, Minshew NJ.** Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain.* 2004 Aug;127(Pt 8):1811-21
- **Kana RK, Keller TA, Cherkassky VL, Minshew NJ, Just MA.** Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. *Brain.* 2006 Sep;129(Pt 9):2484-93

- **Kingston RE, Chen CA, Okayama H, Rose J.** Introduction of Plasmid DNA into Eukaryotic Cells, *Current Protocols in Molecular Biology* (1996) 9.1.1-9.1.11. Copyright © 2003 by John Wiley & Sons, Inc.
- **Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, Descartes M, Holt L, Braddock S, Troxell R, Kaplan L, Volkmar F, Klin A, Tsatsanis K, Harris DJ, Noens I, Pauls DL, Daly MJ, MacDonald ME, Morton CC, Quade BJ, Gusella JF.** Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet.* **2008** Jan;82(1):199-207
- **Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, O'Donovan MC, Erdogan F, Owen MJ, Ropers HH, Ullmann R.** Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet.* **2008** Feb 1;17(3):458-65
- **Knight D, Xie W, Boulianne GL.** Neurexins and neuroligins: recent insights from invertebrates. *Mol Neurobiol.* **2011** Dec;44(3):426-40. Review
- **Lam KS, Aman MG.** The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. *J Autism Dev Disord.* **2007** May;37(5):855-66
- **Laumonnier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, Raynaud M, Ronce N, Lemonnier E, Calvas P, Laudier B, Chelly J, Fryns JP, Ropers HH, Hamel BC, Andres C, Barthélémy C, Moraine C, Briault S.** X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet.* **2004** Mar;74(3):552-7
- **Lauritsen MB.** Autism spectrum disorders. *Eur Child Adolesc Psychiatry.* **2013** Feb;22 Suppl 1:S37-42. Review
- **Lawson-Yuen A, Saldivar JS, Sommer S, Picker J.** Familial deletion within NLGN4 associated with autism and Tourette syndrome. *Eur J Hum Genet.* **2008** May;16(5):614-8
- **Leone P, Comoletti D, Ferracci G, Conrod S, Garcia SU, Taylor P, Bourne Y, Marchot P.** Structural insights into the exquisite selectivity of neurexin/neuroligin synaptic interactions. *EMBO J.* **2010** Jul 21;29(14):2461-71
- **Levinson JN, Chéry N, Huang K, Wong TP, Gerrow K, Kang R, Prange O, Wang YT, El-Husseini A.** Neuroligins mediate excitatory and inhibitory synapse formation: involvement of PSD-95 and neurexin-1beta in neuroligin-induced synaptic specificity. *J Biol Chem.* **2005** Apr 29;280(17):17312-9.

- **Maley F, Trimble RB, Tarentino AL, Plummer TH Jr.** Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal Biochem.* **1989** Aug 1;180(2):195-204. Review
- **Marshall CR, Noor A, Vincent JB et al.** Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet.* **2008** Feb;82(2):477-88
- **Milev P, Chiba A, Häring M, Rauvala H, Schachner M, Ranscht B, Margolis RK, Margolis RU.** High affinity binding and overlapping localization of neurocan and phosphacan/protein-tyrosine phosphatase-zeta/beta with tenascin-R, amphoterin, and the heparin-binding growth-associated molecule. *J Biol Chem.* **1998** Mar 20;273(12):6998-7005
- **Missler M, Zhang W, Rohlmann A, Kattenstroth G, Hammer RE, Gottmann K, Südhof TC.** Alpha-neurexins couple Ca²⁺ channels to synaptic vesicle exocytosis. *Nature* **2003** Jun 26;423(6943):939-48
- **Missler M, Südhof TC.** Neurexins: three genes and 1001 products. *Trends Genet.* **1998** Jan;14(1):20-6. Review
- **Nakabayashi K, Scherer SW.** The human contactin-associated protein-like 2 gene (CNTNAP2) spans over 2 Mb of DNA at chromosome 7q35. *Genomics.* **2001** Apr 1;73(1):108-12
- **Nolen-Hoeksema S, Friedrickson BL, Loftus GR, Wagenaar WA.** Atkinson & Hilgard's Introduction to Psychology, 15th edition **2009**: chapter 15, Cengage Learning EMEA
- **Núñez S, Venhorst J, Kruse CG.** Target-drug interactions: first principles and their application to drug discovery. *Drug Discov Today.* **2012** Jan;17(1-2):10-22
- **Nuutinen T, Suuronen T, Kauppinen A, Salminen A.** Clusterin: a forgotten player in Alzheimer's disease. *Brain Res Rev.* **2009** Oct;61(2):89-104
- **Occhi G, Rampazzo A, Beffagna G, Antonio Danieli G.** Identification and characterization of heart-specific splicing of human neurexin 3 mRNA (NRXN3). *Biochem Biophys Res Commun.* **2002** Oct 18;298(1):151-5
- **Ogawa Y, Horresh I, Trimmer JS, Brecht DS, Peles E, Rasband MN.** Postsynaptic density-93 clusters Kv1 channels at axon initial segments independently of Caspr2. *J Neurosci.* **2008** May 28;28(22):5731-9
- **Panaitof SC, Abrahams BS, Dong H, Geschwind DH, White SA.** Language-related Cntnap2 gene is differentially expressed in sexually dimorphic song nuclei essential for vocal learning in songbirds. *J Comp Neurol.* **2010** Jun 1;518(11):1995-2018

- **Peles E, Salzer JL.** Molecular domains of myelinated axons. *Curr Opin Neurobiol.* **2000** Oct;10(5):558-65. Review
- **Peñagarikano O, Geschwind DH.** What does CNTNAP2 reveal about autism spectrum disorder? *Trends Mol Med.* **2012** Mar;18(3):156-63. Review. Erratum in: *Trends Mol Med.* 2012 Aug;18(8):502
- **Persico AM, Bourgeron T.** Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends Neurosci.* **2006** Jul;29(7):349-58. Review
- **Poliak S, Salomon D, Elhanany H, Sabanay H, Kiernan B, Pevny L, Stewart CL, Xu X, Chiu SY, Shrager P, Furley AJ, Peles E.** Juxtaparanodal clustering of Shaker-like K⁺ channels in myelinated axons depends on Caspr2 and TAG-1. *J Cell Biol.* **2003** Sep 15;162(6):1149-60
- **Poliak S, Gollan L, Martinez R, Custer A, Einheber S, Salzer JL, Trimmer JS, Shrager P, Peles E.** Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K⁺ channels. *Neuron.* **1999** Dec;24(4):1037-47
- **Prange O, Wong TP, Gerrow K, Wang YT, El-Husseini A.** A balance between excitatory and inhibitory synapses is controlled by PSD-95 and neuroligin. *Proc Natl Acad Sci U S A.* **2004** Sep 21;101(38):13915-20
- **Rathjen FG, Wolff JM, Chiquet-Ehrismann R.** Restrictin: a chick neural extracellular matrix protein involved in cell attachment co-purifies with the cell recognition molecule F11. *Development.* **1991** Sep;113(1):151-64
- **Rodriguez J, Esteve P, Weigl C, Ruiz JM, Fermin Y, Trousse F, Dwivedy A, Holt C, Bovolenta P.** SFRP1 regulates the growth of retinal ganglion cell axons through the Fz2 receptor. *Nat Neurosci.* **2005** Oct;8(10):1301-9
- **Rubenstein JL, Merzenich MM.** Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* **2003** Oct;2(5):255-67. Review.
- **Salmivirta M, Elenius K, Vainio S, Hofer U, Chiquet-Ehrismann R, Thesleff I, Jalkanen M.** Syndecan from embryonic tooth mesenchyme binds tenascin. *J Biol Chem.* **1991** Apr 25;266(12):7733-9
- **Santos SD, Iuliano O, Ribeiro L, Veran J, Ferreira JS, Rio P, Mulle C, Duarte CB, Carvalho AL.** Contactin-associated protein 1 (Caspr1) regulates the traffic and synaptic content of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-type glutamate receptors. *J Biol Chem.* **2012** Feb 24;287(9):6868-77

- **Savvaki M, Panagiotaropoulos T, Stamatakis A, Sargiannidou I, Karatzioula P, Watanabe K, Stylianopoulou F, Karagogeos D, Kleopa KA.** Impairment of learning and memory in TAG-1 deficient mice associated with shorter CNS internodes and disrupted juxtaparanodes. *Mol Cell Neurosci.* **2008** Nov;39(3):478-90
- **Scheiffele P, Fan J, Choih J, Fetter R, Serafini T.** Neuroligin expressed in nonneuronal cells triggers presynaptic development in contacting axons. *Cell.* **2000** Jun 9;101(6):657-69
- **Scott-Van Zeeland AA, Abrahams BS, Alvarez-Retuerto AI, Sonnenblick LI, Rudie JD, Ghahremani D, Mumford JA, Poldrack RA, Dapretto M, Geschwind DH, Bookheimer SY.** Altered functional connectivity in frontal lobe circuits is associated with variation in the autism risk gene CNTNAP2. *Sci Transl Med.* **2010** Nov 3;2(56):56ra80
- **Seidman CE, Struhl K, Sheen J, Jessen T.** Introduction of Plasmid DNA into Cells, *Current Protocols in Molecular Biology (1997)* 1.8.1-1.8.10. Copyright © 1997 by John Wiley & Sons, Inc.
- **Sheng M, Hoogenraad CC.** The postsynaptic architecture of excitatory synapses: a more quantitative view. *Annu Rev Biochem.* **2007**;76:823-47. Review
- **Song JY, Ichtchenko K, Südhof TC, Brose N.** Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses. *Proc Natl Acad Sci U S A.* **1999** Feb 2;96(3):1100-5
- **Stewart EM, Aquilina JA, Easterbrook-Smith SB, Murphy-Durland D, Jacobsen C, Moestrup S, Wilson MR.** Effects of glycosylation on the structure and function of the extracellular chaperone clusterin. *Biochemistry.* **2007** Feb 6;46(5):1412-22
- **Strauss KA, Puffenberger EG, Huentelman MJ, Gottlieb S, Dobrin SE, Parod JM, Stephan DA, Morton DH.** Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N Engl J Med.* **2006** Mar 30;354(13):1370-7
- **Südhof TC.** Neuroligins and neurexins link synaptic function to cognitive disease. *Nature.* **2008** Oct 16;455(7215):903-11. Review
- **Szatmari P, Paterson AD, Zwaigenbaum L, et al.** Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet.* **2007** Mar;39(3):319-28
- **Talebizadeh Z, Lam DY, Theodoro MF, Bittel DC, Lushington GH, Butler MG.** Novel splice isoforms for NLGN3 and NLGN4 with possible implications in autism. *J Med Genet.* **2006** May;43(5):e21

- **Traka M, Goutebroze L, Denisenko N, Bessa M, Nifli A, Havaki S, Iwakura Y, Fukamauchi F, Watanabe K, Soliven B, Girault JA, Karagogeos D.** Association of TAG-1 with Caspr2 is essential for the molecular organization of juxtaparanodal regions of myelinated fibers. *J Cell Biol.* **2003** Sep 15;162(6):1161-72
- **Traka M, Dupree JL, Popko B, Karagogeos D.** The neuronal adhesion protein TAG-1 is expressed by Schwann cells and oligodendrocytes and is localized to the juxtaparanodal region of myelinated fibers. *J Neurosci.* **2002** Apr 15;22(8):3016-24
- **Ullrich B, Ushkaryov YA, Südhof TC.** Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. *Neuron.* **1995** Mar;14(3):497-507
- **Ushkaryov YA, Petrenko AG, Geppert M, Südhof TC.** Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor and laminin. *Science.* **1992** Jul 3;257(5066):50-6
- **Vabnick I, Trimmer JS, Schwarz TL, Levinson SR, Risal D, Shrager P.** Dynamic potassium channel distributions during axonal development prevent aberrant firing patterns. *J Neurosci.* **1999** Jan 15;19(2):747-58
- **Varoquaux F, Aramuni G, Rawson RL, Mohrmann R, Missler M, Gottmann K, Zhang W, Südhof TC, Brose N.** Neuroligins determine synapse maturation and function. *Neuron.* **2006** Sep 21;51(6):741-54
- **Varoquaux F, Jamain S, Brose N.** Neuroligin 2 is exclusively localized to inhibitory synapses. *Eur J Cell Biol.* **2004** Sep;83(9):449-56
- **Verkerk AJ, Mathews CA, Joesse M, Eussen BH, Heutink P, Oostra BA; Tourette Syndrome Association International Consortium for Genetics.** CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. *Genomics.* **2003** Jul;82(1):1-9
- **Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcón M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE.** A functional genetic link between distinct developmental language disorders. *N Engl J Med.* **2008** Nov 27;359(22):2337-45
- **Volkmar F, Chawarska K, Klin A.** Autism in infancy and early childhood. *Annu Rev Psychol.* **2005**;56:315-36. Review
- **Werner E, Dawson G.** Validation of the phenomenon of autistic regression using home videotapes. *Arch Gen Psychiatry.* **2005** Aug;62(8):889-95
- **Woodworth A, Pesheva P, Fiete D, Baenziger JU.** Neuronal-specific synthesis and glycosylation of tenascin-R. *J Biol Chem.* **2004** Mar 12;279(11):10413-21

- **Yan J, Noltner K, Feng J, Li W, Schroer R, Skinner C, Zeng W, Schwartz CE, Sommer SS.** Neurexin 1alpha structural variants associated with autism. *Neurosci Lett.* **2008** Jun 27;438(3):368-70
- **Yan J, Oliveira G, Coutinho A, Yang C, Feng J, Katz C, Sram J, Bockholt A, Jones IR, Craddock N, Cook EH Jr, Vicente A, Sommer SS.** Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients. *Mol Psychiatry.* **2005** Apr;10(4):329-32
- **Zacharias U, Leuschner R, Nörenberg U, Rathjen FG.** Tenascin-R induces actin-rich microprocesses and branches along neurite shafts. *Mol Cell Neurosci.* **2002** Dec;21(4):626-33
- **Zafeiriou DI, Ververi A, Dafoulis V, Kalyva E, Vargiami E.** Autism spectrum disorders: the quest for genetic syndromes. *Am J Med Genet B Neuropsychiatr Genet.* **2013** Jun;162B(4):327-66. Review
- **Zahir FR, Baross A, Delaney AD, Eydoux P, Fernandes ND, Pugh T, Marra MA, Friedman JM.** A patient with vertebral, cognitive and behavioural abnormalities and a de novo deletion of NRXN1alpha. *J Med Genet.* **2008** Apr;45(4):239-43
- **Zhang C, Milunsky JM, Newton S, Ko J, Zhao G, Maher TA, Tager-Flusberg H, Bolliger MF, Carter AS, Boucard AA, Powell CM, Südhof TC.** A neuroligin-4 missense mutation associated with autism impairs neuroligin-4 folding and endoplasmic reticulum export. *J Neurosci.* **2009** Sep 2;29(35):10843-54
- **Zweier C, de Jong EK, Zweier M, Orrico A, Ousager LB, Collins AL, Bijlsma EK, Oortveld MA, Ekici AB, Reis A, Schenck A, Rauch A.** CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in *Drosophila*. *Am J Hum Genet.* **2009** Nov;85(5):655-66

ACKNOWLEDGMENTS

At the end of this work, I want to thank all the people that made it possible. First, I would like to thank Prof. Santi Mario Spampinato for letting me spend a year abroad. A special thank you is due to Prof. Davide Comoletti, that welcomed me in his lab and has been day by day a mentor, teaching me all that I needed to work on this project and advising me in all of its parts.

I'd also like to thank Eva N. Rubio Marrero for sharing some precious data (figures 4.3, 4.4, 4.5 and 4.6) and for working with me as a team during part of my stay in the Comoletti Lab.

A thank you goes to Sventja Von Daake for being a good teacher and for kindly offering her valuable help during the experiments.

Finally, I want to thank all the people in Comoletti lab for the support and the friendship that they showed me during my time in the US and especially Irene Pakos (Ειρήνη Πάκου) for her “special role” in my project and my life.

In a word: THANK YOU!