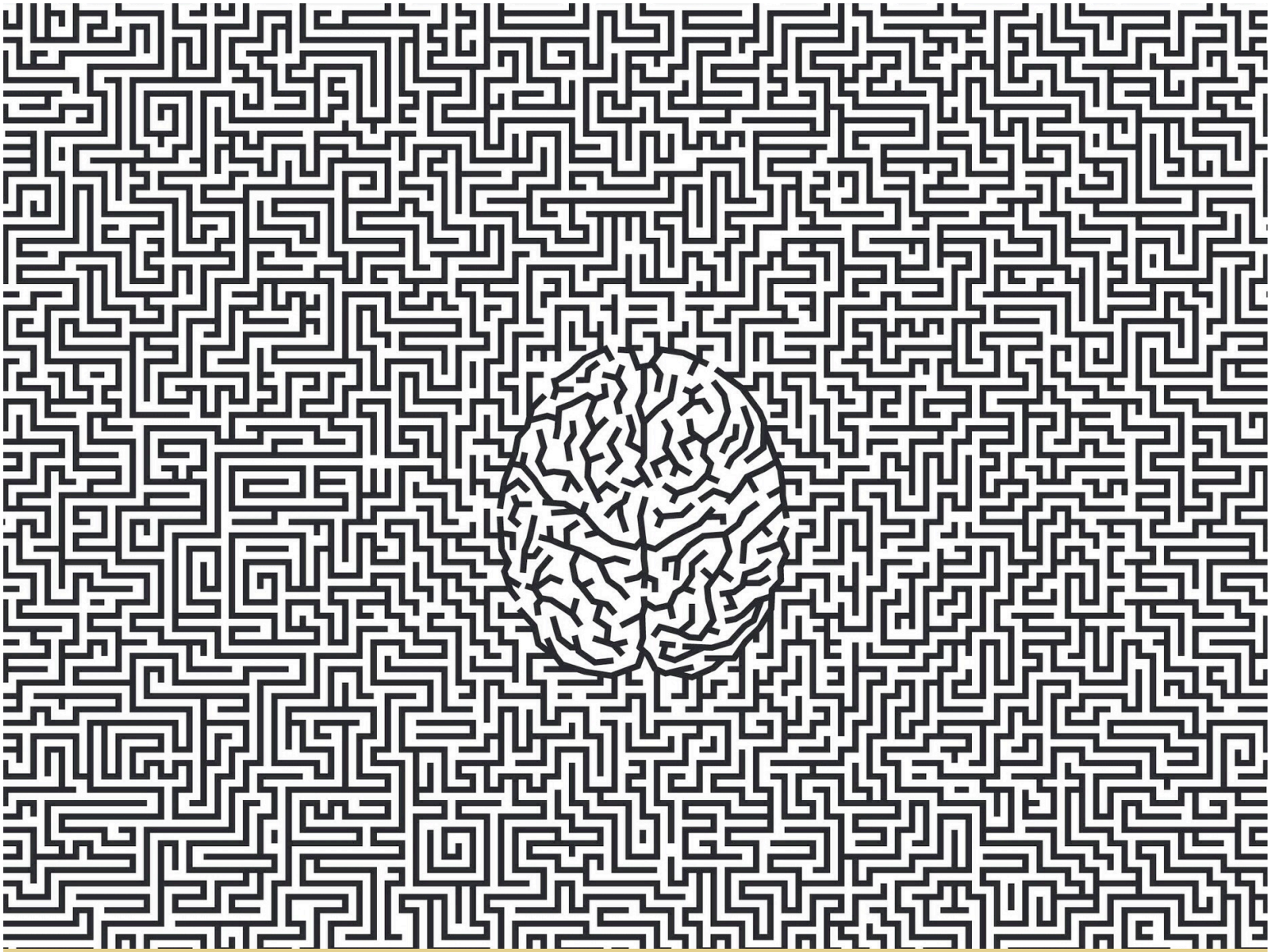


Ana Sofia Tremoceiro Lourenço

Biochemical and interactomic characterization of SAPAP3 – a scaffolding protein involved in obsessive-compulsive disorder

Tese de doutoramento em Biociências, especialização em Biotecnologia, orientada pelo Professor Doutor Euclides Manuel Vieira Pires e pelo Doutor Pedro Miguel Castanheira e apresentada ao departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia Universidade de Coimbra

Outubro de 2015

· U C ·

UNIVERSIDADE DE COIMBRA

Biochemical and interactomic characterization of SAPAP3 – a scaffolding protein involved in obsessive-compulsive disorder

Caracterização bioquímica e do interactoma da SAPAP3 – uma proteína de suporte envolvida no transtorno obsessivo-compulsivo

UNIVERSIDADE DE COIMBRA

Ana Sofia Tremoceiro Lourenço

Tese apresentada ao Departamento de Ciências da Vida (DCV) e à Faculdade de Ciências e Tecnologia da Universidade de Coimbra (FCTUC), para prestação de provas de doutoramento em Biociências, especialização em Biotecnologia.

Este trabalho foi realizado sob orientação do Professor Doutor Euclides Manuel Vieira Pires (Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia da Universidade de Coimbra) e pelo Doutor Pedro Miguel Castanheira (Biocant).

Outubro 2015

Agradecimentos

Eis que chega ao fim mais uma etapa. Não foi fácil, já o sabia. Surgiram vários obstáculos que levaram a uma mudança de direção. Mas com determinação e o apoio incondicional de pessoas inspiradoras foi possível atingir o objectivo. A todas elas gostaria de deixar o meu agradecimento.

Começo por agradecer ao Doutor Pedro Castanheira e ao Professor Doutor Euclides Pires por me terem recebido no laboratório e proporcionado as condições necessárias para a realização deste projeto. Em particular, agradeço ao Doutor Pedro Castanheira pela orientação, ajuda, partilha de conhecimentos, compreensão, disponibilidade, por se preocupar com o meu futuro, e por me ter proporcionado a possibilidade de uma experiência lá fora, o meu obrigada.

Agradeço à Doutora Isaura Simões o entusiasmo contagiante, a visão optimista com que encara a ciência, toda a ajuda, todos os conselhos, todas as discussões científicas, todas as conversas que me ajudaram a erguer e sem as quais este percurso teria sido bem mais difícil, muito obrigada.

I would like to thank Doctor Darren Hart and Doctor Philippe Mas for having received me so well in their lab. I would like to thank their availability and all the support with the ESPRIT experiments.

O meu agradecimento ao Doutor João Peça pela disponibilidade e pelos conselhos científicos que ajudaram a crescer este projeto. Agradeço também ao Mário Carvalho pela sua simpatia e pela preparação dos extratos de sinaptoneurossomas usados neste trabalho.

Ao Doutor Bruno Manadas agradeço a ajuda, a disponibilidade, a partilha de conhecimentos, e a confiança depositada ao longo de todos estes anos. Agradeço à Sandra Anjo pela ajuda nas experiências de interatómica; pela força e por ter partilhado comigo o lado bom e o lado menos bom da vida de um estudante de doutoramento.

O meu agradecimento especial à Carla Almeida por cuidar tão bem do nosso laboratório e manter tudo organizado, pela sua força, pela partilha de conhecimentos, por todos os conselhos e conversas, muito obrigado.

Ao Pedro Curto e ao Rui Cruz, que para além de colegas de bancada se tornaram verdadeiros amigos do peito, obrigada por toda a vossa força e apoio, obrigada pelas conversas, pela paródia, por me fazerem rir, por terem partilhado comigo os bons e os maus momentos, tanto dentro como fora do lab.

Agradeço à Ana Rita Leal, ao André Soares, à Marisa Simões, à Liliana Antunes e à Joana Furtado por terem sido os colegas de trabalho que qualquer pessoa pode desejar, obrigado pela boa disposição e pelo convívio, foi um prazer trabalhar com todos vocês.

À D. Alda quero deixar um beijinho muito especial e agradecer-lhe por me tratar de uma forma tão maternal, obrigada por se preocupar comigo e obrigada por aquele abraço aconchegante. Agradeço à Filipa, à D. Alda e à D. Adelaide a boa disposição, a galhofa das horas de almoço, e por cuidarem tão bem do nosso Biocant.

Um obrigada às meninas Vera Mendes, Cátia Santa, Sandra Anjo, Matilde Melo e Mariana Marcos da Unidade de Proteómica pelo convívio, e pela ajuda que sempre se disponibilizaram a prestar.

Quero deixar um agradecimento muito especial à Marta Vieira, à Ana Rita Santos, à Joana Ferreira, à Joana Fernandes e à Joana Pedro por todo o vosso apoio, por todas as palavras, por serem pessoas que admiro enquanto cientistas, pela vossa perseverança e determinação, pelos skypes, pelo carinho, pelas gargalhadas e claro pela vossa amizade.

Agradeço à Raquel Veiga, à Carla Saraiva, à Andreia Jorge, à Ana Margarida Pestana, à Flávia Vicente e ao Luís Santos por serem os amigos de sempre, amigos que a distância e o tempo não separa, por estarem sempre lá, por todas as conversas, por todos os choros, por todos os risos, pelo vosso apoio e por serem quem são, obrigada.

Agradeço à Cátia Mota, à Cristina Pereira, à Catarina Ferreira e à Rosa Baptista a vossa amizade, as palavras, as parvoíces, a cumplicidade, a compreensão, a vossa presença no meu dia-a-dia, o apoio constante; vocês estão no meu coração.

Um agradecimento muito especial à Isabel Febra e ao João Rosa, nem imaginam o quanto têm sido importantes neste meu percurso. Admiro-vos, admiro a vossa luta e a vossa forma de estar na vida. Obrigado por se preocuparem comigo, obrigada por se terem tornado parte do meu dia-a-dia, obrigado por me arrancarem de casa, e sobretudo obrigado pela vossa amizade.

Agradeço ao David Marques a boa energia que transmite, a boa disposição, toda a parvoíce que me faz rir, o optimismo, a tranquila forma de estar na vida e claro por ter de aturar não uma, mas de vez em quando duas manas Lourenço.

Um agradecimento muito muito muito especial aos meus queridos papás, Maria Anjos Lourenço e António Lourenço, pelo constante apoio, pela amizade, por acreditarem em mim, pelas palavras de conforto, por me terem transmitido os valores certos e por serem o pilar daquilo que sou hoje. Obrigado por tudo.

Agradeço à minha irmã, Paula Lourenço, ou melhor Paulinha, por ser uma das pessoas mais especiais da minha vida, por ser a minha melhor amiga, por ser a pessoa com quem mais me identifico, pela sua força e determinação, pela boa disposição, pelo seu sorriso, pelas suas palavras de conforto, pelo seu optimismo, porque apesar da distância está incondicionalmente presente, pela cumplicidade que nos une, por me chamar à razão quando necessário, por acreditar em mim, e simplesmente por ser quem é.

Termino por dizer que, para mim o melhor da vida são as pessoas que marcam a diferença, por isso do fundo do meu coração, muito obrigado a todos vocês.

Cover image: Brain maze image. Copyright: Grandeduc/Depositphotos.

This work was supported by European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grants N° 3419 for ESPRIT and N°7328 for HTX); by Fundação para a Ciência e Tecnologia (FCT): PTDC/NEU-NMC/0205/2012, and PEst-C/SAU/LA0001/2013-2014, and co-financed by "COMPETE Programa Operacional Factores de Competitividade", QREN, the European Union (FEDER – Fundo Europeu de Desenvolvimento Regional) and by The National Mass Spectrometry Network (RNEM) under the contract REDE/1506/REM/2005., and by FCT Ph.D. fellowship reference grant SFRH/BD/78585/2011.

Table of Contents

Resumo

O gene *SAPAP3* codifica uma proteína de suporte da densidade pós-sináptica, expressa em abundância na região do estriado, desempenhando funções importantes na regulação das sinapses glutamartérgicas cortico-estriatais. Os resultados obtidos a partir da caracterização de um modelo animal no qual este gene foi deletado sugerem o envolvimento da SAPAP3 na etiologia do transtorno obsessivo compulsivo. Os animais homozigóticos para a deleção deste gene apresentam elevados níveis de ansiedade, e é observado um comportamento ritualístico que leva a lesões na pele e à perda de pêlo na região facial. Ambos os comportamentos apresentam melhorias significativas após a administração de inibidores de recaptação de serotonina. Os animais mutantes revelam ainda alterações nas sinapses cortico-estriatais, parte do circuito envolvido no transtorno obsessivo compulsivo, e a expressão seletiva de SAPAP3 no estriado provoca a recuperação a nível sináptico e comportamental. Recentemente, esta proteína foi também associada a esquizofrenia, onde várias mutações foram identificadas em pacientes com esta patologia.

A produção da SAPAP3 em sistema heterólogo não foi descrita até ao momento, provavelmente devido ao seu elevado tamanho e à falta de domínios anotados, limitando a capacidade de obter informação relativa a esta proteína. Por forma a identificar domínios solúveis da SAPAP3 produzidos em *E. coli*, foi utilizada uma tecnologia designada ESPRIT ("expression of soluble proteins by random incremental truncation"). Usando esta tecnologia, 28 000 clones foram investigados tendo sido selecionados 96 para expressão em pequena escala. A partir desta informação, o domínio com maior rendimento de proteína solúvel – o domínio 19, correspondente ao C-terminal – foi produzido em larga escala e o seu protocolo de purificação optimizado. Depois de avaliada a qualidade da proteína purificada por técnicas analíticas, foram também iniciados ensaios preliminares de cristalização para determinação da estrutura tridimensional por difração de raios-X, tendo-se obtido putativos cristais em algumas das condições testadas.

Numa tentativa de obter novos dados relativos ao papel desta proteína, assim como elucidar os mecanismos moleculares que associam a SAPAP3 ao transtorno obsessivo compulsivo e esquizofrenia, foi realizada igualmente uma caracterização funcional avaliando o interactoma do domínio 19 da SAPAP3, assim como de dois mutantes desta proteína, um mutante envolvido no transtorno obsessivo-compulsivo (K910R) e outro em esquizofrenia (R770L). O resultado destas análises permitiram identificar uma associação entre a SAPAP3 e

componentes relacionados com a mitocôndria. Até ao momento este é o primeiro trabalho que descreve este tipo de associação, a qual poderá fornecer novos dados relativos ao envolvimento da SAPAP3 na patofisiologia do transtorno-obsessivo compulsivo e esquizofrenia.

Palavras chave: SAPAP3, transtorno obsessivo-compulsivo, esquizofrenia, ESPRIT e interatómica.

Abstract

SAPAP3 gene encodes a postsynaptic scaffolding protein, highly expressed in the striatum and important in regulating glutamatergic corticostriatal synapses. Results from a knockout animal model suggest that SAPAP3 may be involved in the pathophysiology of obsessive-compulsive disorder (OCD). Mice homozygous for the gene deletion exhibited increased anxiety and excessive grooming leading to facial hair loss and skin lesions; both behaviors were alleviated by administration of a serotonin uptake inhibitor. Additionally, the mutant mice reveals defects on corticostriatal synapses, part of the circuitry implicated in OCD, and the selective expression of SAPAP3 in the striatum rescues the synaptic and behavioral defects. More recently, human SAPAP3 was also associated with schizophrenia, with several missense mutations identified in patients suffering from such conditions.

SAPAP3 protein has not been able to be produced in heterologous systems, probably due to its large size and lack of annotated domains, limiting the ability to obtain more information regarding this protein. In this work, we used the high-throughput screening technology ESPRIT (expression of soluble proteins by random incremental truncation) to identify SAPAP3 soluble domains in *E. coli*. 28 000 clones were initially screened and 96 selected for small-expression assays. From there, one domain with higher yields of accumulation of soluble protein – C-terminal domain 19 – was produced in larger scale and protein purification methods optimized. After quality control assessment by analytical approaches, preliminary crystallization trials were started, in order to determine its three dimensional structure by x-ray diffraction methods, with several crystals being observed in some of the conditions tested.

To get new insights into the role of SAPAP3 and to help in the elucidation of the molecular mechanisms that associate SAPAP3 with OCD and schizophrenia, a functional characterization was performed, by the analysis of SAPAP3 domain 19 interactome, along with the interactome from two SAPAP3 mutants, one mutant associated with OCD (K910R) and another associated with schizophrenia (R770L). Results from these analysis revealed an association between SAPAP3 and mitochondria related components. To our knowledge, this is the first study presenting, what we believe to be, a novel role for SAPAP3 through the identified interaction with mitochondria components, which could also contribute to the understanding of the OCD and schizophrenia pathophysiology.

Key words: SAPAP3, obsessive-compulsive disorder, schizophrenia, ESPRIT and interactomics.

Abbreviations

A	AD	Activating domain
	ACC	Anterior cingulate cortex
	ACN	Acetonitrile
	AKAP	A-kinase anchor protein
	AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
	AP	Affinity purification
B	BAP	Biotin aceptor peptide
	BD	Binding domain
C	CD	Circular dichroism
	CSTC	Cortico-striato-thalamo-cortical
	CTB	Cognitive behaviour therapy
D	DDA	Data dependent acquisition
	DLC	Dynein light chain
	DSM-5	Manual of mental disorders 5
E	ERP	Exposure response prevention
	ESPRIT	Expression of soluble protein by random incremental truncation
	ExoIII	exonuclease III
F	FA	Formic acid
	FAK	Focal adhesion kinase
	FAT	Focal adhesion-targeting
	FDA	Food and drug administration
	FXM	Fragile X syndrome
G	GABA	Gamma-aminobutyric acid
	GK	Guanylate kinase
	GFP	Green fluorescent protein
	GO	Gene ontology
	GPe	Globus pallidus externa
	GPi	Globus pallidus pars internalis
I	ICAT	Isotope code affinity tags
	iTRAQ	Isobaric tags for relative and absolute quantification
L	LC	Liquid chromatography
	LUMIER	Luminescence-based mammalian interactome
M	MAGUIN	Membrane-associated guanylate kinase-interacting protein
	MAPPIT	Mammalian protein-protein interaction trap
	MBN	Mung bean nuclease
	mGluR	Metabotropic glutamate receptor

	MRI	Magnetic renosance imaging
	MS	Mass spectrometry
N	NMDA	N-methyl-D-aspartate
	nNOS	Nitric oxide synthase
O	OCD	Obsessive-compulsive disorder
	OFC	Orbitofrontal cortex
P	PCA	Principal component analysis
	PET	Positron emission tomography
	PPI	Protein-protein interaction
	PR	Proline-rich
	PSD	Postsynaptic density
	PYK2	Proline-rich tyrosine kinase
R	RF	Restriction free
S	SAPAP	Synapse-associated protein 90 (SAP90)/post synaptic density protein 95 (PSD95) - associated protein
	Shank	Src homology 3 (SH3) and multiple ankyrin repeat domain
	SILAC	Stable isotope labelling by amino acids
	SNr	Substantia nigra pars reticulate
	SNP	Single nucleotide polymorphism
	S-SCAM	Synaptic scaffolding molecule
	SSRIs	Selective serotonin reuptake inhibitors
	STAT	Signal transducers and activated of transcription
	STN	Subthalamic nucleus
	SWATH	Sequential window acquisition of all theoretical fragment ion spectra
T	TAP	Tandem affinity purification
	TCA	Tricyclic antidepressant
	TEV	Tobacco etch virus
W	WT	Wild-type
Y	Y2H	Yeast two-hybrid
	Y-BOCS	Yale-Browne obsessive-compulsive scale

1. Introduction

1.1 Obsessive-compulsive disorder

1.1.1 General description

Obsessive-compulsive disorder (OCD) is a chronic, neuropsychiatric condition that affects approximately 2 % of the world population[1-3], and significantly interferes with the person's normal functioning and social relationships[4].

OCD is characterized by the presence of obsessions and compulsions[5] – Obsessions are intrusive and unwanted thoughts, images, or impulses that cause considerable anxiety and distress[6], while compulsions are repetitive behaviours or mental acts performed by patients with OCD, in an attempt to neutralize the anxiety and/or the distress produced by the obsessions[6]. OCD is a symptomatically heterogeneous disorder, in which different forms of obsessions and compulsions exist, with examples of this condition including fears related with illness and germ contamination, unwanted aggressive thoughts, taboo thoughts involving sex and religion, the need for symmetry or exactness; leading to compulsions such as excessive cleaning, arranging, checking, counting, repeating, or reassurance seeking[7].

OCD has frequently an early onset, and a chronic and relapsing course could cause significant distress and disability[8]. Its pathophysiology is still not understood, and treatments available are only moderately effective, including cognitive behavioural psychotherapies and antidepressant drugs, which mechanisms of action are unknown[9]. More recently, an invasive procedures like deep brain stimulation had diversified achievement in improving severe symptoms[10], but this option remains far from the perfect solution, which highlights the importance of continuous research into the detailed causes of OCD for the development of safer and more effective treatments.

1.1.2 Epidemiology

OCD is more common than previously supposed, being the fourth most typical psychiatric disorder in the world with a lifetime prevalence of 2 %[1-3]. It is characterized by a bimodal distribution for age of onset, occurring either in childhood, with a mean age of onset at 10 years; or during adolescence/young adulthood, occurring in average at 21 years[11]. Onset after 35 is unusual[12]; and males tend to have an earlier age of onset than females[3]. In childhood, males are more commonly affected, whereas the opposite is observed in adulthood[1, 3]. Generally the symptoms onset is gradual, and if untreated, the course is

usually chronic with waxing and waning symptoms[6].

OCD has a high comorbidity with other anxiety and mood disorders, suggesting that some patients have impulsive behaviours, including symptoms of childhood conduct disorder and an increased rate of suicide attempts[13]. OCD is associated with considerable direct and indirect costs, as a result of the lack of recognition, under-diagnosis and inappropriate treatment[14].

1.1.3 Clinical features and diagnosis

Obsessive-compulsive disorder presents a heterogeneous symptomatology, in which several different types of obsessions and compulsions occur. Moreover, other mental disorders have similar phenomenological features, can be comorbid with OCD, or sometimes considered as uncommon forms of OCD[15] - these include body dysmorphic disorder, Tourette syndrome, tic disorders, trichotillomania and hoarding disorder[15].

To characterize clinical heterogeneity of OCD several studies based in factor analysis have been performed using the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) data[16, 17]. This scale is used in research and clinical practice to rate the severity of OCD symptoms by scoring: the time taken by obsessions and compulsions; the degree to which they interference with daily life; the subjective distress they cause; the individual's internal resistance to the obsessions and compulsions; and the degree of control to the obsessions and compulsions[17]. It comprises quantitative measures of avoidance, insight, indecisiveness, pathological responsibility, doubt and obsessional slowness. The scale also includes a checklist of around 70 symptoms of obsessions and compulsions such as hoarding, washing, checking, and fear of contamination[17].

Despite some variability among studies to identify a limited number of symptom clusters[18-29], factor analysis indicates that obsessions and compulsions tend to form five main symptom dimensions[30]: (1) symmetry obsessions and repeating, ordering and counting compulsions; (2) obsessions about being responsible for causing or failing to prevent harm and checking compulsion; (3) contamination obsessions and cleaning compulsions; (4) aggressive, sexual and religious obsessions and checking compulsions; and (5) hoarding obsessions and compulsions. These results contributed to the establishment of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) to help in the diagnosis of the disorder. The phenomenological heterogeneity associated with OCD

indicates that different clusters of symptoms could have distinct genetic and neuronal circuitry.

1.1.4 Treatments

The first line treatments to relieve the symptoms of OCD include cognitive behaviour therapy (CBT), pharmacotherapy using selective serotonin reuptake inhibitors (SSRIs), or a combination of both. Many authors suggested that a combined therapy is better to achieve an ideal treatment response than a single approach[31]. Unfortunately, 30 % of all patients do not respond adequately to these therapies and are considered resistant to the treatment[32].

Cognitive behaviour therapy consists in a psychotherapeutic method to address dysfunctional emotions, maladaptive behaviours, and cognitive processes through oriented and specific procedures[33]. The most effective cognitive behaviour therapy in OCD is exposure and response prevention (ERP). Exposure involves systematic, repeated, and continued confrontation with stimuli that create anxiety and the impulse to perform compulsive rituals[9]. In exposure the patients are confronted with images, thoughts, objects, and events leading to anxiety, and in response the patients are abstained from performing the compulsive act[33]. Some studies support the efficacy of this method in OCD, and about 80 % of the patients present a reduction in symptoms[34, 35]. For patients with severe symptoms, additional motivation interventions are sometimes necessary[9]. Although frequently successful, ERP generated anxiety in patients, and approximately 25 % of patients leave the treatment[34], for those who continue the ERP therapy, the effects usually last up to at least 2 years[36].

As already mentioned, efficient pharmacotherapies for OCD include a class of antidepressants called serotonin reuptake inhibitors[37]. Five compounds including clomipramine, a tricyclic antidepressant (TCA), and four selective serotonin reuptake inhibitors fluoxetine, fluvoxamine, paroxetine and sertraline were approved by FDA[6]. Clomipramine was the first approved compound for this indication, which is unique among the TCAs for its effective serotonergic effect[38, 39]. Linking the effectiveness of this compound in OCD, and its selectivity for serotonin led to the suggestion that this neurotransmitter could be implicated in the biology of OCD[39]. This hypothesis prompted the study of selective serotonin reuptake inhibitors in OCD, and today these drugs have become the first line of treatment in OCD, since they present less side effects and identical efficacy when compared with clomipramine[40].

SSRIs primarily block the serotonin transporter located in the plasma membrane of presynaptic neuron inhibiting neurotransmitter uptake[41]. Individual patients may respond well to one SSRI and not to another; the therapy requires at least 10-12 weeks, and the optimal doses typically exceed those that are used for major depression[39]. Nevertheless, SSRIs are only effective in about 40-60 % of patients[39]. In cases in which patients do not respond to the first line treatments, other therapeutic interventions have been proposed, including the use of typical pharmacologic drugs in higher dosages or their administration via an alternative way; combining drug treatments; and the use of novel compounds[42]. Augmentation strategies aimed to improve the efficacy of SRIs by the administration of other class of drugs. Antipsychotics (risperidone, quetiapine, haloperidol) acting as dopamine receptors antagonists are currently the pharmacotherapy augmentation strategy of choice[43]. Other molecules including glutamatergic agents such as NMDA receptor agonist D-cycloserine[9]; or several antidepressants other than SSRIs or clomipramine, as alpha-2 receptor and serotonin receptor 5-$HT_{2/3}$ antagonists, and monoamino-oxidase inhibitor phenelzine are being proposed[38]. For extremely severe cases not responding to pharmacological or psychological treatments, neuromodulation therapies such as deep brain stimulation have produced promising results[10]. With this strategy, implemented electrodes are continuously directing electrical current to the commonly targeted OCD brain regions including the anterior limb of the internal capsule, the nucleus accumbens and ventral striatum, and the subthalamic nucleus[44]; however mixed success in alleviating severe symptoms were obtained with this procedure[45].

Despite the advances concerning the treatment of OCD, a high percentage of patients do not respond to the available therapies, so a deeper understanding of OCD neurobiological mechanisms is fundamental to reach more successful treatments.

1.1.5 Neurobiology

Presently, the specific causal factors for OCD are not known, although converging evidences suggest abnormalities within the cortico-striato-thalamo-cortical (CSTC) brain circuitry in its etiology[45-47]. This circuit is involved in motor cognitive, and affective processes, and is presumed to be involved with the origin of OCD symptoms[48, 49].

CSTC circuitry of the healthy individuals[30, 47, 50] (Figure 1.1), cortical neurons send signals to striatum through glutamatergic cortico-striatal synapses; in turn, the striatum connect the globus pallidus pars internalis (GPi) and the substantia nigra pars reticulate (SNr) by two

distinct pathways - direct and indirect pathways - which are thought to have distinct roles. Through the so-called direct pathway, striatal activation increases inhibitory GABA signals to GPi/SNr, which decreases the inhibitory GABA output from GPi/SNr to thalamus, resulting in excitatory glutamatergic output from the thalamus to the frontal cortex, and thereby promoting the selection of behavioural sequences. In the indirect pathway, the striatum inhibits the globus pallidus externa (GPe), which decreases its inhibition of subthalamic nucleus (STN). The STN can then excite the GPi/SNr and thereby inhibit the thalamus, inhibiting in this way the selection of behavioural sequences. In OCD, the excessive activity of the direct pathway over the indirect pathway leads to hyper-activation of cortical pathway. In this way, exaggerated concerns about danger, hygiene or harm (mediated by cortex) could result in persistent obsessions[30].

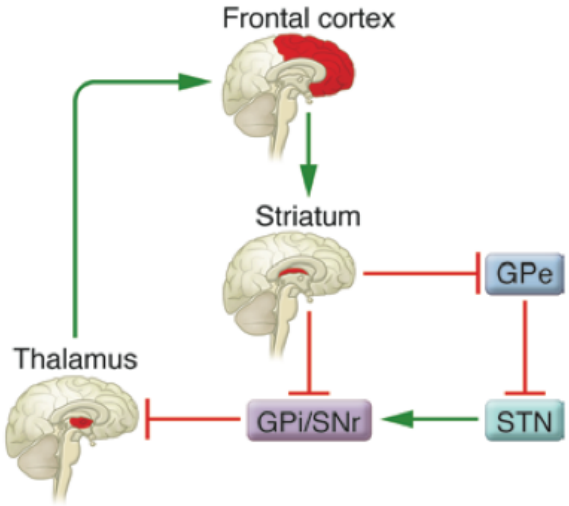


Figure 1.1 – The cortico-striato-thalamo-cortical (CSTC) circuit in a healthy individual. Simplified circuit diagram of the CSTC circuit in mammals. Green arrows depict glutamate (excitatory) pathways and red lines GABAergic (inhibitory) pathways. GPe: globus pallidus pars externa; GPi: globus pallidus pars internalis; SNr: substantia nigra pars reticulate; SPN: subthalamic nucleus. Adapted from Kalra *et al.*[50].

Neuroimaging studies using non-invasive methods such as functional and structural magnetic resonance imaging (MRI) and positron emission tomography (PET) in the brain of OCD patients have provided support for the involvement of CSTC circuit in the disorder[51]. Metabolic hyperactivity and increased blood flow in OCD patients were observed in orbitofrontal cortex (OFC), anterior cingulate cortex (ACC), and striatum during neutral and resting states[52-54], which were heightened during symptom provocation[55-57]. A decrease in the glucose metabolism and blood flow was reported within the same CSTC regions after

treatment using SSRI, clomipramine, behavior therapy and ablative neurosurgical procedures[58-62]. Animal studies using optogenetics further support the involvement of CSTC circuit in OCD: Ahmari *et al*.[63] reported a repetitive behavior in mice after optogenetic excitation of a neuronal region connecting orbitofrontal-cortex and the striatum. Additionally, repetitive behaviors were suppressed by optogenetic excitation of the lateral orbitofrontal-cortex and its terminals in the striatum, in a genetic mouse model of OCD[64].

1.1.6 Genetic and Neurochemistry

Twin and family studies support a genetic contribution to obsessive-compulsive disorder[15], with higher concordance rates among monozygotic twins compared with dizygotics confirming the genetic basis of OCD[15]. Twin studies of OCD using a dimensional approach by structural equation modeling concluded that OCD symptoms are heritable, with genetic influences ranging from 45-65 % in children, and 27-47 % in adults[65]. A large number of studies have included first-degree family members of individuals with OCD and have estimated a recurrence risk among first-degree relatives for lifetime OCD between 10 % and 20 %[15]. Additionally, the first-degree relatives of individuals with childhood-onset have higher rates of OCD compared with first-degree relatives of those with adult-onset[66, 67], suggesting a different etiology between children and adults OCD onset[68].

The involvement of neurotransmitters in OCD pathology in still unclear, but several evidences suggest that serotonergic, dopaminergic and glutamatergic genes may play an important role – the disruption of the serotonin system has been linked to obsessive-compulsive disorder based on the effectiveness of selective serotonin reuptake inhibitors in alleviating OCD symptoms[69]. Furthermore, serotonin transporter and serotonin receptor subtypes are at their highest levels in ventral striatum, where they could influence the functioning of CSTC circuit[70]. Indeed, 5-HT2a and 5-HT2c receptors have been implicated in the pathophysiology of OCD and in the mechanism mediating the therapeutic effect of SSRIs, however, it is not clear whether activation or blockade of these receptors would involve OCD behaviour[71]. Therefore, the specific dysfunctions of the serotonergic system in OCD or the mechanisms by which SSRIs exert their therapeutic effect are not known.

Dopaminergic mechanisms seem to be also implicated in the pathophysiology of the disorder, since antipsychotics are beneficial when added to ongoing SRI treatment[43]. The D_2 receptor appears to be of primary interest, since rats treated with quinpirole (a selective $D_{2/3}$ receptor agonist) develop repetitive checking behaviours[72], and a positron emission

tomography study revealed a reduction of D_2 receptor availability in the whole striatum, particularly in the ventral portion, possibly reflecting dopaminergic hyperactivity[73]. Several other candidate genes including the dopamine transporter and dopamine receptors ($D_1 – D_4$) have been reported, however findings have been varied with no conclusive identified genes[74]. In fact, the reported associations between OCD and genes involving serotonergic/dopaminergic pathways lack genome-wide significance and failed when replicated[74]. Furthermore, it should be pointed out that a considerable fraction of patients are refractory to the modulation of serotonin and dopamine systems.

Recently, glutamatergic synaptic dysfunction within the CSTC circuit has likewise been implicated in the pathophysiology of OCD[46, 47], supported by neurochemical investigations, neuroimaging studies, medication, genetics, and animal model. Biochemical studies of cerebrospinal fluid showed elevated levels of glutamate in OCD patients when compared to controls[75], providing a support for a dysregulation in glutamate homeostasis. Despite some conflicting results, magnetic resonance spectroscopy of glutamate and related compounds also presents abnormalities in OCD conditions[46]. These evidences have motivated the interest of using glutamate-modulating agents in the treatment of OCD, particularly in refractory cases to standard pharmacotherapy and psychotherapy, where symptoms improvement were observed in open-label trials[76-78]. Several genetic studies associating genes involving glutamatergic neurotransmission and OCD have been reported[79,80,81]. Results from a genome-wide linkage OCD study suggested the glutamate transporter gene SlcA1 as an OCD candidate gene, localized in the chromosomal region 9p24[79]. Since then, several linkage studies have supported OCD association with this genomic region[80, 81]. *GRIN2B* gene encoding-methyl-D-aspartate (NMDA) receptor subunit 2 was also reported as a positive association between glutamatergic system and OCD, but this study needs to be replicated and thus should be regarded as preliminary[82].

Association between OCD and the ionotropic kainate receptor *GRIK2* gene was also described, where a single nucleotide polymorphism was identified[83], later a second study reporting such association found a different single nucleotide polymorphism in the same gene[84]. Consequently, further studies with independent samples and large sample sizes are required.

Several animal models including *SLITRK5*, *SLC1A1* and *SAPAP3* knockout animals have likewise provided evidences that glutamatergic alterations are related with OCD-like behaviours. Mice lacking *SLITRK5* – a gene coding for a protein involved in neurite outgrowth – showed excessive anxiety and self-grooming behaviours, moreover it presented

substantial alterations in striatal ionotropic glutamate receptor expression and in the distribution of corticostriatal glutamatergic transmission[85]. SLC1A, which regulates neuronal function by limiting the glutamate diffusion in extra-synaptic regions[86], presents a presynaptic localization and a strong expression in the CSTC circuitry[87], where the gene knockout mice exhibit increased aggression and excessive self-grooming[88]. Finally, the *SAPAP3* gene gained substantial interest in OCD, based on its function at the glutamatergic synapse, and on the intriguing phenotype observed in *SAPAP3* knockout mice[89]. The findings of *SAPAP3* association with OCD as well as its role at the glutamatergic synapses will be systematically described in the next topic, and this gene will be the focus of the present work.

In summary, OCD pathophysiology has been focused on corticostriatal circuit and on serotonin, dopamine and glutamate neurotransmitter systems. According to the genetic data obtained so far, and considering the symptoms heterogeneity associated with OCD, it is unlikely that a unique gene contribute substantially to the risk of OCD phenotype, but rather a group of genes that function in a brain network. In this way, a better understanding of the pathophysiology, neurocircuitry and neurotransmitter pathways will be crucial to develop more specific and targeted therapies.

1.2 SAPAP (SAP90/PSD-95-Associated Protein)

1.2.1 SAPAP family - scaffolding protein family of postsynaptic density

In the mammalian brain, neurons communicate with each other through specialized junctions called synapses (Figure 1.2). The chemical synapse is a microscopic structure that transports electrical signals within brain circuits from presynaptic to postsynaptic neurons by the release and action of neurotransmitters[90].

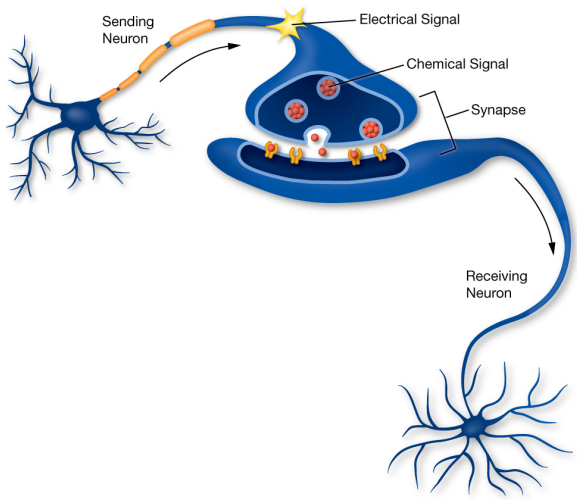


Figure 1.2 – A schematic representation of a synapse. Neurons communicate with each other through a small gap called synapse. At the synapse, information in the form of electrical signals is translated into chemical signals in order to cross the gap. Adapted from[91].

Normal synaptic transmission is dependent on the proper localization and arrangement of specific proteins on both sides of the synapses[92]. The majority of these synapses use glutamate as the excitatory neurotransmitter[92], which is packaged into vesicles at the synaptic termini of glutamatergic neurons[46], when a neuron fires an action potential, vesicles containing glutamate fuse with the presynaptic membrane, releasing their content into the synaptic cleft. In turn, glutamate diffuses across the synaptic cleft, where it can bind to postsynaptic receptors or to presynaptic receptors to limit further glutamate release[46]. Glutamate receptors can be divided in two different types: ionotropic and metabotropic. Ionotropic glutamate receptors (AMPA, kainate, and NMDA receptors) are cation channels that open when glutamate binds and consequently electrically excite the postsynaptic cell. Unlike ionotropic receptors, metabotropic glutamate receptors (mGluRs) are not ion channels; instead they activate signalling cascades that modulate neuronal function. At postsynaptic site the glutamate receptors are inserted in a dense proteinaceous matrix of

signalling proteins, and cytoskeletal components, all assembled by a diversity of scaffolding proteins in a very organized structure called the postsynaptic density (PSD) (Figure 1.3)[92].

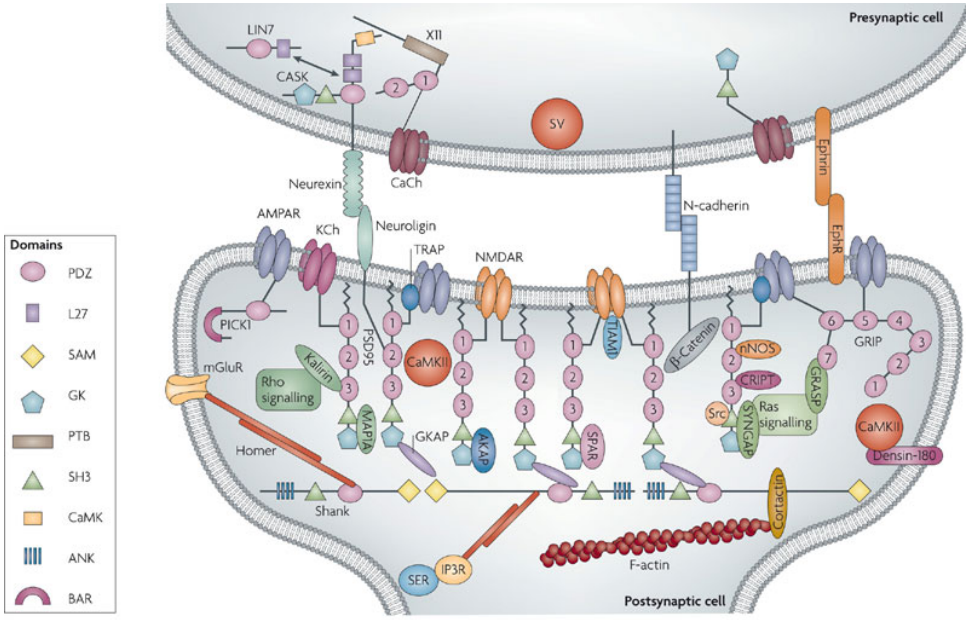


Figure 1.3 – A schematic representation of the protein organization in the glutamatergic postsynaptic density. The postsynaptic density is composed by several components including membrane receptors and ion channels, scaffold and adaptor proteins, signalling proteins, cell-adhesion molecules and components of the cytoskeleton. AKAP, adenylate-kinase anchoring protein; AMPAR, α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors; CaCh, Ca^{2+} channel; CaMKII, calcium/calmodulin-dependent protein kinase II; CRIPT, cysteine-rich PDZ-binding protein; EphR, ephrin receptor; GKAP, guanylate kinase-associated protein; GRASP, GRIP-associated protein; GRIP, glutamate receptor interacting protein; IP3R, inositol-1,4,5-trisphosphate receptor; KCh, K^+ channel; MAP1A, microtubule-associated protein 1A; mGluR, metabotropic glutamate receptor; NMDAR, N-methyl-d-aspartate receptor; nNOS, neuronal nitric oxide synthase; PICK1, protein interacting with PRKCA1; SER, smooth endoplasmic reticulum; SPAR, spine-associated RAPGAP; SV, synaptic vesicle; SyNGAP, synaptic Ras GTPase-activating protein; TIAM1, T-cell lymphoma invasion and metastasis 1; TRAP, C-terminal receptor-binding region. Adapted from Feng *et al.*[93].

The dynamic nature and precise topographical organization of all of these components in the PSD can be viewed as a supramolecular organelle specialized in postsynaptic signal transduction and processing[90]. Moreover, PSD is not just a signal-processing machinery that mediates synaptic transmission, but can also regulate synaptic plasticity changing its own properties to effect patterns of synaptic transmission[92], thus, it can be viewed as a dynamic structure, changing its characteristics (size, shape and composition) depending on synaptic inputs. PSDs can be biochemically purified from mammalian brain by differential centrifugation and sucrose gradient sedimentation to obtain synaptosomes, followed by detergent extraction with non-ionic detergents, such as Triton X-100, which does not

solubilize the PSD[94]. Several PSD proteins have been identified by means of mass spectrometry, which can be classified into a wide range of functional categories including cellular communication and signal transduction, cellular organization, energy, and protein synthesis and processing[95].

Among the PSD components, scaffolding proteins are extremely abundant and crucial players of this organelle (Figure 1.4).

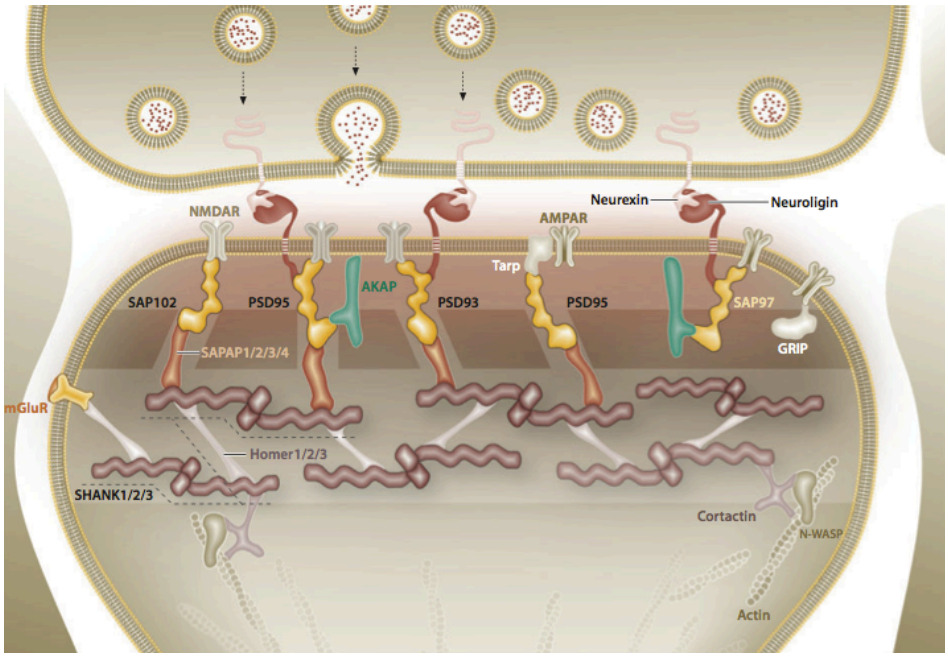


Figure 1.4 – Schematic representation of scaffolding protein arrangements at the postsynaptic density. The main family members of PSD scaffolding proteins at excitatory synapses: PSD95 family (SAP102, PSD95, PSD93 and SAP97); A-kinase anchoring protein (AKAP) family; Homer family; SAP90/PSD95-associated protein (SAPAP) family, and the SH3 and multiple ankyrin repeat domain (Shank) family. Adapted from Ting *et al.*[90]**.**

They act in several synaptic functions including the trafficking, anchoring, and clustering of receptors and adhesion molecules[96]. Scaffolding proteins are commonly defined by a highly conserved organization of domains for specific protein-protein interactions, and can interact simultaneously with multiple binding partners to physically link the PSD components[90]. They may influence the function and availability of receptors by physically clustering and connecting them with other PSD proteins[97]. Scaffolding proteins are represented by multiple family members and the most well studied include members of PSD95 family, A-kinase anchoring protein (AKAP) family, Homer family, SAP90/PSD95-associated protein (SAPAP) family, and the SH3 and multiple ankyrin repeat domain (Shank) family. Findings of *in vivo* functional studies together with human genetics data strongly suggests that mutations in a

variety of these postsynaptic scaffolding proteins may contribute to the etiology of diverse human psychiatric disorders such as schizophrenia, autism spectrum disorders, and obsessive-compulsive spectrum disorders[90].

One of the members of the SAP90/PSD95-associated protein (SAPAP) or guanylate kinase associated protein (GKAP) family is of particular interest in OCD, as will be described later. Four homologous genes encoding for SAPAP 1-4 compose this family of proteins (Figure 1.5), with each member exhibiting approximately 40 % identity at the amino acid level, and no homology to other proteins found in sequence database searches. The 200 amino acids in the C-terminus are particularly well conserved, but regions in the N-terminus and in the middle region are also preserved. The amino acid sequences analysis predicted hydrophilic proteins without a transmembrane region or a signal peptide[98]. Secondary structure prediction shows that SAPAP is mostly unstructured except the C-terminus region[99].

```
O14490 DLGP1_HUMAN          1    ------------------------------------------------------------          0
Q9P1A6 DLGP2_HUMAN          1    MGTAQVLPGILQKHCCILPDRNTESQCTLCGEPEEEEAGDLVQPGISFPGPAEEDLDPQY         60
O95886 DLGP3_HUMAN          1    ------------------------------------------------------------          0
Q9Y2H0 DLGP4_HUMAN          1    ------------------------------------------------------------          0

O14490 DLGP1_HUMAN          1    -------------------MKGLSGSRSH-------HHGVTCDSACDSLSHHS-DRKPYL         33
Q9P1A6 DLGP2_HUMAN         61    SWSPTQHFNEERYSPAPRSMKGLSGSRTQPPLCSGHTCGLAPPEDCEHLHHGPDARPPYL        120
O95886 DLGP3_HUMAN          1    -------------------MRGYHGDRGSHPRPA-------RFADQQHMDVGPAARAPYL         34
Q9Y2H0 DLGP4_HUMAN          1    -------------------MKGLGDSRPRH--------LSDSLDPPHEPLF-AGTDRNPYL         33
                                                    *:*   .*                  : :      * ***

O14490 DLGP1_HUMAN         34    LSPVEHHPADHPYYTQRNSFQAECVG---P---------------FSDPLASSTFPRRHY         75
Q9P1A6 DLGP2_HUMAN        121    LSPADSCPGGRHRCSPRSSVHSECVMMPVV---------------LGDHVSSSTFPRMHY        165
O95886 DLGP3_HUMAN         35    LGSREAFSTEPRFCAPRAGLGHISPEGPLSLSEGPSVGPEGGPAGAGVGGGSSTFPRMYP         94
Q9Y2H0 DLGP4_HUMAN         34    LSPTEAFAREARFPG-QNTLPGD-GLFPLN---------------NQLPPPSSTFPRIHY         76
                                 *.  :          :  .                          ****** :

O14490 DLGP1_HUMAN         76    T-SQQELKDECALVPRTLATKANRIPANLLDQFERQLPLSRDGYHTLQYKRTA-------        127
Q9P1A6 DLGP2_HUMAN        166    S-SHYDTRDDCAVAH--AGAKINRIPANLLDQFEKQLPLHRDGFHTLQYQRTSA------        216
O95886 DLGP3_HUMAN         95    GQGPFDTCEDCV--GHPQGKGAPRLPPTLLDQFEKQLPVQQDGFHTLPYQRGPAGAGPGP        152
Q9Y2H0 DLGP4_HUMAN         77    N-SHFEVPEESPFPSHAQATKINRLPANLLDQFEKQLPIHRDGFSTLQFPRGE-------        128
                                   .  :  ::.       .    *:* .******:***: :**: ** : *

O14490 DLGP1_HUMAN        128    -------VEHRSDSPGRIRHLVHSVQKLFT-KSHSLEGPSKGSV-NGGKAS---------        169
Q9P1A6 DLGP2_HUMAN        217    ------AAEQRSESPGRIRHLVHSVQKLFT-KSHSLEGSSKSNA-NGTKAD---------        259
O95886 DLGP3_HUMAN        153    APGTGTAPEPRSESPSRIRHLVHSVQKLFA-KSHSLEAPGKRDY-NGPKAEGRGGSGGDS        210
Q9Y2H0 DLGP4_HUMAN        129    -------AKARGESPGRIRHLVHSVQRLFFTKAPSLEGTAGKVGGNGSKKG---------        172
                                        : *.:**.**********:**  *: ***. .     ** *

O14490 DLGP1_HUMAN        170    --------------PD-----EAQAARYGKRSKSKERRAEPKA---RPSTSPGWWSSDDN        207
Q9P1A6 DLGP2_HUMAN        260    --------------GRADDHHHAHHAKHSKRSKSKERKPEGKP---RPG-MSSWWSSDDN        301
O95886 DLGP3_HUMAN        211    YPGPGSGGPHTSHHHHHHHHHHHHQSRHGKRSKSKDRKGDGRH----QAKSTGWWSSDDN        266
Q9Y2H0 DLGP4_HUMAN        173    --------------GM--------EDGKGRRAKSKERAKAGEPKRRSRSNISGWWSSDDN        210
                                                          .:*:***:*    .      .   .*******

O14490 DLGP1_HUMAN        208    LDGDMCIYHA---PSGVMTMGRCPDRSASQYFLEAYNTISEQAVKASRSNND--------        256
Q9P1A6 DLGP2_HUMAN        302    LDSDSTY-RT---PSVLNRHHLGP--VAHCYPDALQSPFGDLSLKTSKSNND--------        347
O95886 DLGP3_HUMAN        267    LDSDSGFLAGGRPPG-EPGGPFCLEGPD--------GSYRDLSFKGRSGGSE--------        309
Q9Y2H0 DLGP4_HUMAN        211    LDGEAGAFRSSGPASGLMTLGRQAERSQPRYFMHAYNTISGHMLKTTKNNTTELTAPPPP        270
                                 **.:          .                          .*     .

O14490 DLGP1_HUMAN        257    -VKCSTCANLPVSLDTPLLKKSAWSSTLTVSRAREVYQKASVNMDQAMVKSESCQQERSC        315
Q9P1A6 DLGP2_HUMAN        348    -VKCSACEGLALTPDAKYLKRSSWS-TLTVSQAKEAYRKSSLNLDKPLLHQDAKPALRPC        405
O95886 DLGP3_HUMAN        310    -GRCLACTGMSMSLDGQSVKRSAWH-TMMVSQGRDGYPGA--GPGKGLLGPETKAKARTY        365
Q9Y2H0 DLGP4_HUMAN        271    PAPPATCPSLGVGTDTNYVKRGSWS-TLTLSHAHEVCQKTSATLDKSLLKSKSCHQGLAY        329
                                      :*  : :  *   :*:.:*  *: :*:.::    :     : ::  .:
```

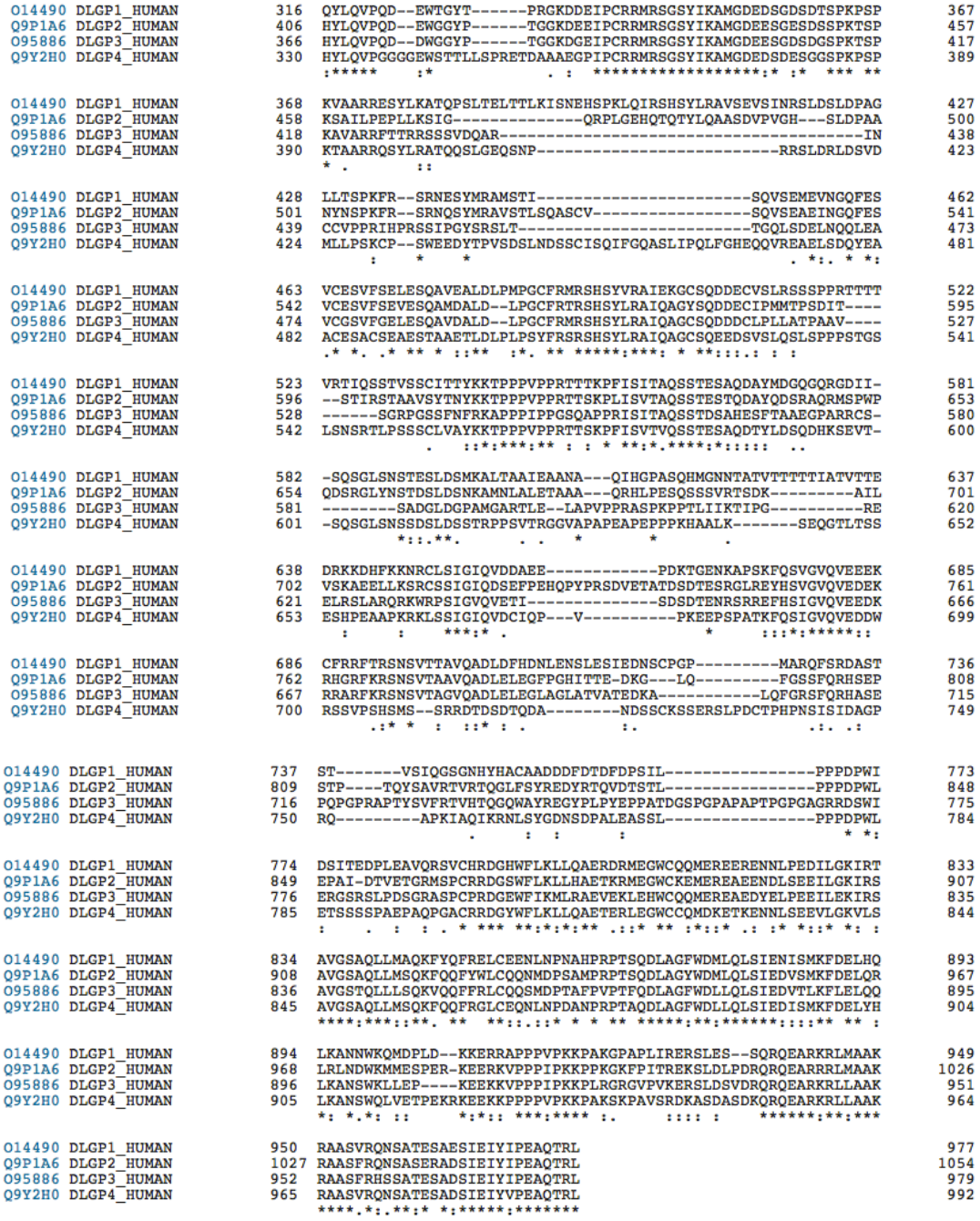


Figure 1.5 – Alignment of SAPAP 1-4 amino acid sequences. Sequences are shown in single letter code and numbered on the right. Positions that have a single, fully conserved residue are signed with an asterisk; conservation between groups of weakly similar properties is signed with one dot; and conservation between groups of strongly similar properties is signed with a colon. Accession numbers and gene names: SAPAP1: O14490, DLG1; SAPAP2: Q9P1A6, DLG2; *SAPAP3*: O95886; SAPAP4: Q9Y2H0. The alignment was performed with Clustal Omega through Uniprot website[100].

Recently, the structure of the C-terminus (residues 807-916) of SAPAP1 from *Rattus Norvegicus* was resolved[99], displaying a three-helix bundle connected by short flexible loops (Figure 1.6), which is stabilized by extensive hydrophobic interaction, three salt bridges, and

two hydrogen bonds. The structure of the protein does not contain a pocket on surface for enzymatic function indicating that this domain might serve for protein-protein interaction[99].

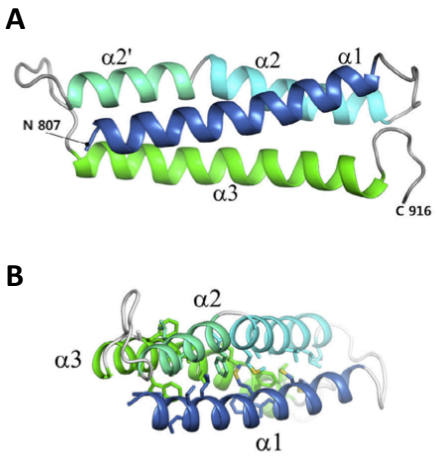


Figure 1.6 – Structure of C-terminus domain 807-916 of SAPAP1. (A) Structure of the C-terminus domain of SAPAP1 was determined at 2.0 Å resolution. The helices were coloured from blue to green. (B) Hydrophobic core of SAPAP1 C-terminus domain. Adapted from Tong J. *et al.*[99].

SAPAPs were originally identified as a family of proteins that interacts with PSD95 family in yeast two-hybrid (Y2H) screenings[98, 101], which led to a proposed function of SAPAPs in the anchoring/signalling complex as linker proteins between glutamate receptor binding proteins and the cytoskeleton. SAPAPs are highly and differentially expressed in many areas of the nervous system, in addition to glutamatergic synapses. SAPAPs also localize to cholinergic synapses, raising the possibility that the SAPAP proteins could be general central components of all excitatory synapses[102].

SAPAP3 mRNA and protein is targeted to the dendrites, whereas SAPAP 1, 2 and 4 mRNAs and proteins are detected mainly in cell bodies[102]. This differential regulation in the expression of SAPAPs suggests that these proteins may play separate *in vivo* molecular roles. Furthermore, the dendritic localization of SAPAP3 indicates that this protein can be locally synthesized playing a role in synaptic plasticity requiring local synthesis/degradation[102], since regulated turnover of PSD proteins may be one mechanism for determining synaptic strength and plasticity. Supporting this idea, Hung *et al.*[103] have shown that stimulating neuronal activity *in vitro* and *in vivo* induces the ubiquitination and degradation of SAPAP by the TRIM3 ubiquitin ligase in hippocampal neurons. SAPAP3 was also associated in the regulation of postsynaptic mGluRs and endocannabinoids-mediated synaptic plasticity, a mechanism for activity-dependent inhibition of synaptic strength in the brain[104]. A role in inhibiting mGluR5 activity and the downstream triggering of AMPA receptors endocytosis

was also proposed[105]. In addition to these functions, protein-protein interaction studies have indicated that SAPAPs also bind to signalling molecules and may thus have multiple roles at synapses[102].

1.2.2 SAPAPs known interacting partners

SAPAPs were originally identified as a family of proteins that interacts with PSD-95 through its guanylate kinase domain, but further studies have indicated that they interact with other PSD proteins as well, suggesting their multiple roles at excitatory synapses.

In 1997, Kim *et al.*[101] and Takeuchi *et al.*[98] stated that SAPAP protein family members bind directly to the GK domain of PSD-95 family members, in yeast two hybrid experiments. Binding to PSD95 is mediated by the N-terminal region of SAPAP containing multiple 14 amino acids repeats, which are highly conserved in all SAPAPs[101]. The most important function of PSD-95 family is to organize signalling complexes at the postsynaptic membrane[96], and members of this family have a molecular structure organization composed by three PDZ domains, one SH3 domain and one guanylate kinase (GK) domain for protein-protein interactions (Figure 1.7)[90]. PSD-95 interacts with a large variety of proteins functioning as molecular scaffolds for the clustering of receptors/ion channels and their associated downstream signalling molecules[90].

Later, two independent reports using yeast two-hybrid screens described the interaction between the members of SHANK family with members of SAPAP family[106, 107]. SHANK family members also contain multiple sites for protein-protein interaction including ankyrin repeats, SH3 domain, a PDZ domain, a long proline-rich region, and a SAM domain (Figure 1.7)[108]. Shank binds through its PDZ domain to the C-terminus of SAPAPs, more precisely via the carboxy-terminal QRTL motif[106]; this binding seems to be regulated by SAPAP phosphorylation with Ca^{2+}/calmodulin-dependent protein kinase II according to Dosemeci *et al.*[109]. In turn, both cortactin and homer proteins were also identified as binding partners to SHANK proteins via its proline-rich region[108], providing a way to connect with the cytoskeleton and the metabotropic receptor (mGluR), respectively. Therefore, PSD-95/SAPAP/SHANK interactions play an important role in organizing the postsynaptic signalling at glutamatergic synapses, associating in an indirect way NMDA with the cytoskeleton and metabotropic glutamate receptors.

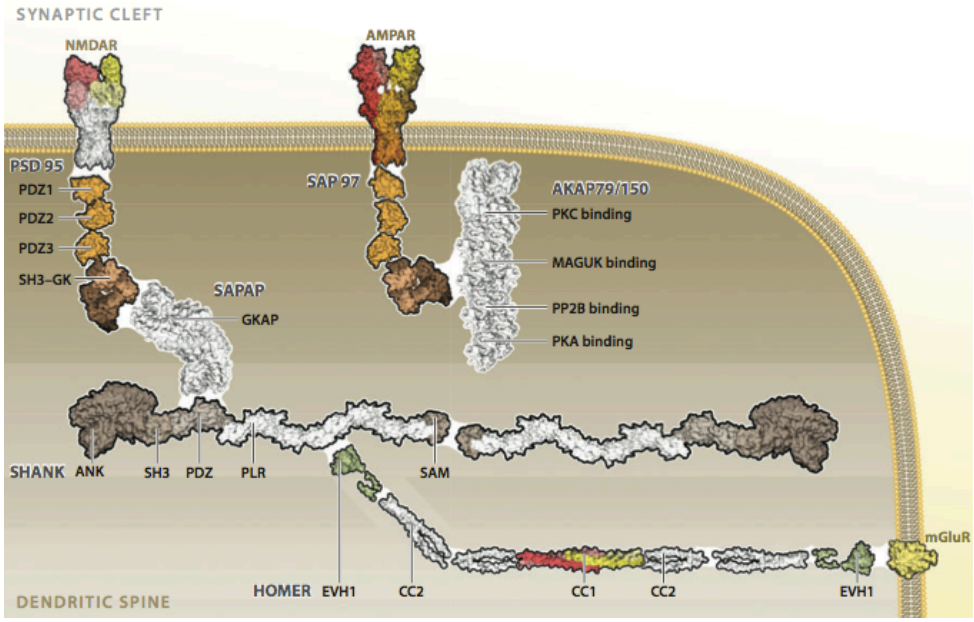


Figure 1.7 – Physical interactions of scaffolding proteins in PSD via discrete binding domains. Represented domains include PDZ and SH3 domains of PSD95 and SHANK family proteins; the GK domain in PSD95 family proteins, the ANK and SAM domains of SHANK family proteins, and the EVH1 and coiled-coil (CC) domains in HOMER family proteins. Adapted from Ting T. *et al.*[90].

Hirao *et al.*[110] found that SAPAP also interacts with the GK domain of synaptic scaffolding molecule (S-SCAM), which has a molecular structure similar to that of PSD-95 comprised of six PDZ domains, two WW domains, and one GK domain. S-SCAM also interacts with NMDA receptors; neuroligin, a cell adhesion protein on the postsynaptic membrane that mediates the formation and maintenance of synapses between neurons; and membrane-associated guanylate kinase-interacting protein (MAGUIN-1) that act as a scaffold molecule at synaptic junctions by assembling neurotransmitter receptors and cell adhesion proteins, showing that S-SCAM is implicated in the assembly of synaptic components[110].

In 1999, Kawabe *et al.*[111] identified the adapter protein nArgBP2 as a SAPAP interacting partner, whose interaction is mediated via the third SH3 domain of nArgBP2 and the proline rich region of SAPAP. It was proposed that this protein might form a complex with PSD-95, S-SCAM and Shank to link the actin cytoskeleton to the components of PSD[110]. Neurofilaments were shown to associate with SAPAP by co-immunoprecipitation from rat brain and by co-localization with neurofilaments in transfected cells. Neurofilaments are components of neuronal cytoskeleton that are important for the maturation of axons and for the maintenance of axonal integrity, that in interaction with SAPAP are also linked to the components of the PSD[112].

Later, it was shown that SAPAP interacts with neuronal nitric oxide synthase (nNOS)[113] and dynein light chain (DLC)[113, 114], linking these proteins to the NMDA receptor/PSD-95 complex. nNOS contains an N-terminal PDZ-binding domain that binds to a similar PDZ domain from the PSD-95, which in turn binds to the cytosolic tail of the NMDAR[115]. Functionally, nNOS might represent a central component that regulates synaptic transmission and intercellular signalling, through negative regulation of NMDAR by S-nitrosylation[115]. DLC, a light chain shared by myosin-V and cytoplasmic dynein, was first isolated as 8 kDa light chain of cytoplasmic dynein, which belongs to a superfamily of molecular motors have fundamental roles in neuronal function, plasticity, morphogenesis, and survival by transporting cargos such as synaptic vesicle precursors, neurotransmitter and neurotrophic factor receptors, and mRNAs within axons, dendrites, and synapses[116]. The interaction between SAPAP and DLC could mediate the interaction between SAPAP and microtubule-based neurons, raising the possibility that SAPAP and its associated proteins can be translocated as a cargo[114]. Later, it was shown that DLC2 isoform interaction with SAPAP activates accumulation of SAPAP and PSD-95 in dendritic spines, and potentiates NMDA currents[117], and that by increasing PSD-95 location in spines, SAPAP-DLC2 interaction could therefore control the AMPA receptors function[117]. Additionally, this interaction also affects the targeting of SAPAP partners such as SHANK3, HOMER3 and PSD-95 to the spine[117].

In 2003, tamalin a scaffold protein that comprises multiple protein-interacting domains, including one PDZ domain, a leucine-zipper region, a proline-rich region, and a carboxyl-terminal PDZ binding motif was also proposed to bind to SAPAP1/3, mGluR1/mGluR5 through the PDZ domain[118]. It also binds the S-SCAM and tamalin itself through the carboxyl-terminal PDZ binding motif, and PSD-95 and MINT-2, a protein with a putative function in synaptic vesicle exocytosis[118]. The interaction of tamalin with all of these proteins involved in postsynaptic organization and protein trafficking in neurons indicates that this protein may participate in receptor clustering, trafficking and intracellular signalling[118].

Through yeast two-hybrid screening and S-transferase pull-down assay, focal adhesion kinase (FAK) and proline-rich tyrosine kinase (PYK2) were described as interactors of SAPAP3[119]. FAK and PYK2 are two related non-receptor tyrosine kinase highly expressed in the brain, and reported to be involved in synaptic plasticity. They contain a central tyrosine kinase domain, an N-terminal FERM domain and a C-terminal region encompassing several proline-rich (PR) motifs, and a focal adhesion-targeting (FAT) domain[119]. FAK plays an important role in regulating actin cytoskeleton, axon guidance, particularly axon outgrowth[120], and PYK2 is involved in the regulation of NMDA glutamate receptors[121].

SAPAP3 was suggested to act as an anchor for these tyrosine kinases in PSD, providing a support for their proposed function in synaptic transmission[119]. Altogether these data demonstrates the importance of SAPAPs as scaffolding proteins in PSD compartments, and evidences how changes in the integrity of these proteins can influence excitatory synaptic transmission through glutamate receptor redistribution and signalling. Therefore extending the knowledge of SAPAPs interactors network will be crucial to gain insight into the role of this protein in excitatory and also in disease synapses.

1.2.3 *SAPAP3* knockout mice – an OCD animal model

SAPAPs are unique to excitatory synapses and well portioned to influence synaptic activity of ionotropic and metabotropic glutamate receptors. They are differentially expressed in many regions of the nervous system, with SAPAP3 being the only family member that is strongly expressed in the striatum[102], thus presenting the possibility to studying its specific function at glutamatergic synapses in this brain region.

An experimental study accomplished by Welch *et al*.[89] found that a genetically engineered mouse with a *SAPAP3* gene deletion shows behavioral phenotypes similar to those observed in obsessive-compulsive disorder in humans. The knockout mouse exhibited increased anxiety and compulsive grooming behavior leading to facial hair loss and skin lesions, being these behaviors significantly reduced after subchronic administration of the SSRI fluoxetine[89]. This observation was particularly interesting since the defect in glutamate neurotransmission of this mouse was responsive to an enhancer of the serotonin neurotransmission, suggesting that alterations in the serotonergic system may modulate the action of the glutamatergic mechanism[89].

SAPAP3 is the only member of the SAPAP family highly expressed in the striatum, and the mice abnormalities observed were rescued by the local expression of SAPAP3 in this brain region[89]. Furthermore, the animals presented reduced activity at corticostriatal synapses, which are part of the circuitry implicated in OCD[89]. They presented an NMDAR subunit composition suggestive of immature corticostriatal synapses, since the levels of the NR2B were significantly increased when compared to the wild-type mice, indicating that SAPAP3 may also be important for the switching and maturation of the NMDAR subunit at cortico-striatal synapses[89]. Noteworthy, no significant differences were found in the levels of PSD95 or Shank protein in the striatal PSD in the *SAPAP3* knockout mice experiments[89]. The authors also stated that the anxiety-like behavior exhibited by the knockout mice is originated from

striatal defects and not from the amygdala and ventral hippocampus defects, to which the anxiety is primarily attributed; but they also considered that some functions of amygdala-striatal projections could be impaired in the mutant mice, since amygdala projects to the striatum[89].

Later on, using optogenetics, which combines genetics and optics to manipulate temporally and spatially precise electrical and biochemical events using fiber-optic light in living organisms[122], Burguière *et al.*[64] demonstrated that in the *SAPAP3* knockout animal model excessive grooming was caused by an impaired pathway in the ventromedial striatum in the orbitofrontal circuitry. The selective stimulation of the orbitofrontal cortex-striatal pathway can restore a behavioural inhibition signal in the knockout mice and can prevent overexpression of repetitive grooming[64].

In addition, using this same animal model, Wan *et al.*[123] showed defects in corticostriatal, but not in thalamostriatal synapses. This finding defines the circuit-level neurotransmission defects in the *SAPAP3* knockout mice, focusing attention on the corticostriatal circuit for mediating the behavioural abnormalities[123]. Moreover, it was also shown that SAPAP4 is present at thalamostriatal synapses, but not corticostriatal, synapses, which provides a molecular basis for the functional divergence observed between thalamic and cortical striatal circuits in *SAPAP3* knockout mice[123].

Together these findings suggest a critical role of SAPAP3 at cortico-striatal synapses, highlighting the involvement of corticostriatal circuitry dysfunction in the pathophysiology of OCD.

1.2.4 Human genetic studies of *SAPAP3* in OCD

The findings that *SAPAP3* null mice exhibits OCD-like behaviours turned *SAPAP3* into an interesting functional susceptibility gene candidate for human OCD, and prompted several human genetic studies of *SAPAP3* in OCD.

The first report involved an association study of *SAPAP3* in OCD and grooming disorders (pathological nail biting, pathological skin picking and trichotillomania) in 383 families, where six SNPs spanning region 1p35 of the *SAPAP3* gene were identified[124]. The preliminary evidence suggested that variations in *SAPAP3* gene are associated with human grooming disorders, which seemed familial and substantially comorbid with each other and with OCD. No clear association between SAPAP3 variants and OCD was established, however the authors pointed out to the fact that grooming disorders without OCD were not common in

the population tested, suggesting the possibility that SAPAP3 variants may be involved in a subtype of OCD involving pathological grooming behaviours[124].

In the second study, *SAPAP3* gene resequencing analysis was performed in OCD and trichotillomania patients[125] where seven sequence variants were identified (Figure 1.8); the variants were present in 4.2 % of the patients and only in 1.1 % of control individuals. The majority of the variants represented missense mutations, and according to computer based algorithms some of the mutations could be damaging to protein function[125]. Two variants were found in patients with OCD alone (A148insGPAGA and P606T), providing the first clinical evidence of potentially associating alterations of this gene with OCD itself, however these findings await further validation[125].

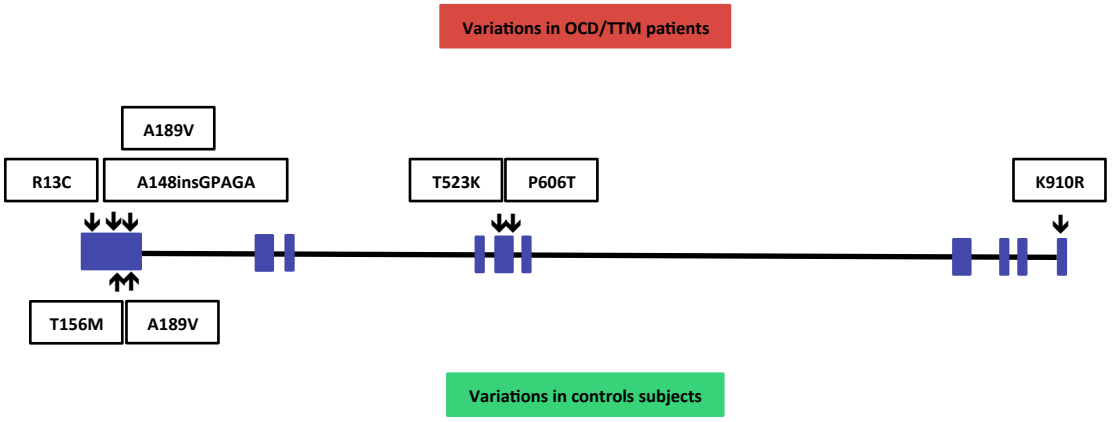


Figure 1.8 – Identified rare nonsynonymous polymorphisms in *SAPAP3* gene. Schematic representation of *SAPAP3* gene with 10 coding exons (blue boxes). Six sequence variants were identified in OCD/ trichotillomania patients (upper part) and two in controls (lower part). Adapted from Zuchner S. *et al.*[125].

A third study evaluated *SAPAP3* genetic SNPs in the etiology of OCD and trichotillomania in South African population[126] - this study found evidences for the association between SNP rs11583978 in *SAPAP3* and trichotillomania, however this association was not significant after correction for multiple testing. For the OCD group, a significant association was observed for earlier onset OCD[126].

So far, the mutations identified are undoubtedly rare, but support the possibility that alterations in SAPAP3 may contribute to OCD and related disorders.

1.2.5 Evidences of SAPAP3 alterations in other neurologic disorders

Taking into account the possible multiple roles of SAPAP3 at excitatory synapses, defects on this scaffold protein might underlie several neurologic disorders. Searching for reports associating SAPAP3 and brain dysfunctions other than OCD, two studies come out. One of these studies examined *SAPAP3* genetic variations associated with schizophrenia[127], and the other reported decreased levels of SAPAP3 in a mouse model of fragile X syndrome[128]. Schizophrenia is a chronic, disabling mental disorder that affects about 1 % of the world population, presenting positive symptoms (hallucinations and delusions, disorganized speech and behavior), negative symptoms (flattened affect, social withdrawal, avolition and anhedonia) as well as cognitive deficits (including impaired attention, working memory, executive function and verbal memory)[129] late in adolescence or in early adulthood[130]. Aberrant synaptic connectivity and synaptogenesis seems to be involved in the pathogenesis of the disorder[131, 132], and although initial neurobiology of schizophrenia mainly focused on excessive dopamine, more recent findings pointed towards the importance of other neurotransmitters such as glutamate in the disorder[133].

Recent animal models[134] and human genetic studies[135] suggested that hypofunction of the NMDA glutamatergic receptor may underlie schizophrenia, since NMDA receptor antagonists like ketamine and phencyclidine, can cause psychotic and cognitive abnormalities reminiscent of schizophrenia[134]. In addition, genetic epidemiological studies have demonstrated that schizophrenia is a complex disorder with a high heritability[135], thus the genes encoding for the components of NMDA receptor-associated proteins can be considered as candidate genes for schizophrenia[127]; in the case of SAPAP3, given its involvement in the stabilization of synaptic junction and regulation of neurotransmitter receptors, it could be also considered as a potential candidate gene for schizophrenia[127]. Accordingly, the authors of the mentioned work searched for genetic variations in the coding region of *SAPAP3* gene in patients with schizophrenia when compared with non-psychotic subjects[127]. As a result, seven known SNPs were identified, but no significant differences were obtained between schizophrenia and controls. Nevertheless, seven rare missense mutations were identified including: P45N, R46C, R116H, P162S, G381S, G587R and R770L, where variants G381S, G587R, and R770L were only detected in schizophrenia, suggesting that these mutations may be patient-specific. The authors speculated that these three missense mutations might affect the interaction of SAPAP3 with their protein kinases partners, and thus affecting synaptic transmission through no phosphorylation of possible targets in schizophrenia[127].

Fragile X syndrome (FXM) is a common form of inherited mental retardation and a leading known cause of autism[136]. It is caused by transcriptional silencing of the *FMR1* gene, which leads to absence or significant reduction of the fragile X mental retardation protein (FMRP) [136]. This protein is important for suitable plasticity, neuronal morphology, and cognitive development, and its absence leads to different levels of intellectual disability[136]. In the work of Krueger *et al.*[128], it was studied the molecular, cellular, and behaviour consequences of the loss of the fragile X mental retardation protein in the prefrontal cortex *FMR1* gene knockout mice. One of the findings indicated that the levels of proteins involved in synaptic function including NMDA receptors NR1, NR2A, and NR2B; and the scaffolding proteins PSD95 and SAPAP3 are decreased in the prefrontal cortex of the knockout mice[128]. Given the loss of SAPAP3 in this animal model, the authors hypothesized that changes in SAPAP3 levels could contribute to the cognitive phenotype in the *FMR1* null mice, based on the phenotype observed in the *SAPAP3* knockout mice.

These studies further reinforce the importance of SAPAP3 in the synaptic function.

1.3 Objectives

SAPAP3, a postsynaptic scaffolding protein highly expressed in the striatum, interacts with a variety of PSD proteins, linking together the NMDA receptor/PSD95 complex and the metabotropic receptor/HOMER/SHANK complex, playing an important role in the regulation of glutamatergic corticostriatal synapses. Results from a knockout animal model suggested that SAPAP3 may be involved in the pathophysiology of obsessive-compulsive disorder (OCD), and demonstrated its critical role at cortico-striatal synapses. These findings prompted several human genetics studies of SAPAP3 in OCD, where several point mutations were identified. More recently, this protein was also associated with schizophrenia, where several missense mutations were identified in patients with this pathology.

The molecular characterization of SAPAP3 has been difficult by the lack of protein production protocols in heterologous systems, most likely due to its large size and lack of annotated domains. The main goals of this work are the establishment of the molecular tools necessary for the characterization of SAPAP3, namely the identification of soluble SAPAP3 domains amenable to be produced in *E. coli*, along with a detailed structural and interactomic characterization of such domains, which will extend the knowledge of the structure-function relationship of SAPAP3 domains, in order to understand the role of this protein at excitatory synapses and its potential involvement in diseases like OCD and schizophrenia.

To achieve such goals, it will be described the use of a high-throughput screening technology called ESPRIT (expression of soluble proteins by random incremental truncation) to identify soluble domains of SAPAP3 (Chapter 2), this part of the work was done in collaboration with Dr. Darren Hart (EMBL-Grenoble, France). It will also be presented the recombinant expression and purification of the best SAPAP3 domain found with ESPRIT (the C-terminal domain 19), as well as, the recombinant production of two mutant forms of this SAPAP3 domain, one associated with OCD (K910R) and the other with schizophrenia (R770L), along with the first crystallization trials done at HTX lab (EMBL-Grenoble). Finally, to get new insights into the role of SAPAP3, Chapter 3 provides a functional interactomic characterization of WT SAPAP3 domain 19, K910R and R770L using an affinity-purification assay combined with SWATH-MS analysis (in collaboration with Dr. Bruno Manadas, CNC – Coimbra, Portugal and Dr. João Peça, CNC – Coimbra, Portugal).

2. Identification of soluble SAPAP3 domains

2.1 Introduction

Numerous protein research strategies largely depend on the ability to produce enough quantities of soluble purified material[137]. While some proteins can be successfully expressed in eukaryotic systems such as mammalian, yeast or insect cells, these are far from universal solutions and present their specific set of problems such as yield, cost, or time[137, 138]. Heterologous protein expression in *Escherichia coli* is many times used as the preferred system due to the well-established recombinant DNA techniques, low cost, high production yields and the simplicity of biomass processing[137].

Mammalian proteins are often more difficult to express in bacterial systems than those from bacteria - they are commonly larger and comprise multiple domains connected by linkers or low-complexity regions[139]. In order to improve yield and solubility of these proteins, it is common to sub-clone shorter constructs corresponding to single or multiple protein domains. The classical approach to predict protein domains consists in the identification of conserved regions using multiple sequence alignments guiding the design of sub-cloning trials[140]. However, the success of this approach is dependent upon the presence of protein sequences with sufficient similarities. Furthermore, even when similar sequences are available, the production of protein domains whose sequence prediction was based on sequence analysis frequently fail to yield sufficient amounts of soluble protein[141]. In these cases, bioinformatics tools for predicting secondary structure and ordered/disordered regions may suggest which regions of the protein are folded or unfolded[140], but they are only of limited accuracy and are unable to reliably define the domain boundaries within the protein primary structure. For such problematic targets, several methods have been described employing strategies common to directed evolution projects, whereby a random library is generated encompassing construct diversity, which is then screened to identify the desired constructs[142]. The random construct library screening-based methods differ from prediction-based strategies in that they identify soluble domains by experimentation, and they often remove the need for iterative cycles of prediction and testing since all genetic constructs may be synthetized and screened in a single experiment[137].

Such libraries can be prepared by enzymatic cleavage using for example exonuclease III[143], uracil-DNA glycosylase[144] or endonuclease V[141]; or by non-specific random priming PCR[145] and physical breakage using sonication or point-sink shearing[146]. These methods produce a multitude of different sized gene fragments that can be sized-fractionated within desired ranges by agarose gel electrophoresis. The selection or screening process for identification of soluble protein domains is critical to the success of the method, where the majority of

library-compatible solubility screens are indirect, in the sense that they report solubility via the detection of C-terminally fused partners such as fluorescent proteins as green fluorescent protein (GFP)[147] or antibiotic resistance proteins as homotrimeric chloramphenicol acetyltransferase[148]. However, false positives are frequent since constructs identified are often soluble only in the presence of the fused reporter. One strategy to reduce the number of false positives is to implement a screen using small linear peptide fusion tags, as they should not perturb significantly the protein solubility[149]. These fusion tags are only detected if fused to a stable domain, as aggregated protein will bury these tags, which will no longer be detected[138]. One of such strategies is based on the use of a split-GFP strategy[150, 151]. GFP has a typical β-barrel structure, comprising 11 anti-parallel β-strands, and a truncated non-fluorescent form comprising strands 1-10 (GFP 1-10) can be produced, which will become fluorescent in the presence of the eleventh strand (GFP 11), even when present in *trans*[150]. This is the basis for the slit-GFP strategy, where GFP 11, a 15 amino acid peptide, is fused to the constructs C-terminus and the co-expressed detector GFP 1-10 acts as a fluorescent solubility detector[150].

Alternatively to the use of split-GFP system as indicator of domain expression and solubility, the short C-terminal extension biotin acceptor peptide (BAP) can also be used. BAP is post-translationally biotynylated *in vivo* in *E. coli* by cytoplasmic BriA enzyme[138]. This is the case of ESPRIT (Expression of Soluble Protein by Incremental Truncation) technology (depicted in Figure 2.1), used in this work and further described below.

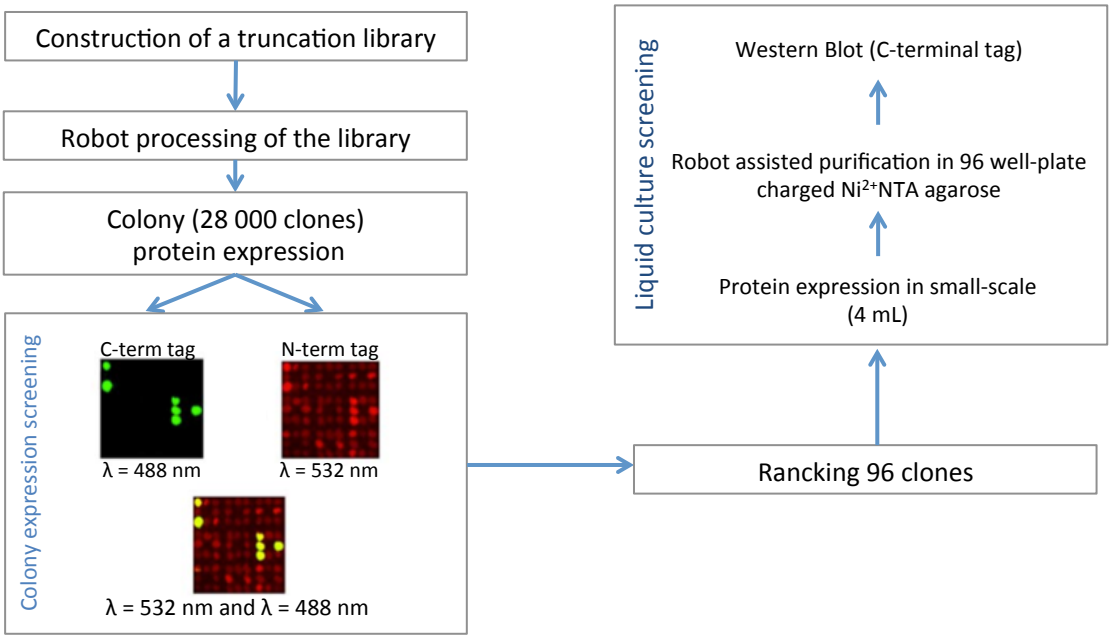


Figure 2.1 – ESPRIT process workflow. A random truncation library of a target protein is generated with exonuclease III and mung bean nuclease, producing tens of thousands of variant clones after transformation in *E. coli*. In the first screening step (colony expression screening), putative soluble constructs are assessed by intensity of N-terminal hexahistidine tag (red) and C-terminal biotynylated biotin acceptor peptide (green) signals on the colony blot. The 96 highest ranked clones are expressed in 4 mL liquid expression cultures, purified by Ni^{2+} NTA affinity chromatography and eluted fractions analysed by Western blot to identify biotinylation signals (liquid culture screening). Adapted from Yumerefendi *et al.*[138].

Briefly, the first step of ESPRIT consists in generating a large library of random genetic constructs by the action of exonuclease III and mung bean nuclease. The gene of interest is cloned into a vector encoding an N-terminal hexahistidine tag, and a C-terminal biotin acceptor peptide; and at the edges of the insert to be truncated exist a pair of restriction sites that leave exonuclease III sensitive 5' and resistant 3' overhangs (Figure 2.2).

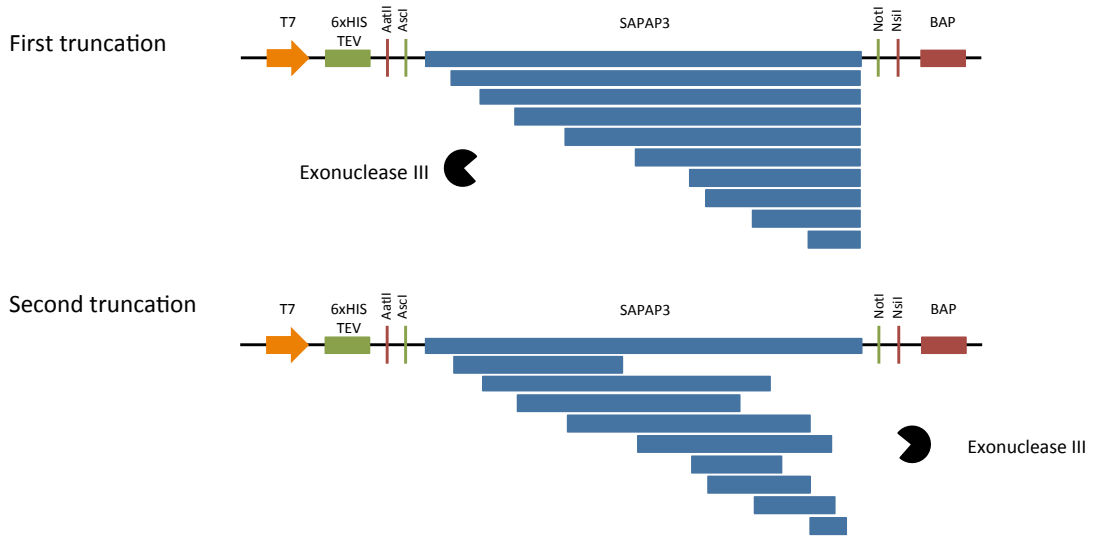


Figure 2.2 – ESPRIT truncation molecular biology. The truncation vector combines two pairs of restriction enzyme sites (AatII and AscI, NotI and NsiI) to allow sequential unidirectional reactions that generate internal fragments by the action of exonuclease III and mung bean nuclease in a two step process. BAP: biotin aceptor peptide; ExoIII: exonuclease III; MBN: mung bean nuclease; 6xHis: hexahistidine tag; TEV: Tobacco Etch Virus protease cleavage site. Adapted from An *et al.*[140].

In a two step process (Figure 2.2), one end is truncated randomly and a plasmid library recovered, a second reaction is then performed on a pool of these plasmids which are then religated; producing tens of thousands of variant clones after transformation in *E. Coli*. Colonies are then robotically picked and gridded onto nitrocellulose membranes on top of agar to generate colony arrays, in which protein expression is induced. Colony blots are prepared by in situ lysis on nitrocellulose membranes and blots are probed with fluorescent streptavidin to detect the biotinylation status of the C-terminus BAP. The BAP biotinylated will function as an indicator of solubility, since soluble proteins are better biotinylated by *E. coli* than those that are insoluble/aggregated (Figure 2.3). Simultaneously, the array is probed with a monoclonal antibody against the N-terminal hexahistidine tag, and only clones exhibiting both N- and C-terminal tags are further analysed since they comprise undegraded, intact protein constructs (Figure 2.3). Identification of the most efficiently biotinylated constructs often correlates with the most soluble clones, that in the second stage of screening, are expressed in small-scale liquid format, lysed and purified on Ni^{2+} affinity resin. Soluble, purifiable constructs are visualised by streptavidin-Western blot and their boundaries determined by DNA sequencing (Figure 2.1).

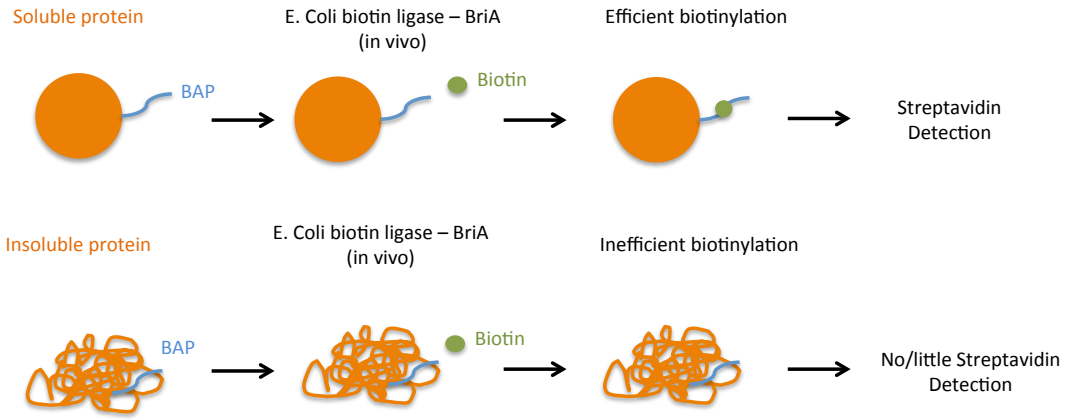


Figure 2.3 – *In vivo* biotynylation as a solubility marker. The target of interest is fused to the biotin aceptor peptide (BAP) via a short linker. Soluble proteins are better biotynalated by *E. coli* than those that are degraded or insoluble. Adapted from Yumerefendi *et al.*[138].

Once positive clones have been detected, protein required for the downstream applications can be expressed and purified using standard procedures, with no need for subcloning into different vectors.

The work presented in this chapter provides a detailed description on how ESPRIT technology led to the identification and of SAPAP3 soluble domains when expressed in *E. coli*. Part of the work included in this chapter (sections 2.3.2 – 2.3.4) was performed at Dr. Darren Hart's group - EMBL Grenoble (France).

2.2 Materials and Methods

2.2.1 Generation of plasmids for *SAPAP3* library construction

The synthetic cDNA encoding for SAPAP3 (in pUC57) was sintetised by GenScript (USA), with codons optimized for *E. coli* expression, with AscI and NotI sites in 5' and 3´terminals, respectively (cDNA sequence in Supplementary Figure 6.1) and avoiding the AatII, AscI, NotI and NsiI sites within the *SAPAP3* sequence. The vector was digested with AscI and NotI-HF restriction enzymes (New England Biolabs) and inserted into the expression vector pESPRIT. This vector encodes a N-terminal hexahistidine tag and a C-terminal biotin acceptor peptide and the cloning site is flanked by a pair of restriction sites that originate exonuclease III sensitive 5' and resistant 3' overhangs. The vector was also digested with AscI and NotI-HF and treated with alkaline phosphatase (New England Biolabs) before ligation with AscI and NotI-HF digested *SAPAP3* cDNA. The fragments obtained after the digestion of *SAPAP3* gene and pESPRIT vector were separated by agarose gel electrophoresis, and the bands of interest excised from the gel and purified with Nucleospin Gel and PCR Clean-up (Macherey-Nagel). The digested DNA fragments were ligated with T4 DNA Ligase (New England Biolabs) according to the manufacturer's instructions. The ligated DNA construct was transformed in *E. coli* Top10F' competent cells and positive clones selected by restriction analysis, using HindIII and NdeI, with positive clones yielding fragments of 1717 bp and 3360 bp size.

2.2.2. Internal His rich region removal

The PCR reaction mixture for histidine rich region deletion contained 10 ng of DNA; 10 mM dNTP´s; 10 µM phosphorylated of forward primer (5'-CAGTCTCGTCATGGTAAACGC-3'); 10 µM of phosphorylated reverse primer (5'-TGAGGTGTGCGGGCC-3´), 5 x Phusion HF Buffer and 1.0 units Phusion DNA polymerase to a final volume of 50 µL. The PCR conditions for the deletion consisted in an initial denaturation step at 98 °C for 3 minutes, followed by 30 cycles of denaturation at 98 °C for 30 seconds, annealing at 65 °C for 30 seconds and extension at 72 °C for 3 minutes, with a final extension step at 72 °C for 10 minutes. To digest the parental DNA, the PCR product was treated with 20 units of DpnI restriction enzyme (New England Biolabs) at 37 °C for 3 hours, and purified using the Nucleospin Gel and PCR Clean-up (Macherey-Nagel). The DNA was re-circularised using 10 units of T4 DNA ligase (New England Biolabs), transformed into *E. coli* Top10F' competent cells and purified using Nucleospin Plasmid Quick-pure (Macherey-Nagel). The DNA sequence was confirmed

by Sanger DNA Sequencing at Macrogen (Netherlands), and the results analyzed with VectorNTI sequence alignment software (Thermo Fisher Scientific).

2.2.3 Construction of *SAPAP3* random truncation libraries

Exonuclease III and mung bean nuclease were used to construct incremental truncation libraries. First, high quality plasmid was prepared from 200 mL overnight culture of cells by standard alkaline lysis and phenol-chloroform extraction protocols, which was then further purified using QIAprep Spin Miniprep Kit (Qiagen). Ten micrograms of plasmid were digested with AatII (yielding an exonuclease III insensitive end adjacent to the promoter) and AscI (yielding an exonuclease sensitive end adjacent to the gene insert), both sites located upstream of the target gene.

Four micrograms of purified, linearised plasmid were mixed with reaction buffer (1 x Buffer 1 New England Biolabs supplemented with 45 mM NaCl) and 400 U of exonuclease III, in a final volume of 120 µL that was incubated at 22 °C. To ensure even fragment distribution, 0.5 µL of the reaction was removed every 30 sec over 2 h and pooled in a tube containing 200 µL of 3 M NaCl on ice. The quenched reaction was denatured at 70 °C for 20 min and DNA purified using Nucleospin Extract II kit (Macherey-Nagel). In order to remove the 5' overhang after exonuclease III digestion, the DNA was incubated with 5 U of mung bean nuclease in a final volume of 55 µL at 30 °C for 30 min, which was then purified with the Nucleospin Extract II kit. The ends of the DNA molecules were polished by incubation with *Pfu* polymerase (Stratagene) in a final volume of 50 µL (1 x *Pfu* polymerase native buffer, 2.5 mM dNTPs and 5 U of enzyme) at 72 °C for 20 min. The reaction was then separated by electrophoresis in a 0.5 % agarose gel and slices of gel containing linearized plasmid with inserts of size 300-3000 bp were excised. DNA fragments were recovered from the gel using QIAEXII resin (Qiagen) and re-circularised using T4 DNA Ligation Kit (Roche), and then used to transform *E. coli* Mach1 T1 competent cells (Thermo Fisher Scientific). Transformation mixes were recovered in SOC medium and plated on LB (Luria Broth) agar (supplemented with kanamycin 50 µg/mL) in 22 cm QTrays (Genetix). After overnight growth at 37 °C, approximately 4.0 x 10^4 colonies were scraped from the agar resuspended in phosphate saline (PBS) buffer and the plasmid DNA mixture extracted from the cell pellets by miniprep kit (Quiagen).

The random C-terminal truncation library of *SAPAP3* was constructed from the DNA mixture obtained from the N-terminal truncation using a similar approach, but differing in the

cleavage of the plasmid with NsiI and NotI located at the 3' end of the *SAPAP3* gene insert, and in the agarose-gel purification of vector constructs with inserts with sizes ranging from 300-1200 bp. After library transformation into *E. coli* Match T1 competent cells (Thermo Fisher Scientific) approximately 5 x 10^4 colonies were pooled for DNA extraction.

2.2.4 Preparation of colony arrays for expression screening

Competent cells of *E. coli* BL21-AI™ (Thermo Fisher Scientific) (an arabinose-inducible strain designed to give the maximum protein expression with the tightest regulation available from a T7 expression system) were transformed with the target library and plated on 22 cm LB agar QTrays (supplemented with 50 µg/mL kanamycin) at a density of approximately 3 000 colonies/tray, and incubated overnight at 30 °C. Colonies were isolated using a colony-picking robot (KBiosystems, Basildon, UK) into 384 well plates containing 70 µL TB medium per well supplemented with 50 µg/mL kanamycin. Liquid cultures were grown overnight at 37 °C in a HiGro shaker incubator (Genomic Solutions), and then all clones were arrayed robotically onto a nitrocellulose membrane (Amersham) laid over LB agar plates supplemented with antibiotic. Plates were incubated overnight at 25 °C until colonies on the membrane were just visible. The membrane was then carefully moved onto a fresh LB agar plate (supplemented with antibiotic, 0.2 % arabinose and 50 µM biotin) to induce recombinant protein expression within the colonies during overnight incubation at 25 °C.

2.2.5 High-throughput screening of SAPAP3 putative soluble domains

Membranes were incubated for 4h at 30 °C, lifted from the inducing agar and laid over filter paper soaked in denaturing buffer (0.5 M NaOH, 1.5 M NaCl) for 10 min at room temperature. The membranes were neutralized 2 x 5 min in neutralisation buffer (1 M Tris-HCl, pH 7.5, 1.5 M NaCl) and then for 15 min in 2 x SSC buffer (30 mM Trisodium citrate dihydrate, 0.3 M NaCl). The cellular debris of colonies on the membrane was carefully removed with a glass spreader, and the membrane was blocked overnight in Superblock (Pierce, Chicago, USA) at 4 °C. The membranes were washed 3 times with PSB-T (PBS with 0.05 % Tween 20) buffer for 5 min each, and incubated in 50 mL of PBS-T containing 1/3125 dilution of mouse anti-hexahistidine antibody (Amersham) for 1 h at 4 °C. After washing with PBS-T buffer, the membranes were further incubated in 50 mL of PBS-T containing 1/5000

dilution of streptavidin Alexa Fluor 488 (Thermo Fisher Scientific) and 1/1000 dilution of Alexa Fluor 532 rabbit anti-mouse IgG (Thermo Fisher Scientific) for 1 h. After washing with PBS-T buffer, the membranes were scanned with a Typhoon 9400 fluorescence imager (Amersham) for hexahistidine tag and biotin acceptor peptide signals intensities, respectively. Signals from the array were quantified from digitalized images using Visual Gird software (GPC Biotech) and exported to Microsoft Excel for analysis. Clones were sorted according to hexahistidine signals and those exhibiting no clear signal were eliminated from further analysis. Those that remained were ranked according to their biotinylation signals in order to identify putatively soluble expressing clones.

2.2.6 Small-scale expression and purification screening of SAPAP3 soluble domains

The 96 most intensely biotinylated clones from the library were selected for small-scale protein expression at 4 mL scale in 24 well plates with TB medium supplemented with kanamycin. After the cultures reached OD_{600} = 0.8, 0.2 % arabinose and 50 µM biotin were added to induce protein expression and enhance protein biotinylation respectively. After overnight induction at 25 °C, the cells were pelleted by centrifugation, and resuspended in 4 mL of sphaeroplast buffer (20 mM Tris pH 8.0, 140 mM NaCl, 20 % sucrose and 1 mg/mL lysozyme). The sphaeroplasts were harvested by centrifugation and then resuspended in 800 µL lysis solution per well, containing 10 mM Tris pH 7.5, 0.5 % Brij35, 0.25 U/ µL Benzonase (Novagen) and 0.8 µL Protease Inhibitor Cocktail (Sigma). Ninety-six protein samples were simultaneously purified on a liquid handling robot with samples loaded into a 96 well filter plate with 60 µL Ni^{2+} NTA agarose (Qiagen) in each well and mixed by orbital shaking at 4 °C for 30 min. The samples were then washed with washing buffer (50 mM phosphate buffer, pH 7.0, 140 mM NaCl, 5 mM Imidazole) and eluted with elution buffer (50 mM phosphate buffer, pH 7.0, 140 mM NaCl, 300 mM Imidazole). The 96 samples were subjected to SDS-PAGE, and then electroblotted onto nitrocellulose membrane. The membrane was probed with a 1:1000 dilution of a Streptavidin 488; washed with PBS-T, and scanned using a fluorescence imager (Typhoon scanner, GE Healthcare) to analyse biotinylation of the biotin acceptor peptide.

Plasmids were purified from the clones identified as soluble hits, and the DNA inserts characterized by sequencing to identify domain boundaries.

2.2.7 Medium-scale expression and purification of selected SAPAP3 clones

The positive clones were grown at 37 °C in 50 mL of LB media supplemented with kanamycin, and when the culture reached an OD_{600} = 0.8, 0.2 % (m/v) L-arabinose was added to induce protein expression at 25 °C and 170 rpm for 16 hours. The cells were harvested by centrifugation and each cell pellet was resuspended in 2 mL of resuspension solution (20 mM sodium phosphate, 500 mM NaCl, 20 mM Imidazole, pH 7.4) and lysed by freeze and thaw. Twenty five microliters of DNase was added to each sample; and after centrifugation, the supernatant was collected and loaded into columns supplemented with 300 µL Ni^{2+}NTA resin (GE Healthcare), washed (20 mM sodium phosphate, 500 mM NaCl, 20 mM Imidazole, pH 7.4), and first eluted with 100 mM Imidazole (20 mM sodium phosphate, 500 mM NaCl, 100 mM Imidazole, pH 7.4) followed by elution with 500 mM Imidazole (20 mM sodium phosphate, 500 mM NaCl, 500 mM Imidazole, pH 7.4). Eluted fractions were analysed by SDS-PAGE and Western blot with anti-histidine tag mouse monoclonal antibody (GenScript).

2.2.8 SAPAP3 domain 19 removal of BAP tag

Two hundred nanograms of the vector encoding SAPAP3 domain 19, which contained a C-terminal biotin acceptor peptide flanked by two BspEI sites, was digested with 5 units (0.5 µL) of this enzyme in 20 µL reaction volume (New England Biolabs) at 37 °C for 3 hours, to remove the BAP tag, and purified using Nucleospin Gel and PCR Clean-up (Macherey-Nagel). The digested DNA was quantified by measuring the absorbance at 260 nm on a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). One hundred nanograms of digested DNA was ligated with 1 µL of T4 DNA Ligase (New England Biolabs) in a total reaction volume of 20 µL, overnight at 16 °C in T4 DNA ligase buffer (50 mM Tris-HCl, pH 7.5; 10 mM $MgCl_2$, 1 mM ATP, 10 mM Dithiothreitol). After ligation, 10 µL were transformed into DH5α competent cells - for transformation, the competent cells were thawed on ice and the ligated DNA was added to 50 µL of cells. The transformation reaction was incubated on ice for 30 minutes and heat-shocked at 42 °C for 1 minute. After 3 minutes of incubation on ice, 1 mL of LB medium was added and cells were recovered at 37 °C for 1 hour with shaking. The mixture was centrifuged for 1 minute at 14 000 *g*, the supernatant discarded and the pellet re-suspended in the remaining medium (approximately 150 µL). The culture was spread on LB/agar plate containing 50 µg/mL of kanamycin and the plate was incubated overnight at 37 °C. A single colony from the transformation was grown overnight at 37 °C with shaking in 10 mL of LB medium containing 50 µg of kanamycin. The culture was

harvested by centrifugation for 10 min at 4 000 *g* at 16 °C and purified using Nucleospin Plasmid Quick-pure (Macherey-Nagel) according to the manufacturer's instructions. The plasmid DNA was quantified by measuring the absorbance at 260 nm on a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific).

2.2.9 Mutagenesis of *SPAPAP3* domain 19 - K910R and R770L mutants

The SAPAP3 K910R mutation was obtained by restriction free (RF) cloning [152] using SAPAP3 domain 19 WT without BAP tag as template. To do that, the PCR reaction mixture contained 100 ng of template DNA; 0.2 mM dNTPs; 0.5 µM of forward primer (5'-AGGAAAGTGCCGCCGC-3'); 0.5 µM of reverse primer (5'-TTCTTCTTTCGGTTCCAGCA-3´), 5x Phusion HF Buffer and 1.0 unit of Phusion DNA polymerase (New England Biolabs) in a final volume of 50 µL. The PCR conditions for the mutation consisted in an initial denaturation step at 98 °C for 30 seconds, followed by 35 cycles of denaturation at 98 °C for 10 seconds, annealing at 65 °C for 10 seconds and extension at 72 °C for 3 minutes, with a final extension step at 72 °C for 10 minutes. The PCR amplified product was gel-purified and re-circularized using T4 DNA ligase (New England Biolabs).

The SAPAP3 R770L mutation was obtained using the QuickChange Site-directed mutagenesis kit (Agilent) of the SAPAP3 domain 19 WT without BAP tag. The PCR reaction mixture contained 50 ng of template DNA; 0.2 mM dNTP´s; 0.2 µM of forward primer (5'-GGGTCCGGGTGCGGGCCTGCGCGATAGTTGGATTGAACG-3'); 0.2 µM of reverse primer (5'-CGTTCAATCCAACTATCGCGCAGGCCCGCACCCGGACCC-3'), 10x Pfu Turbo DNA polymerase buffer and 2.5 units Pfu Turbo DNA polymerase (Agilent) in a final volume of 50 µL. The DNA was denatured by initial denaturation step at 95°C for 30 seconds, followed by 18 cycles of denaturation steps at 95 °C for 30 seconds, annealing at 60 °C for 1 minute and extension at 68 °C for 1 minute. The PCR products were treated with 20 units of DpnI restriction enzyme (New England Biolabs) at 37 °C for 3 hours, and purified using the Nucleospin Gel and PCR Clean-up (Macherey-Nagel). Both DNAs were transformed in DH5α competent cells and purified using Nucleospin Plasminc Quick-pure (Macherey-Nagel) as described in section 3.2.1. The DNA sequence was confirmed by Sanger DNA Sequencing at Macrogen (Netherlands), and the results analyzed with VectorNTI 11 sequence alignment software (Thermo Fisher Scientific).

2.2.10 Expression and purification of WT SAPAP3 domain 19, K910R and R770L

For larger scale expression of SAPAP3 domain 19 and their mutant forms, the constructs were transformed into competent E. coli BL21 AI strain and plated into LB/agar plates supplemented with 50 µg/mL kanamycin. One colony was used to inoculate 50 mL of LB supplemented with 50 µg/mL kanamycin, that was grown overnight at 37 °C. To inoculate 1 L of LB supplemented with 50 µg/mL kanamycin, 25 mL of the previous cell culture were used. The cells were grown at 37 °C in incubator with shaking, until an OD_{600nm} of 0.6 – 0.8, and protein expression was induced with 0.2 % of L-arabinose, at 25 °C, for about 16 hours. Cells were harvested by centrifugation at 4 000 *g* for 20 min at 4 °C, re-suspended in lysis buffer (20 mM sodium phosphate pH 7.4, 500 mM NaCl, 10 % (v/v) glycerol, 25 mM imidazole, 0.1 M phenylmethylsulphonyl fluoride (PMSF)), and disrupted through a high pressure homogenizer - Emulsiflex-C3 (Avestin) with three passages at 1 000 bar.

The cellular extract was clarified by centrifugation (35 000 *g* for 20 min, at 4°C) and the supernatant applied to a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with lysis buffer at 2 mL/min. Protein elution was performed by a four-step gradient of imidazole (50 mM, 150 mM, 300 mM and 500 mM) at a flow rate of 5 mL/min and fractions containing the protein of interest (150 mM imidazole gradient step) were pooled and applied to a HiLoad Superdex 200 26/60 prep-grade column (GE Healthcare) pre-equilibrated in 20 mM sodium phosphate pH 7.4, 150 mM NaCl and 10 % (v/v) glycerol at 2 mL/min. Protein was further purified by cation exchange chromatography with a Mono S 5/50 GL column (GE Healthcare), equilibrated in 20 mM sodium phosphate pH 7.4 and 10 % (v/v) glycerol. Elution was carried out by a gradient of NaCl (0 - 1 M), at a flow rate of 0.75 mL/min. Protein was finally concentrated using a 10 kDa MWCO centrifugal concentrator (Amicon Ultra 4 mL centrifugal filters, Millipore).

2.2.11 SDS-PAGE and Western Blotting

SDS-PAGE analysis was performed in a Bio-Rad Mini Protean III electrophoresis apparatus using 12.5 % polyacrylamide gels, at 150 V and at room temperature, for approximately 45 min. Samples were treated with loading buffer (0.35 M Tris-HCl, 0.28 % SDS buffer pH 6.8, 30 % glycerol, 10 % SDS, 0.6 M DTT and 0.012 % Bromophenol Blue) and incubated at 95 °C for 5 minutes before gel loading. After electrophoresis separation, gels were stained with Coomassie Brilliant Blue (0.2 % Coomassie Brilliant Blue R-250 (Sigma), 10 % acetic acid and 50 % methanol) for 5 min, with agitation at room temperature. For Western blot analysis,

protein samples were resolved by SDS-PAGE and electro-transferred onto polyvinylidene fluoride (PVDF) using the trans-blot turbo (Bio-Rad) apparatus, in transfer buffer (25 mM Tris, 192 mM Glycine, 20 % methanol and 0.025 % SDS). Membranes were blocked for one hour in standard TBS containing 1 % (v/v) Tween-20 supplemented with 5 % (w/v) skim milk and incubated with anti-histidine tag mouse monoclonal antibody (GenScript) diluted in a 1:10 000 ratio in TBS-T buffer with 0.5 % skim milk, for 1 h. Membranes were washed in TBS containing 0.1 % (v/v) Tween-20, incubated with secondary anti-mouse alkaline phosphatase-conjugated antibody (GE Healthcare) and revealed using ECF chemifluorescence detection kit (GE Healthcare) in a Molecular Imager FX (Bio-Rad).

2.2.12 Protein quantification

Purified proteins were quantified by measuring absorbance at 280 nm using a Nanodrop ND-1000 Spectrometer (Thermo Fisher Scientific). The molar extinction coefficient for each protein was estimated from the respective amino acid sequence using Vector NTI software (Thermo Fisher Scientific).

2.2.13 Analytical size exclusion chromatography

Analytical size exclusion chromatography was performed using a Superdex200 5/150 GL column (GE Healthcare). The proteins were injected through and automated injector apparatus, and eluted using 20 mM sodium phosphate pH 7.4, 300 mM NaCl and 10 % (v/v) glycerol at 0.4 mL/min and monitoring absorbance at 220 nm. Apparent molecular weights of eluted proteins were determined using a calibration curve established with the following standards from the gel filtration low- and high-molecular weight calibration kits (GE Healthcare): ribonuclease A (13.7 kDa), carbonic anhydrase (29.0 kDa), ovalbumin (44 kDa), conalbumin (75.0 kDa), aldolase (158 kDa) and blue dextran (≈ 2000 kDa).

2.2.14 Circular dichroism

Circular dichroism measurements were performed at 20 °C on a Jasco J-815 spectrometer fitted with a Peltier temperature controller. Spectra were acquired between 190 and 260 nm, set up to 1 nm band width, 1 s response, 100 nm/min scanning speed and 6 accumulations.

Protein samples were in 20 mM sodium phosphate pH 7.4, 300 mM NaCl and 10 % (v/v) glycerol, which were diluted with water to the desired concentration in a final volume of 300 µL. Each spectrum was the average of six scans with the average buffer control spectrum subtracted. All measurements were obtained in ellipticities in mdeg and converted to molar ellipticities ([θ], deg cm2 dmol-1) by normalizing for the concentration of peptide bonds. Deconvolution analysis of the data was done with CDNN software using the complex CD spectra 33x algorithm.

2.3 Results

2.3.1 Generation of plasmid for *SAPAP3* library construction

The synthetic cDNA encoding for SAPAP3 (in pUC57) with codon optimized for expression in *E. coli* (Supplementary Figure 6.1) cloned into pUC57 between AscI and NotI restriction sites (Figure 2.4-A) was cleaved and inserted into pESPRIT vector. For this purpose pUC57-*SAPAP3* was digested with AscI and NotI restriction enzymes, and the resulting fragments separated by agarose gel electrophoresis (Figure 2.4-B). The gel showed the presence of two bands, where the lower band correspond to the *SAPAP3* (2956 bp) and the upper band to incompletely digested vector (5666 bp). The pESPRIT was also digested with AscI and NotI restriction enzymes in order to open the vector by removing the MBP insert. After agarose gel electrophoresis (Figure 2.5-B) the desired fragment of pESPRIT with 2100 bp, correspondent to the upper band in the gel was excised and purified.

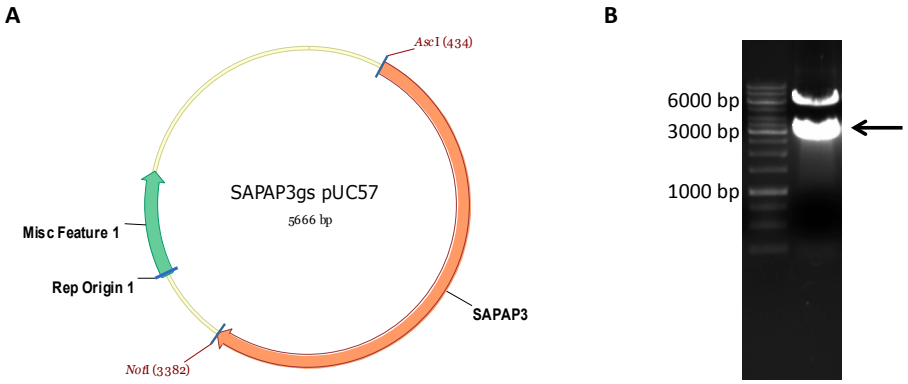


Figure 2.4 – pUC57-*SAPAP3* digestion with AscI and NotI restriction enzymes. (A) Representation of *SAPAP3* gene in pUC57 vector with AscI and NotI restriction sites. (B) Agarose gel electrophoresis separates two fragments corresponding to incomplete digested pUC57 with 5666 bp and *SAPAP3* cDNA with 2956 bp.

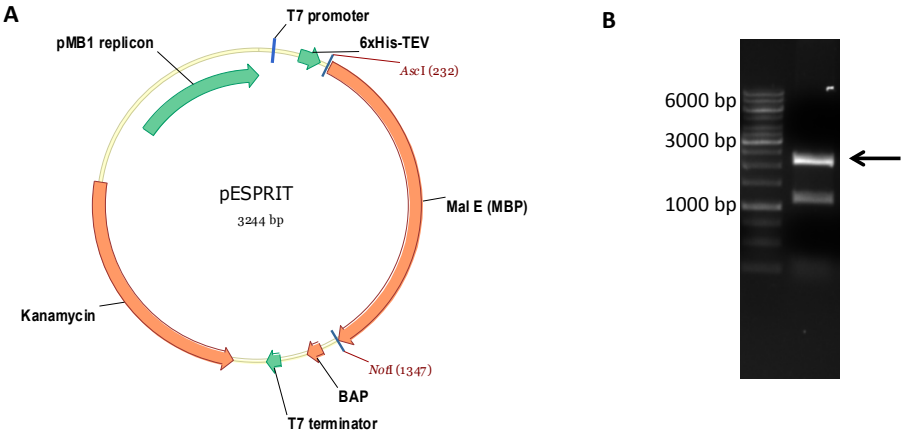


Figure 2.5 – pESPRIT digestion with AscI and NotI restriction enzymes. (A) Representation of pESPRIT vector with AscI and NotI restriction sites. Agarose gel electrophoresis separates two fragments corresponding to pESPRIT with 2100 bp and MBP with 1100 bp.

The purified fragments were ligated (Figure 2.6-A), and the construct verified by restriction analysis with HindIII/NdeI. The digested fragments were separated by agarose gel electrophoresis (Figure 2.6-B) resulted in two expected bands, one with 3360 bp corresponding to SAPAP3 insert and the other with 1717 bp corresponding to the pESPRIT vector.

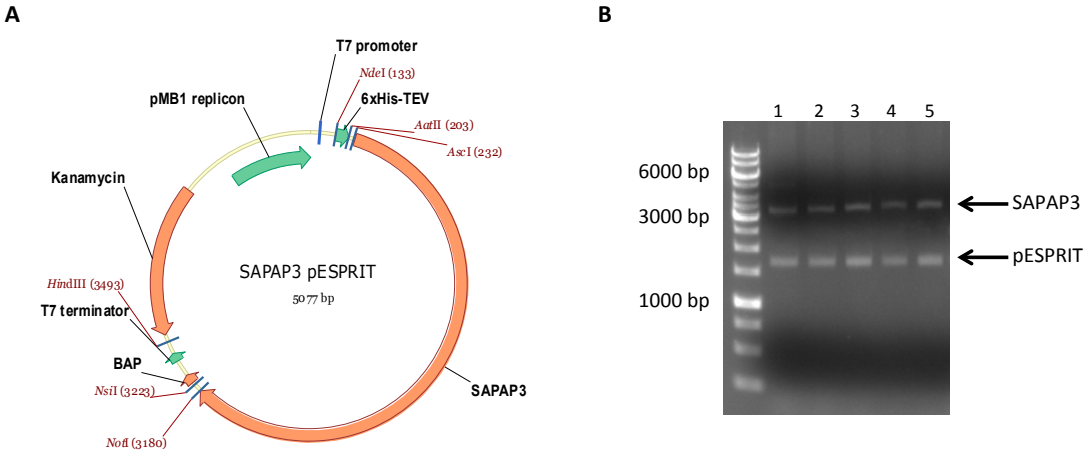


Figure 2.6 – *SAPAP3*_pESPRIT cloning. (A) Representation of *SAPAP3* gene in pESPRIT vector which contains: 6xHis-TEV: N-terminal hexahistidine tag followed by Tobacco Etch Virus protease cleavage site; BAP: Biotin acceptor peptide; NdeI and HindIII restriction site; AatII/AscI and NsiI/NotI: restriction sites for ESPRIT library preparation; Kan – kanamycin resistance gene;

A positive clone, containing the insert with the predicted size, was sent for sequencing and the results showed that *SAPAP3* cDNA was successfully inserted into pESPRIT vector in the correct orientation.

SAPAP3 sequence contains an internal 12 His sequence (Figure 2.7), which was deleted from the *SAPAP3*_pESPRIT construct by PCR, as it would interfere with the ESPRIT screening experimental strategy. After PCR based deletion, the obtained plasmid was sequenced, in order to confirm the deletion (data not shown).

MRGYHGDRGSHPRPARFADQQHMDVGPAARAPYLLGSREAFSTEPRFCAPRAGLGHISPEGPLSLSEGPSVG
PEGGPAGAGVGGGSSTFPRMYPGQGPFDTCEDCVGHPQGKGAPRLPPTLLDQFEKQLPVQQDGFHTLPYQR
GPAGAGPGPAPGTGTAPEPRSESPSRIRHLVHSVQKLFAKSHSLEAPGKRDYNGPKAEGRGGSGGDSYPGPGS
GGPHTSHHHHHHHHHHHHQSRHGKRSKSKDRKGDGRHQAKSTGWWSSDDNLDSDSGFLAGGRPPGEPGG
PFCLEGPDGSYRDLSFKGRSGGSEGRCLACTGMSMSLDGQSVKRSAWHTMMVSQGRDGYPGAGPGKGLLG
PETKAKARTYHYLQVPQDDWGGYPTGGKDGEIPCRRMRSGSYIKAMGDEESGDSDGSPKTSPKAVARRFTTR
RSSSVDQARINCCVPPRIHPRSSIPGYSRSLTTGQLSDELNQQLEAVCGSVFGELESQAVDALDLPGCFRMRSHS
YLRAIQAGCSQDDDCLPLLATPAAVSGRPGSSFNFRKAPPPIPPGSQAPPRISITAQSSTDSAHESFTAAEGPARR
CSSADGLDGPAMGARTLELAPVPPRASPKPPTLIIKTIPGREELRSLARQRKWRPSIGVQVETISDSDTENRSRRE
FHSIGVQVEEDKRRARFKRSNSVTAGVQADLELEGLAGLATVATEDKALQFGRSFQRHASEPQPGPRAPTYSVF
RTVHTQGQWAYREGYPLPYEPPATDGSPGPAPAPTPGPGAGRRDSWIERGSRSLPDSGRASPCPRDGEWFIK
MLRAEVEKLEHWCQQMEREAEDYELPEEILEKIRSAVGSTQLLLSQKVQQFFRLCQQSMDPTAFPVPTFQDLA
GFWDLLQLSIEDVTLKFLELQQLKANSWKLLEPKEEKKVPPPIPKKPLRGRGVPVKERSLDSVDRQRQEARKRLL
AAKRAASFRHSSATESADSIEIYIPEAQTRL

Figure 2.7 – SAPAP3 amino-acid sequence. Representation of the amino-acid sequence of SAPAP3 with the 12 histidine region highlighted in blue.

2.3.2 *SAPAP3* truncation library production

Comprehensive *SAPAP3* truncation library was generated using exonuclease III and mung bean nuclease (Figure 2.2) in a two step process, as previously described. After the first truncation reaction, as a mean of evaluating the library quality, the insert size range of the library constructs was analysed by colony PCR with flaking primers on 70 clones. The colony PCR result showed a linear distribution of the insert size between 100-3 000 bp (Figure 2.8) validating the first truncation step.

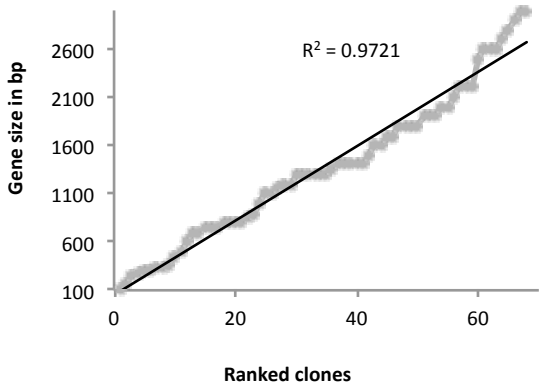


Figure 2.8 – Colony PCR analysis for library quality assessment after first truncation. Colony PCR with flaking primers on 70 clones from the library after N-terminal truncation

After the second truncation step, to reduce the number of clones to screen, and thereby improving the sampling efficiency, a subset of clones with inserts in the size range of 300-1200 bp (100-400 AA), corresponding to plasmids with a size range of 2400-bp, were isolated on agarose gels (Figure 2.9).

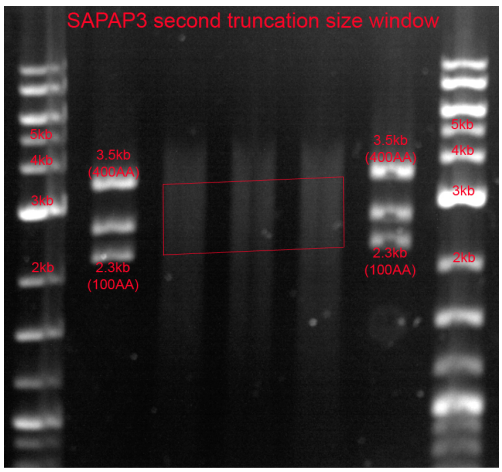


Figure 2.9 – DNA fragments separated on agarose gel electrophoresis. DNA fragments corresponding to gene inserts in the size range of 300-1200 bp (red box) were excised.

To validate the C-terminal truncation reaction, colony PCR with flanking primers on 40 clones shows a homogeneous distribution of clones (Figure 2.10). The histogram also indicates a higher and lower population of clones outside the expected range, nevertheless,

the main population of constructs lay within the range of 100-400 AA. In this way, *SAPAP3* double truncation library was validated and became ready to be screened.

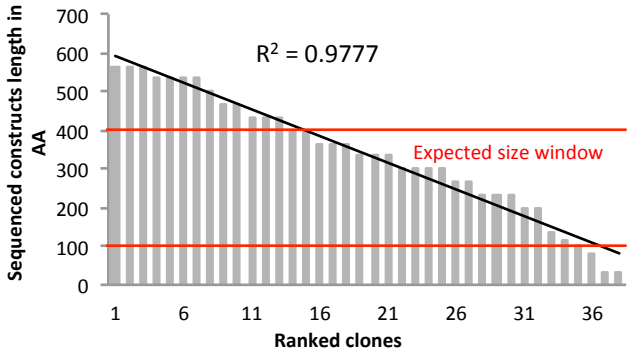


Figure 2.10 – Colony PCR analysis for library quality assessment after second truncation. Colony PCR analysis with flaking primers on 40 clones from the library after the C-terminal truncation; red lines between 100 AA and 400 AA represent the expected size range.

2.3.3 High-throughput screening of SAPAP3 putative soluble domains

The SAPAP3 truncation library was screened using standard colony picking and liquid handling robots in a workflow designed to identify any forms of the soluble library target. In a first step, 28 000 colonies from the library were picked and gridded onto a nitrocellulose membrane for expression analysis. The fluorescent signals from streptavidin detecting the C-terminal biotin acceptor peptide, and anti-hexahistidine antibodies detecting the N-terminal tag were quantified and their distribution determined (Figure 2.11).

To obtain a list of putative soluble expressing clones for further investigation, only colonies exhibiting visible hexahistidine tag signals were selected, thus eliminating truncated or degraded protein products that, even if soluble, would not be purifiable. Based upon this, the 96 positive clones with significantly higher anti-histidine tag and biotinylation derived signals were selected from the large bank of clones. Ranking of clones for their *in vivo* biotinylation efficiency allows enrichment of a population of soluble expression clones from the library, as soluble constructs are generally better substrates for this cytoplasmic post-translation modification than insoluble/aggregated targets.

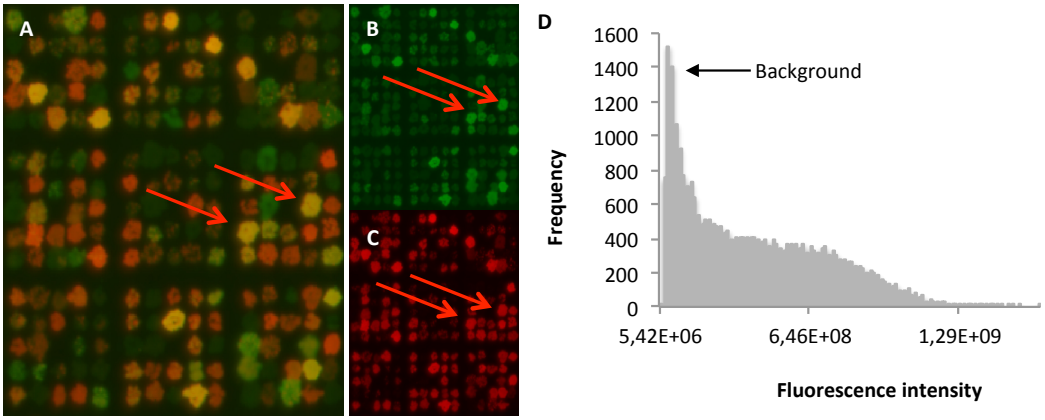


Figure 2.11 – Measurement of putative solubility and intactness of SAPAP3 constructs. Panel A: Close-up view of the array containing 28 000 clones, in which simultaneous hybridization of fluorescent streptavidin (red) and hexahistidine antibody (green) is shown; yellow colonies (such as the ones pointed with the arrows) contain both signals indicating an elevated potential for protein solubility. Panel B: Hybridization of the array with fluorescently labeled streptavidin against the C-terminal biotin acceptor peptide. Panel C: Hybridisation of the array with fluorescently labeled N-terminal antihexahistidine tag monoclonal antibody. Panel D: Frequency histogram analysis of total biotinylation signal: the graph displays the frequency of clones according to their streptavidin signals, the 96 most intensely biotinylated clones selected for subsequent expression tests exhibitied higher signals than the background.

2.3.4 Small-scale expression and purification screening of SAPAP3 soluble domains

In the second level of the screenning, the ninety-six most intense biotinylated, hexahistidine-positive clones were grown in 4 mL and protein expression induced by the addition of arabinose. After robot assisted loading into 96-well filter plates charged with Ni^{2+} NTA agarose, eluted fractions were obtained by incubation with 500 mM imidazole buffer, and analysed by streptavidin-Western blot to identify the biotinylation signals (Figure 2.12). The Western blot result showed that the 96 constructs selected as putative domains of SAPAP3 presented varying levels of protein, and that their molecular weights ranging from 15 kDa and 55 kDa are within the expected size range for domains between 100 and 400 amino acids. Using MBP as experiment control (well 96 in Figure 2.12), and although none of the clones showed a similar yield of expression as MBP, several constructs exhibited strong bands that are labelled with a red arrow in Figure 2.12. In the panels where clones 1 to 24 and 73 to 95 are displayed, it seems to exist a contamination between 25 kDa and 35 kDa, which presumably was due to a leakage from the well with domain 19 and domain 86, respectively. Nevertheless, and despite several clones present low levels of protein

intensities as already mentioned, others exhibited high amounts of detectable protein, which further validates the result of the initial screening.

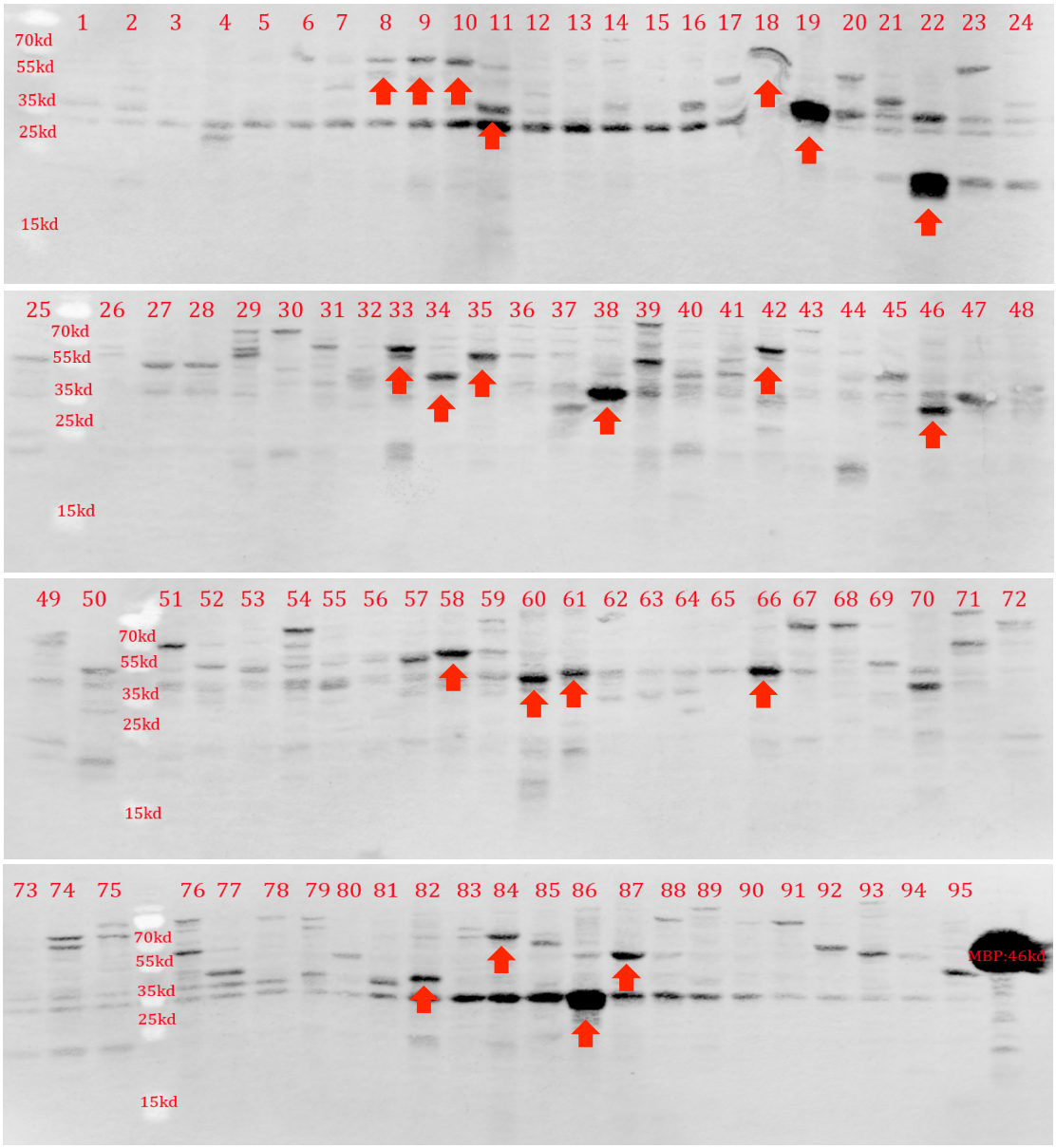


Figure 2.12 – Streptavidin-Western blot analysis of the highest ranking biotinylated domains selected in the first screen. Ninety-six constructs grown in 4 mL liquid cultures were purified by Ni^{2+} NTA purification system, and analysed by Western blot to identify the biotinylation signals through streptavidin probing. The constructs showing the highest levels of expression and purifiability are signed with a red arrow and were selected for further analysis. Biotinylated maltose binding protein (MBP) in the position 96 was used as an internal control. The molecular weight markers in kiloDaltons (kDa) are shown in the image.

The expression and purifiability of the best twenty-one constructs obtained from the small-scale screening were also assessed by SDS-PAGE. For this, the selected constructs were expressed in 4 mL liquid cultures, Ni^{2+} NTA purified and electrophoretically separated. The result of this analysis (Figure 2.13) showed a similar profile comparing to the one obtained for the Western blot (Figure 2.12). For such comparison, it must be noted that the intensities of the SDS-PAGE bands should be comparable to the intensities of the streptavidin-Western blot analysis, in which the protein signal is not amplified by a conjugated detection, but by directly detection of the protein through streptavidin binding.

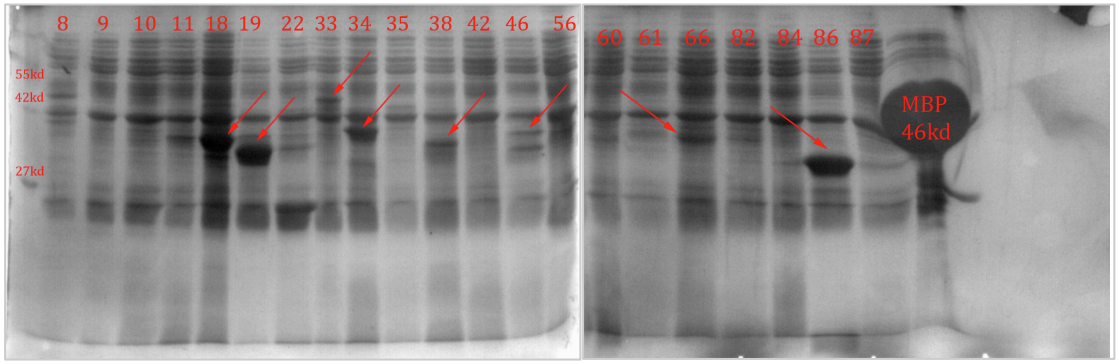


Figure 2.13 – SDS-PAGE analysis of purifiable domains selected from the streptavidin-Western blot analysis. Twenty-one constructs grown in 4 mL liquid cultures were Ni^{2+}NTA purified and analysed by SDS-PAGE to identify the highest levels of expression and purifiability. Purified proteins were stained with Coomassie blue after SDS-PAGE. The best constructs are indicated with a red arrow. Maltose binding protein (MBP) was used as an internal control. The molecular weight markers in kiloDaltons (kDa) are shown on the left.

From the 21 domains selected, the best domains obtained in the SDS-PAGE analysis were the domains 18, 19, 33, 34, 38, 46, 66 and 86 (Figure 2.13), where a specific band for each case stood out from the background. Comparing with the streptavidin-Western blot result, the best ones were coincident with the SDS-PAGE result in terms of intensities, with the exception of domains 18 and 22. The result for domain 18 in the Western blot was inconclusive as a result of a technical problem in that gel region; however, the SDS-PAGE result showed a reasonable amount of soluble protein. The domain 22 presented inconsistent results, since a high amount of protein was observed in the Western blot but not on SDS-PAGE analysis, requiring further evaluation. Nevertheless, gathering the information from the two experiments the domains dysplaying best yields of purifiable protein were the domains 18, 19, 34 and 86.

The sequence boundaries (Table 2.1) of constructs showing the highest levels of purified protein in the previous trial were determined by DNA sequencing, with the exception of constructs 11 and 46, for which the sequencing trials failed and were, therefore, excluded from further analysis.

Table 2.1 – Sequence boundaries and respective molecular weight of the selected domains.

Domain	Sequence boundries (aa.)	Molecular Weight (kDa) without tags
8	F41-439C	42.11
9	L129-511C	24.29
10	P212-606P	41.90
18	G61-438N	39.79
19	G754-968I	24.30
22	L467-613I	15.23
34	G706-967S	29.65
35	H366-708S	36.94
38	M23-449S	45.04
42	K179-595L	44.21
56	D104-244D	15.07
58	S257-511C	25.24
60	T138-433D	31.42
61	G197-474V	29.73
66	V337-599P	27.72
82	S300-578R	29.62
84	S184-610T	45.16
86	L745-941Q	22.37
87	K245-589A	36.43

The translated sequences were aligned against the full-length SAPAP3 polypeptide sequence, which is represented in Figue 2.14. The sequencing results evidenced the identification of putative soluble protein domains evenly distributed across the target gene, with a cluster of the best clones present in the C-terminus of the SAPAP3 sequence.

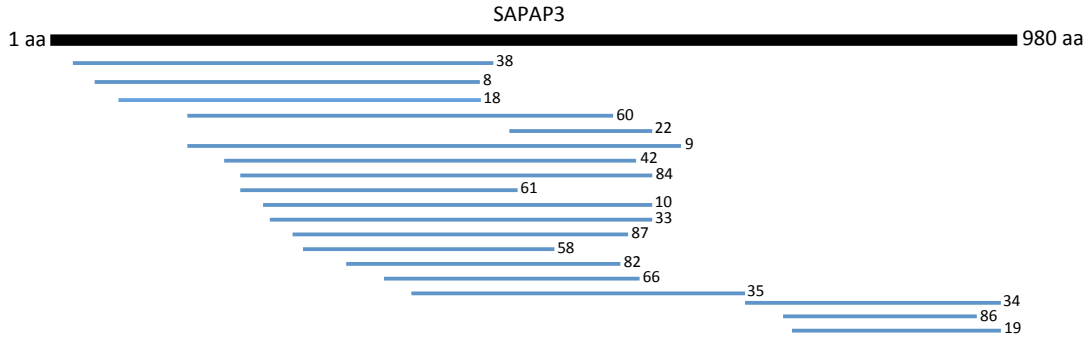


Figure 2.14 – Construct map of purifiable domains identified by sequencing analysis. The constructs identified in the small scale expression and purification screening (blue) were aligned against the full-length SAPAP3 polypeptide sequence (black).

2.3.5 Medium-scale expression and purification of selected SAPAP3 domains

To further confirm the results obtained in the small-scale expression trials, the nineteen domains sequenced and identified in the small-scale expression screening were tested in 50 mL cultures and purified on Ni^{2+} NTA agarose resin. The protein was eluted in two imidazole concentration steps (100 mM and 300 mM) and analysed by SDS-PAGE (Figure 2.15) and Western blot using an anti-histidine tag antibody (Figure 2.16).

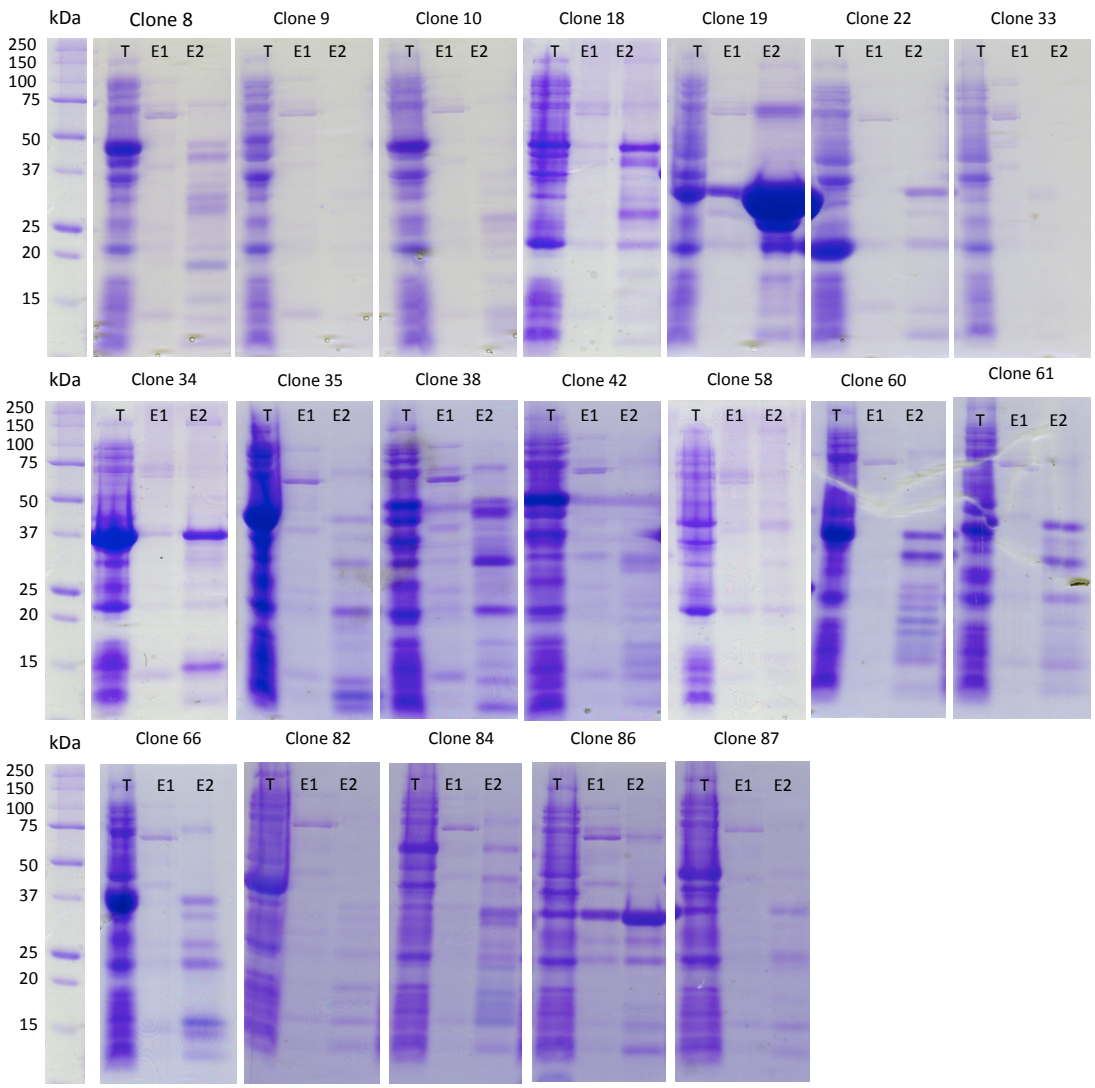


Figure 2.15 – SDS-PAGE analysis after medium-scale expression and purification of the best 19 SAPAP3 clones selected from ESPRIT technology. Selected clones were expressed in 50 mL cultures, purified by nickel affinity chromatography and analysed by SDS-PAGE. Annotations are: T - total fraction; E1 - eluted fraction with 100 mM imidazole; E2 - eluted fraction with 300 mM imidazole. The molecular weight markers in kiloDaltons (kDa) are shown on the left.

The results from SDS-PAGE analysis showed that domains 18 (44 kDa), 19 (30 kDa), 34 (35 kDa), 56 (19 kDa) and 86 (28 kDa) were the ones where larger amounts of purified soluble protein were obtained. This result is in accordance with the results of the small-scale expression and purification, where these same domains were identified as the most promising. The domain 22, whose results were controversial in the small-scale test, revealed the presence of a low intensity protein band. In general, the remaining clones presented lower amounts of soluble protein when compared to the small-scale tests. Moreover, the

SDS-PAGE analysis of the medium-scale experiments revealed additional bands when compared to the SDS-PAGE of small scale, probably resulting from protein truncation.

The yields of soluble protein and the presence of truncation products were also examined by anti-histidine Western blot analysis (Figure 2.16).

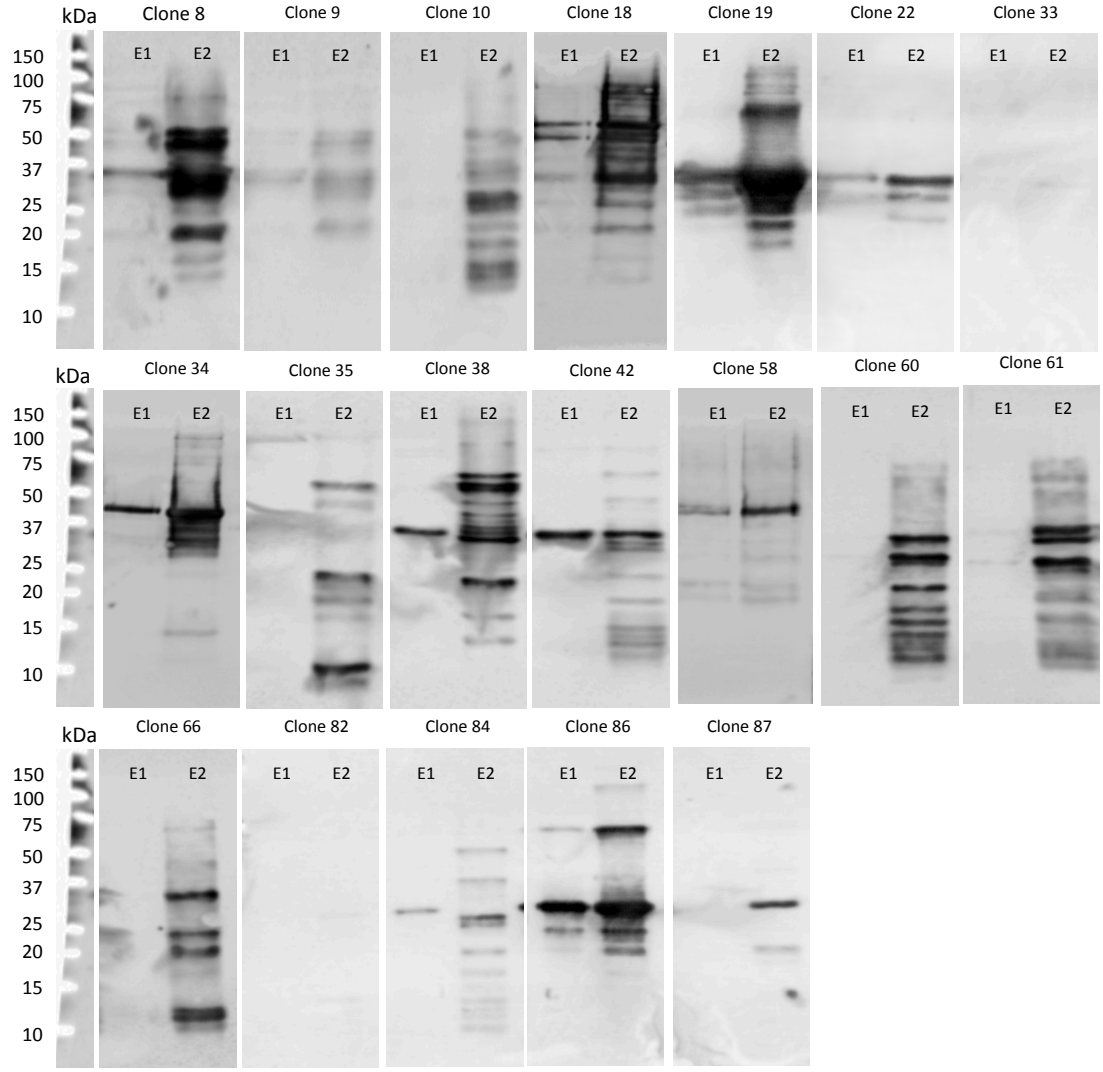


Figure 2.16 – Western blot analysis with anti-histidine antibody after medium scale expression and purification of the best 19 clones of SAPAP3 selected from ESPRIT technology. Selected clones were expressed in 50 mL cultures, purified by nickel affinity chromatography and analysed by Western blot with anti-histidine antibody. Annotations are: E1 - elected fraction with 100 mM imidazole; E2 - eluted fraction with 300 mM imidazole. The molecular weight markers in kilodaltons (kDa) are shown on the left.

The Western blot using anti-hexahistidine antibody showed that domains 18, 19, 34 and 86 presented the best yields of purifiable protein, and the presence of various degradation products. The domain 8 also showed a high Western blot signal, however this result does not agree neither with the SDS-PAGE result nor with the results of the small-scale

experiments. The other clones present faint or lower amounts of soluble protein as well as the presence of several degradation products. While in the SDS-PAGE of the first small-scale test (Figure 2.13) some extra protein bands were also detected, the initial streptavidin-Western blot analysis (Figure 2.12) was not able to identify protein degradation variants, as the detection was being done through the C-terminal tag.

On top of that, in the medium scale experiment, the cell lysis and protein extraction and purification was done in the absence of protease inhibitors, which were present in the small-scale, indicating the need to include such compounds in buffers during protein purification.

Nevertheless, both experiments revealed that domains 18, 19, 34 and 86 were the best candidates in terms of yield of purifiable protein.

Sequence alignment of these five constructs against the full-length SAPAP3 polypeptide sequence (Figure 2.17) revealed that domains 18 cover part of SAPAP3 N-terminus while domains 19, 34 and 86 cover part of the C-terminus of SAPAP3.

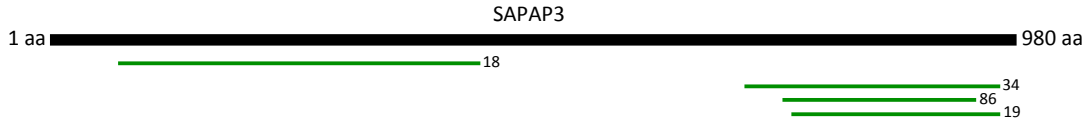


Figure 2.17 – Alignment of domains showing the highest levels of expression and purification. The constructs (green) were aligned against the full-length SAPAP3 polypeptide sequence (black).

In conclusion, since the results of the domains from the C-terminus of SAPAP3 (domains 19, 34 and 86) were consistent over the entire analysis perform, and since domain 19 was the domain that presented the highest yield of purified soluble protein, it was the domain selected for scale-up expression and purification.

2.3.6 Scale-up expression and purification of SAPAP3 domain 19

A protocol for the expression and purification of SAPAP3 domain 19 was optimized involving three chromatografic steps. In the first step, as all the SAPAP3 ESPRIT derived constructs coded for an hexahistidine N-terminal tag, an affinity chromatography was used - the protein resulting from 1 L expression was captured in a 5 mL HisTrap HP column (GE Healthcare) and a major peak was consistently eluted at 150 mM imidazole elution step, as can be observed in the representative chromatogram of Figure 2.18-A. The SDS-PAGE analysis of the protein fraction eluted with 150 mM imidazole (Figure 2.18-D) revealed the

presence of a major band with an aparent molecular mass of 27 kDa, which is the predicted molecular mass of the domain 19 of SAPAP3; however the protein is co-eluting with several other proteins, justifying the need of additional purification steps.

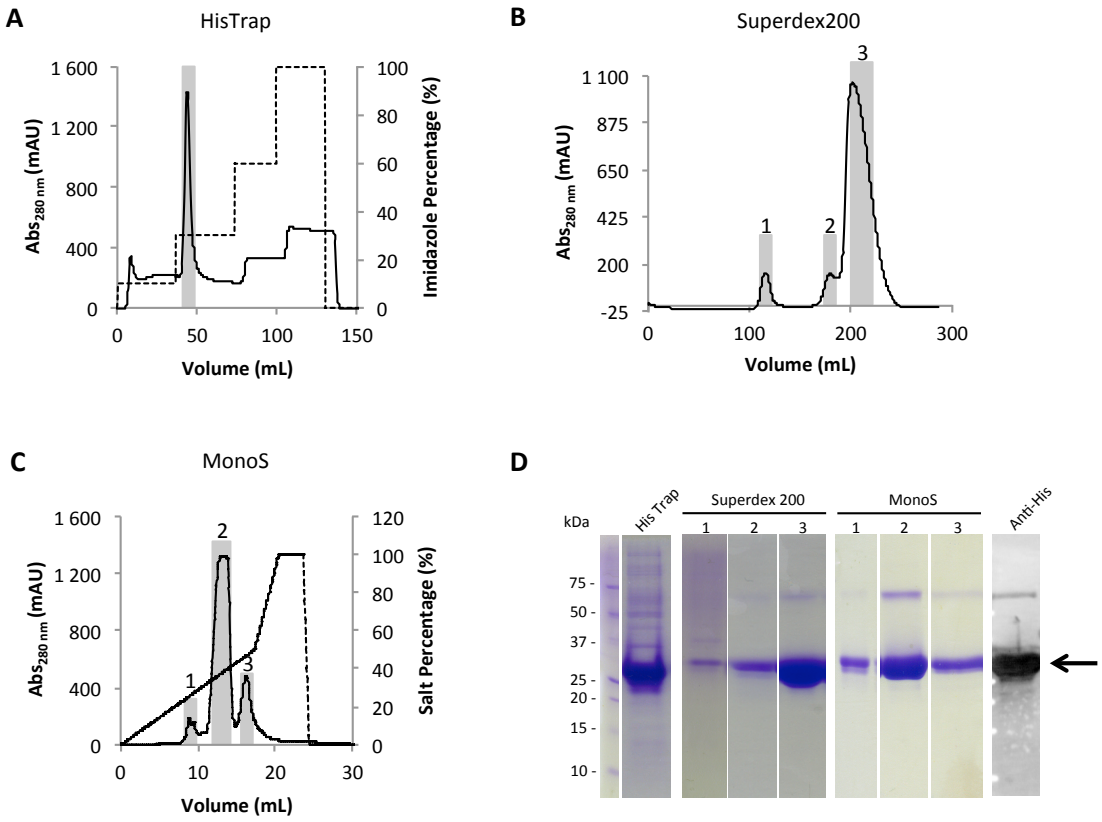


Figure 2.18 – Purification of domain 19 from SAPAP3. (A) Chromatogram of purification of domain 19 from SAPAP3 using a HisTrap 5 mL column equilibrated with 20 mM phosphate buffer pH 7.4, 500 mM NaCl, 10 % glycerol and 25 mM imidazole. Protein was eluted with increasing concentrations of imidazole (50 mM, 150 mM, 300 mM and 500 mM) at a flow rate of 5 mL/min and monitored by measuring the Abs_{280nm}. Fractions from HisTrap eluted with 150 mM imidazole (grey bar) were applied to a Superdex 200 HiLoad 26/60 column **(B),** equilibrated with 20 mM phosphate buffer pH 7.4, 150 mM NaCl and 10 % glycerol at a flow rate of 2 mL/min. Protein elution was monitored by measuring the A_{280nm}. **(C)** Protein fractions from Superdex 200 HiLoad 26/60 (grey bar 3) were applied to a MonoS 5/50 GL column equilibrated in 20 mM phosphate buffer pH 7.4 and 10 % glycerol. The elution was carried out in a gradient of NaCl (0 - 1 M), at a flow rate of 0.75 mL/min and monitored by measuring the A_{280nm}. **(D)** The protein fractions were analysed by SDS-PAGE. The positions of the molecular weight markers, and their sizes in kiloDaltons (kDa) are shown on the left. The eluted protein from ionic exchange chromatography was also analysed by Western blot with anti-histidine antibody. The expected molecular weight for SAPAP3 domain 19 is indicated with the black arrow.

The chosen method for the intermediate purification step was a SEC (size exclusion chromatography) using a HiLoad Superdex 200 26/60 pg preparative column (GE Healthcare), which was used not only to remove the contaminant proteins co-eluted in the

previous affinity chromatography step, but also to evaluate the oligomerization state of the protein and to remove possible aggregated forms.

The SEC chromatogram (Figure 2.18-B) has shown that the protein was mainly eluted, according to the calibration of the column, at a volume compatible with the monomeric form (grey box number 3 in Figure 2.18-B) and a small percentage of the protein eluted at a volume compatible with the dimeric form of the protein (grey box number 2 in Figure 2.18-B); some protein was also eluted in the void volume of the column (grey box number 1 in Figure 2.18-B), showing that protein aggregates could be separated. The SDS-PAGE analysis of the protein sample from the third peak (eluting at 210 mL) of this chromatographic step (Figure 2.18-D) revealed that SEC was effective in removing the observed contaminants co-eluting in the affinity step.

The fractions from the last SEC peak corresponding to the monomeric form of the protein were then applied to a Mono S 5/50 GL column, and eluted by a gradient of 0 - 1 M NaCl (Figure 2.18-C). The cationic exchange chromatogram revealed that the protein applied to the column eluted mainly as a homogeneous peak (peak 2) with salt concentration between 350-400 mM. Two other minor peaks were eluted (peaks 1 and 3), which presented similar SDS-PAGE profiles (Figure 2.18-D), indicating different conformational states of the protein population. This also revealed the usefulness of this last high-resolution polishing step to obtain highly pure and highly homogeneous protein.

The Western blot analysis of the protein eluted in the peak 2 of the cation exchange step with anti-hexahistidine antibody (Figure 2.18-D) revealed the presence of two recognized protein bands: one apparently corresponding to the monomeric form of the SAPAP3 domain 19 with an apparent molecular mass around 27 kDa, and the other with twice that molecular weight. To assay if the extra band could be a possible dimer resiliant to denaturation, it was tested the effect of removing the reducing agent during protein denaturation, as well as the effect of the presense of DTT (0.1 M) and beta-mercaptoethanol in a higher concentration (5 % (v/v)) than the previously used (1.8 % (v/v)), again the SDS-PAGE revealed two bands for all the conditions (data not shown). Considering these results and since the protein is eluted in the monomeric form during the SEC phase, we believe that the extra band may be an artifact that is known to occur during the SDS-PAGE by the formation of reactive oxygen species (ROS)[153]. In SDS-PAGE method, there are several sources of ROS including (i) APS, a strong oxidant used for gel polymerization; (ii) the molecular oxygen present during gel polymerization; and (iii) trace oxidative contaminants present in the gel or formed during electrophoresis[153]. These species can oxidize protein residues that consequently might

promote the formation of covalent intermolecular bonds and the presence of SDS-PAGE bands with non expectable molecular weight. Nevertheless, the monomeric nature of SAPAP domain 19 should be further validated using complementary techniques, such as light scattering, analytical ultracentrifugation or mass spectrometry under native conditions.

2.3.7 Expression and purification of SAPAP3 domain 19 K910R and R770L mutants

Within the sequence of SAPAP3 domain 19 (754-968 aa) falls one missense mutation reported to be associated with OCD (K910R)[125] and one associated with schizophrenia (R770L)[127]. These two mutants were independently generated as described in section 2.2.9, expressed in *E. coli* BL21 AI strain and purified in the same way as the wild-type (WT) SAPAP3 domain 19. For the mutants, the chromatographic and SDS-PAGE profiles of each purification step were very similar to those obtained for the wild-type form (Figure 2.19 and 2.20), including the appearance of the high molecular weight band around 60 kDa, which is believe to be an experimental artifact, as previously discussed.

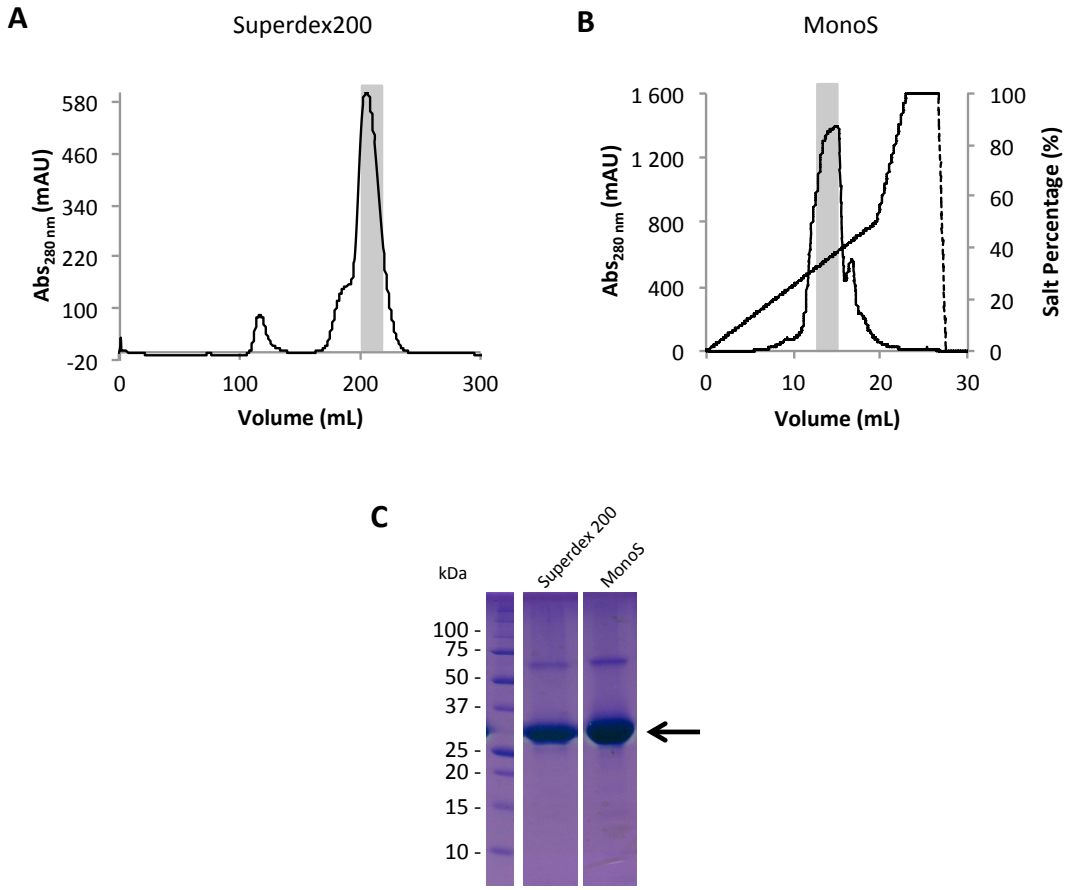


Figure 2.19 – Purification of mutant K910R of SAPAP3 domain 19. Recombinant K910R of SAPAP3 domain 19 was first purified on a HisTrap 5 mL column, followed by **(A)** Superdex 200 HiLoad 26/60 purification step, as described for wild-type domain (Figure 2.18). **(B)** Protein elution on MonoS 5/50 GL column (cation exchange chromatography) was carried out by a salt gradient (0-1 M NaCl), at a flow rate of 0.75 mL/min and monitored by measuring the A_{280nm}. **(C)** The eluted protein outlined by the grey box in (A) and the grey box in (B) was analysed by SDS-PAGE stained with Coomassie blue. The molecular weight marker in kilodaltons (kDa) is shown on the left. The expected molecular weight for SAPAP3 domain 19 is indicated with the black arrow.

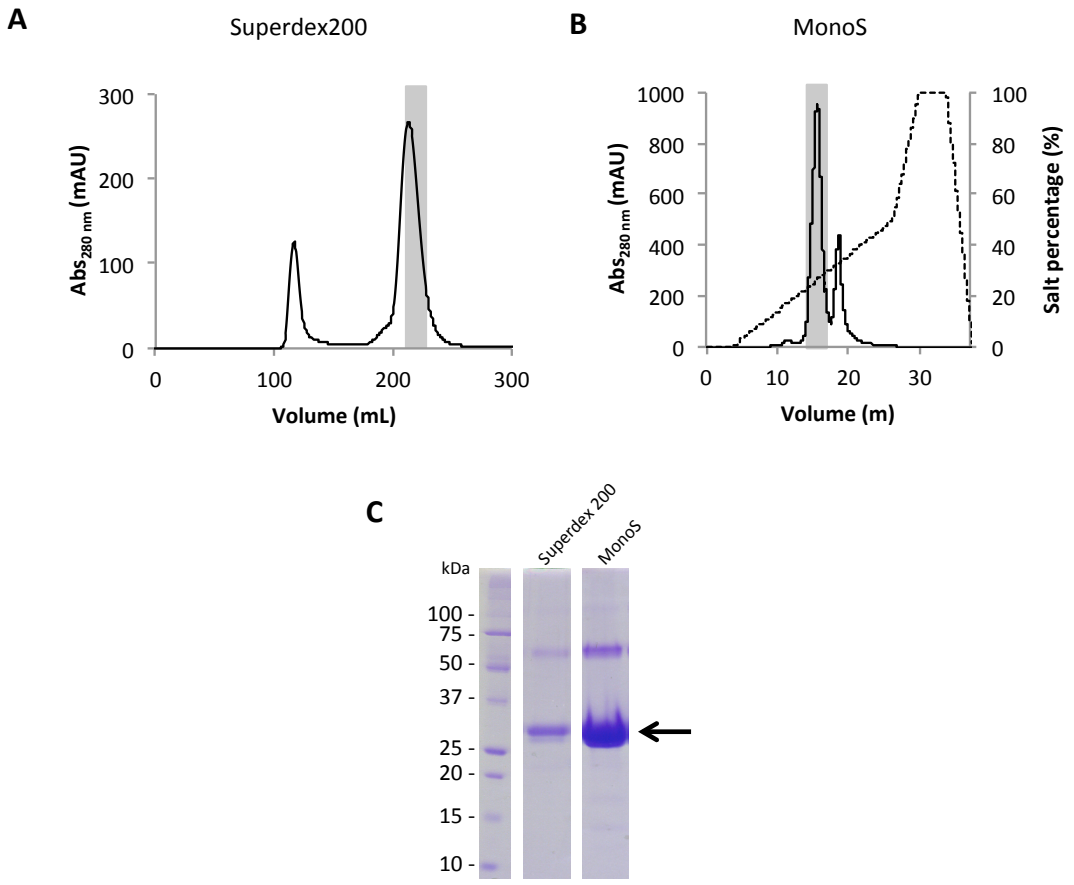


Figure 2.20 – Purification of mutant R770L of SAPAP3 domain 19. Recombinant R770L of SAPAP3 domain 19 was first purified on a HisTrap 5 mL column, as described for wild-type domain (Figure 2.18), and the eluted protein was loaded on a **(A)** Superdex 200 HiLoad 26/60 column equilibrated with 20 mM phosphate buffer pH 7.4, 150 mM NaCl and 10 % glycerol at a flow rate of 2 mL/min. **(B)** Protein corresponding to the peak outlined by the grey box in (A) was applied on MonoS 5/50 GL column and the protein eluted by a salt gradient (0-1 M NaCl), at a flow rate of 0.75 mL/min and monitored by A_{280nm}. **(C)** The eluted protein from the different purification steps was analysed by SDS-PAGE stained with Coomassie blue. The molecular weight marker in kilodaltons (kDa) is shown on the left. The expected molecular weight for SAPAP3 domain 19 is indicated with the black arrow.

2.3.8 Characterization of WT SAPAP3 domain 19, and the K910R and R770L mutant forms

To further characterize the recombinant form of WT SAPAP3 domain 19, K910R and R770L mutant forms for downstream applications, the oligomeric state of the proteins was evaluated by analytical size exclusion chromatography and secondary structure content evaluated by circular dichroism.

The analytical size exclusion chromatography results (Figure 2.21) clearly evidence that both forms of SAPAP3 domain 19 are present as monodisperse proteins, and that, according to the column calibration, all forms of the protein eluted as monomers, with a calculated

apparent molecular weight of 28 kDa, which along with the very similar chromatographic profiles during protein purification, is a strong evidence that the two mutations have not affected significantly the protein conformation.

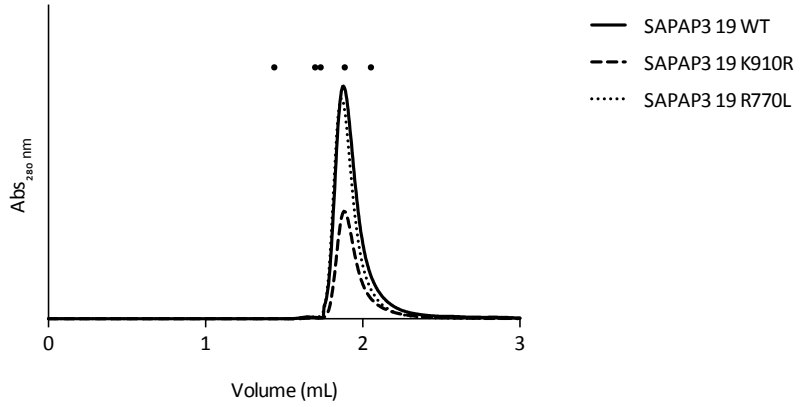


Figure 2.21 – Analytical size exclusion chromatogram of purified SAPAP3 domain 19 WT, K910 and R770L. Twenty micrograms of WT SAPAP3 and SAPAP3 R770L and ten micrograms of SAPAP3 R770L were applied in a Superdex 200 5/150 GL column (GE Healthcare) equilibrated with 20 mM phosphate buffer pH 7.4, 300 mM NaCl and 10 % glycerol at a flow rate of 0.4 mL/min. Protein elution was monitored by measuring the A_{280nm}. The black dots refer to elution volumes of molecular mass markers used for calibration. From left to right: aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa).

To assess WT SAPAP3 domain 19 secondary structure content and to further validate the absence of major structural changes induced by the disease related mutations on the folded state of SAPAP3 domain, far-UV circular dichroism analysis was performed. This spectroscopic technique is able to quantify protein secondary structure, as different secondary structural elements present characteristic CD spectra (Figure 2.22).

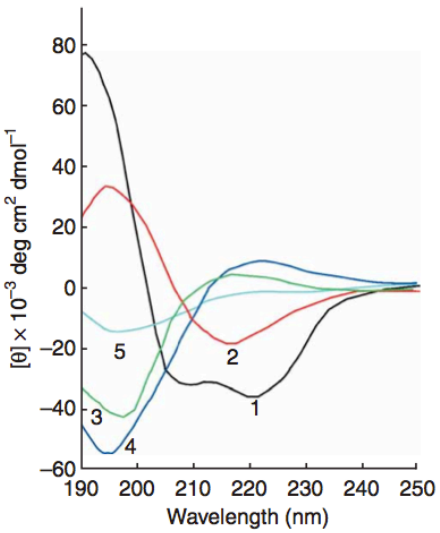


Figure 2.22 – Characteristic CD spectrum of proteins. Representative secondary structures of (1) α-helix; (2) β-sheet; (3) random coil structures. As an example placental collagen in its native triple-helical (4) and denatured forms (5) are also presented. Adapted from Greenfield *et al.*[154].

The CD spectra of WT SAPAP3 domain 19 suggests that the secondary structure is mainly composed of α-helix, since two negative peaks at 222 nm and 208 nm and a positive peak at 193 nm are present (Figure 2.23).

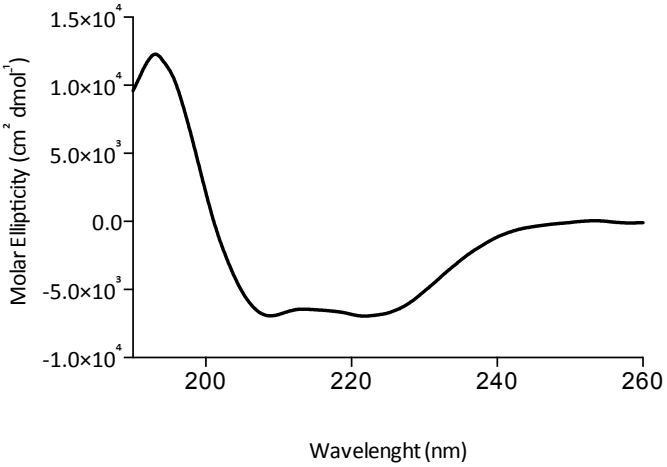


Figure 2.23 – Circular dichroism spectra of recombinant WT SAPAP3 domain 19. Molar Ellipticity in the far-UV range (190 – 260 nm) is plotted for SAPAP3 domain 19 WT at 0.4 mg/mL. The spectrum is an average of six different experiments with baseline correction.

The deconvolution analysis of the CD spectra using CDNN software, which consisted in the comparison of the CD spectrum with reference spectra of proteins with known three dimensional structures, confirmed the presence of mostly alpha helical elements of secondary structure, accounting for 37.1 % - 38.2 %, the remaining secondary structural elements accounted for 2.6 % - 9.0 % antiparallel beta sheet; 6.4 % - 8.5 % parallel beta sheet; 18.0 % - 18.0 % beta turn and 28.7 % - 31.8 % random coil.

The CD analysis of the protein mutants displays a very similar spectra pattern (Figure 2.24) when compared to the wild-type form presenting also two negative bands at 222 nm and 208 nm and a positive band at 193 nm, suggesting that the content of the secondary structure is maintained for the K901R and R770L mutants.

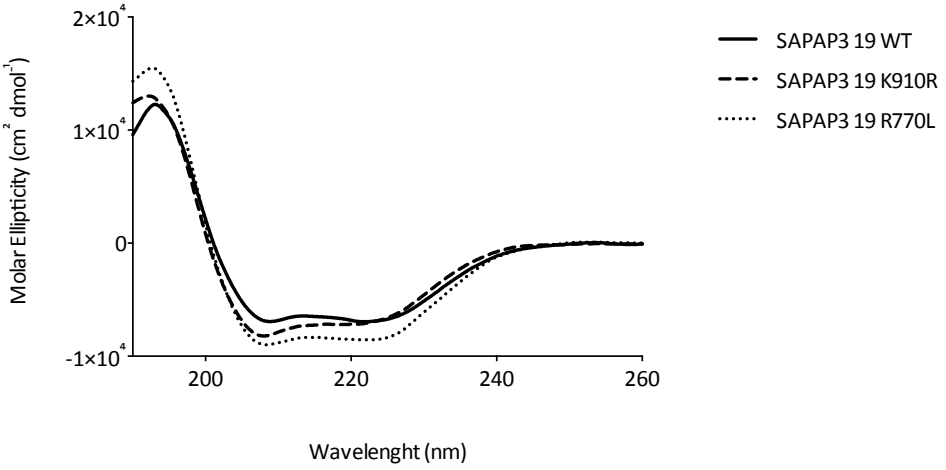


Figure 2.24 – Circular dichroism spectrums of recombinant SAPAP3 domain 19 WT, K910R and R770L comparison. Molar Ellipticity in the far-UV range (190 – 260 nm) is plotted for SAPAP3 domain 19 WT, K910R and R770L at 0.1 mg/mL. The spectrum is an average of three different experiments with baseline correction.

2.3.9 Preliminary results of crystallization screenings of SAPAP3 domain 19

The quality assessment of SAPAP3 domain 19 suggested that the preparation had the high high quality in terms of purity, homogeneity and folding needed to set up crystallization trials for x-ray diffraction structural determination. In this way, it was established a collaboration with the High Throughput Crystallization (HTX) Laboratory, at EMBL Grenoble for screening SAPAP3 domain 19 crystallization conditions. The experiments were performed with a Cartesian PixSys 4200 crystallization robot at 20 °C (Genomics Solutions, U.K.) using the sitting drop method, to test a collection of diverse conditions of buffers, precipitants and other additives using comercial screen plates (Table 2.2).

Table 2.2 – Crystallization conditions screened in High Throughput Crystallization Laboratory at EMBL Grenoble with SAPAP3 domain 19.

Screen	Supplier
The Classics	Qiagen/Nextal
Crystal Screen Lite & PEG/Ion	Hampton Research
MembFac & Natrix	Hampton Research
QuickScreen & Grid screens	Hampton Research – Home Made
Grid screens PEG 6K, PEG/LiCl	Hampton Research – Home Made
Index Screen	Hampton Research

So far, initial results showed that two promising conditions from the Classics Suit screen plate were obtained for SAPAP3 domain 19 at 6.7 mg/mL. In one of these conditions using 0.1 M HEPES pH 7.5 and 1.5 M Lithium Sulfate, some putative crystals appeared at day 34 (Figure 2.25-A and B); in the other using 0.1 M Tris pH 8.5, 1 M Lithium Sulfate; 0.01 M Nickel (II) Chloride at the thirty-fourth day putative crystals were also start growing (Figure 2.26-A and B). However, these are preliminary data and attempts are now being carried out to confirm the protein nature of the putative crystals obtained, to assess if those crystals can diffract X-ray beams.

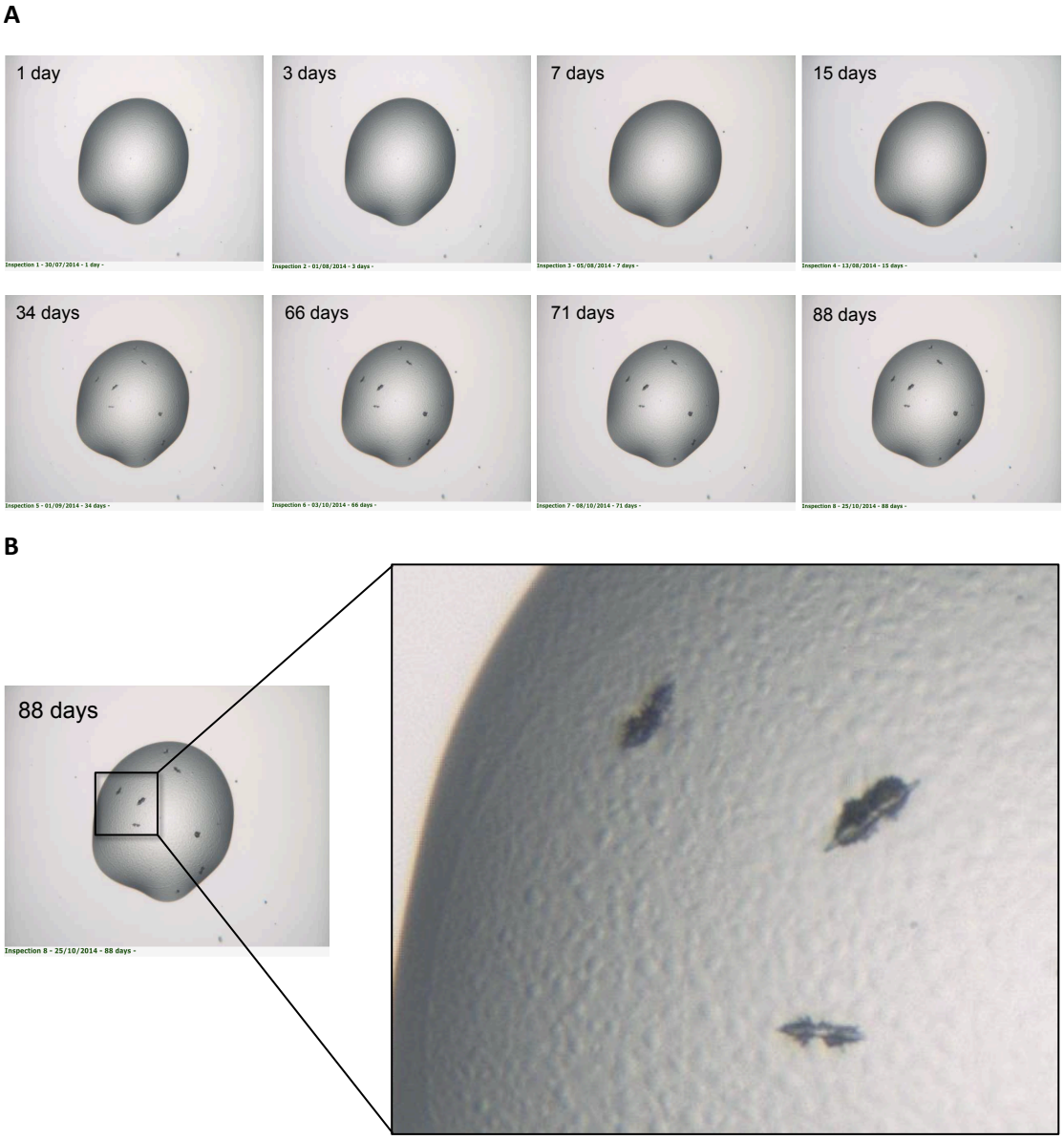
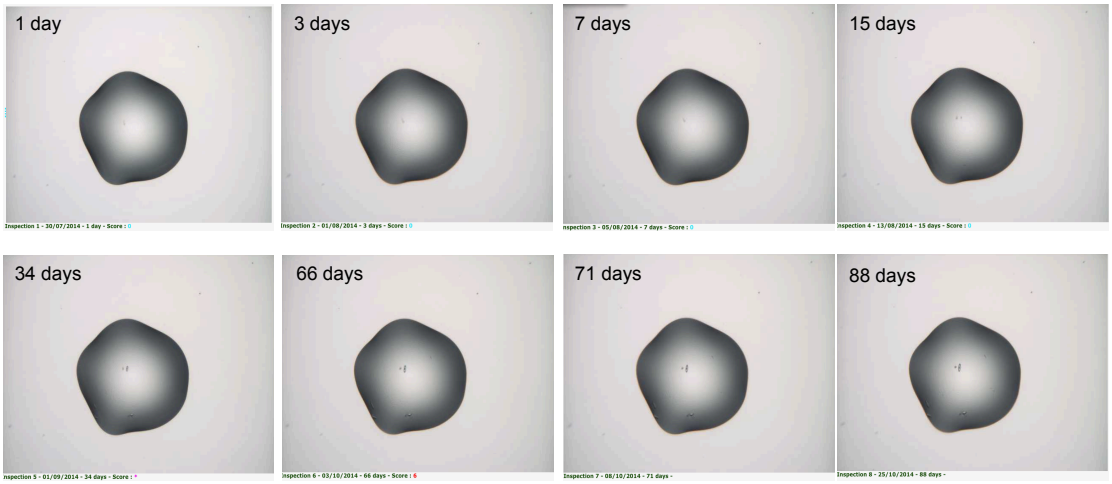


Figure 2.25 – Putative crystals of SAPAP3 domain 19: (A) Periodic images of crystallization drop of condition 0.1 M HEPES pH 7.5 and 1.5 M Lithium Sulfate for SAPAP3 domain 19 at 6.7 mg/mL acquired with the 20 °C RockImager (Formulatrix, Inc., U.S.). **(B)** Magnified view of the crystallization condition.

A



B

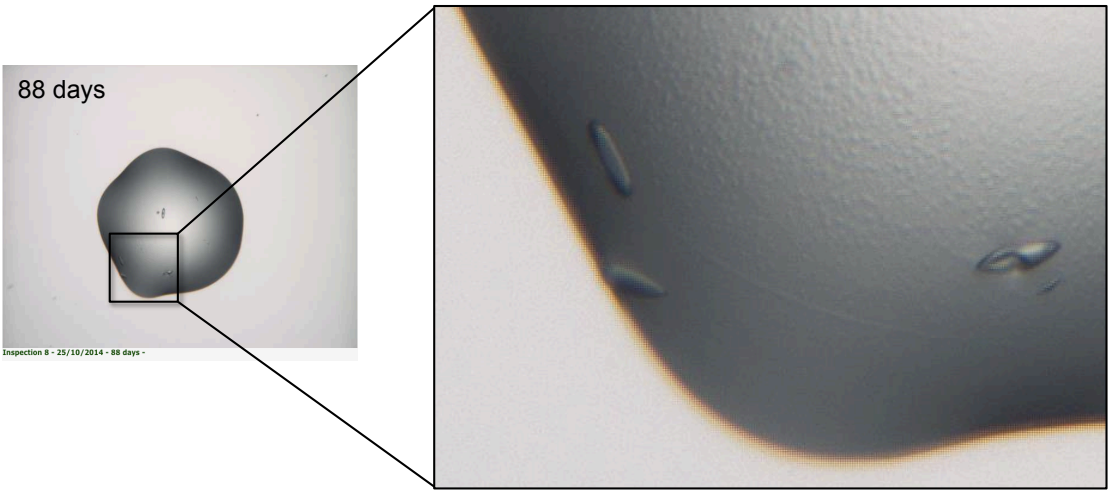


Figure 2.26 – Putative crystals of SAPAP3 domain 19: (A) Periodic images of crystallization drop ofcondition 0.1M Tris pH 8.5, 1M Lithium Sulfate; 0.01M Nickel (II) Chloride for SAPAP3 domain 19 at 6.7 mg/mL acquired with the 20 °C RockImager (Formulatrix, Inc., U.S.). **(B)** Magnified view of the crystallization condition.

2.4 Discussion

SAPAP3, a postsynaptic density scaffolding protein, has not yet been produced in heterologous systems. The reason for this fact could be due to its large size and lack of annotated domains when its protein sequence is analyzed by protein sequence motif analysis algorithms. One common strategy to overcome this limitation is to produce shorter genetic constructs corresponding to single or multi-domains of the protein, however for the case of SAPAP3 it was not clear where those domain boundaries would be located.

To address this question, the access to ESPRIT technology at Dr. Darren Hart laboratory was of pivotal importance for the success of this work. The technology is based on directed evolution strategies, and produces a diversity of clones that would be extremely difficult to achieve using classic PCR cloning strategies, both because of the need to handle tens of thousands of clones, and also of the price of reagents. With ESPRIT, all constructs of the target gene are obtained in a single reaction tube and plasmid molecules clonally separated by bacterial transformation and robotic colony picking.

The random nature of the library leads to all possible reading frame insertion combinations, which renders only one third of the 28 000 clones in frame with the initial Met and hexahistidine tag. For those, only one third is in frame with the 3' end BAP tag coding sequence, rendering the number of valuable clones to one ninth of the 28 000, roughly above 3 000 clones. This is only a small fraction (1.4 %) of the variability needed to cover all the possible combinations of putative domains with sizes between 100-400 residues in a 979 amino acid protein, which would need over 200 000 clones for a 1 x coverage. Nevertheless, as domain solubility is many times observed for a relatively wide range of domain boundaries combinations, the coverage we got was somehow suitable to find a reasonable number of domains covering an extensive region of SAPAP3 sequence, again reinforcing the potential of ESPRIT technique to uncover soluble domains out of not annotated proteins, as was the case of SAPAP3.

To overcome the low coverage of ESPRIT, which is due to the large size of SAPAP3, this work can be pursued in the future with a second round of ESPRIT screenings, focused on the N-terminal and central part of SAPAP sequence, and/or by using an improved version of ESPRIT technology, that was not available for the type of access we got through BioStruct-X funding, designated by ESPRIT ORF-selector[155], which uses an intein-based technology to select for the genetic inserts in the proper reading frame, thus increasing the library variability coverage.

At the end of the screening it was possible to have a good coverage of several domains amenable to expression (clones 18, 19, 34 and 86), with clone 19 from the C-terminal domain residues 754-968, the one yielding more protein, and thus selected for scale-up production, purification and down-stream characterization.

During the development of the current project, a structure of SAPAP1 from Rattus Norvegicus was resolved covering the C-terminus region (residues 807-971) of the protein (Figure 2.27), which was shown to display a three helix bundle, connected by short flexible loops[99].

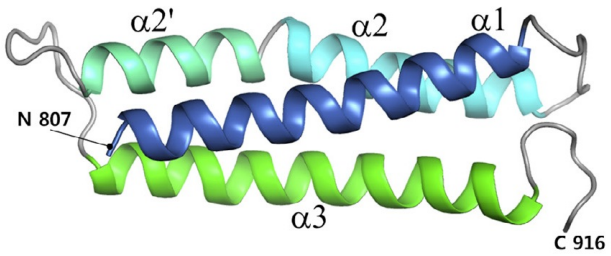


Figure 2.27 – Crystallographic structure of *Rattus norvegicus* SAPAP1 C-terminal domain. Structure of the C-terminus domain of SAPAP1 was determined at 2.0 Å resolution. The helices were coloured from blue to green. Adapted from Tong *et al.*[99].

The authors suggested to study this region of the protein based on conservation of the C-terminus across all the members of SAPAP family, indicating a presumably discrete domain with a functional significance. Interestingly, the domain selected in the current project (residues 754-968) encompasses a closely related domain with the one proposed by Tong *et al.*[99] for SAPAP1 protein (residues 807-971). The alignment between the two sequences shown an identity of 46 % and is represented in Figure 2.28.

```
SAPAP1_RAT    807   ------------------------------------DGHWFLKLLQAERDRMEGWC 826
SAPAP3_HUMAN  754   GSPGPAPAPTPGPGAGRRDSWIERGSRSLPDSGRASPCPRDGEWFIKMLRAEVEKLEHWC 813
                                                        **.**:*:*:** :::* **

SAPAP1_RAT    827   KQMEREERENNLPEDILGKIRTAVGSAQLLMAQKFYQFRELCEENLNPNAHPRPTSQDLA 886
SAPAP3_HUMAN  814   QQMEREAEDYELPEEILEKIRSAVGSTQLLLSQKVQQFFRLCQQSMDPTAFPVPTFQDLA 873
                    :***** .: :***:** ***:****:***::**. ** .**::.::*.*.* ** ****

SAPAP1_RAT    887   GFWDMLQLSIENISMKFDELHQLKANNWKQMDPLDKKERRAPPPVPKKPAKGPAPLIRER 946
SAPAP3_HUMAN  874   GFWDLLQLSIEDVTLKFLELQQLKANSWKLLEPKE--EKKVPPPIPKKPLRGRGVPVKER 931
                    ****:******::::** **:*****.** ::* :  *::.***:**** :* .  ::**

SAPAP1_RAT    947   SLESS--QRQEARKRLMAAKRAASVRQ---------- 971
SAPAP3_HUMAN  932   SLDSVDRQRQEARKRLLAAKRAASFRHSSATESADSI 968
                    **:*   *********:*******.*:
```

Figure 2.28 – Alignment of *Rattus norvegicus* SAPAP1 and *H. sapiens* SAPAP3 C-terminus regions. Sequences are shown in single letter code and numbered on the right. Positions that have a single, fully conserved residue are signed with an asterisk; conservation between groups of weakly similar properties is signed with one dot; and conservation between groups of strongly similar properties is signed with a colon. The alignment was performed in Clustal Omega through Uniprot website[100].

The scale-up production of WT SAPAP3 domain 19 and its mutant forms was successfully accomplished. Next, as no known physiological activity is known for the C-terminal domain of SAPAP, which could be used as an assay to evaluate protein quality, we performed a protein quality characterization, using analytical size exclusion chromatography and circular dichroism. The analytical SEC results confirmed the monodispersity of the purified protein, which was somehow expected from the isolated peak purified under the high-performance and high-resolution cation exchange chromatography. The analytical SEC, also confirmed the monomeric state of SAPAP3 domain 19 and related mutants (Figure 2.21), which is in accordance with the three dimensional crystallographic results obtained by Tong *et al.*[99] for SAPAP1 C-terminal domain.

The analytical SEC analysis was complemented by the analysis of secondary structure of SAPAP3 domain 19 (Figure 2.23). The spectral deconvolution has shown that SAPAP3 domain 19 is mostly comprised by alpha-helical elements, which is in accordance both with the structural results obtained by Tong *et al.*[99] for SAPAP1 C-terminal domain, and with the predicted secondary structure obtained with Quick2D prediction tool (Max-Planck Institute for Developmental Biology) for SAPAP3 domain 19 (Figure 2.29).

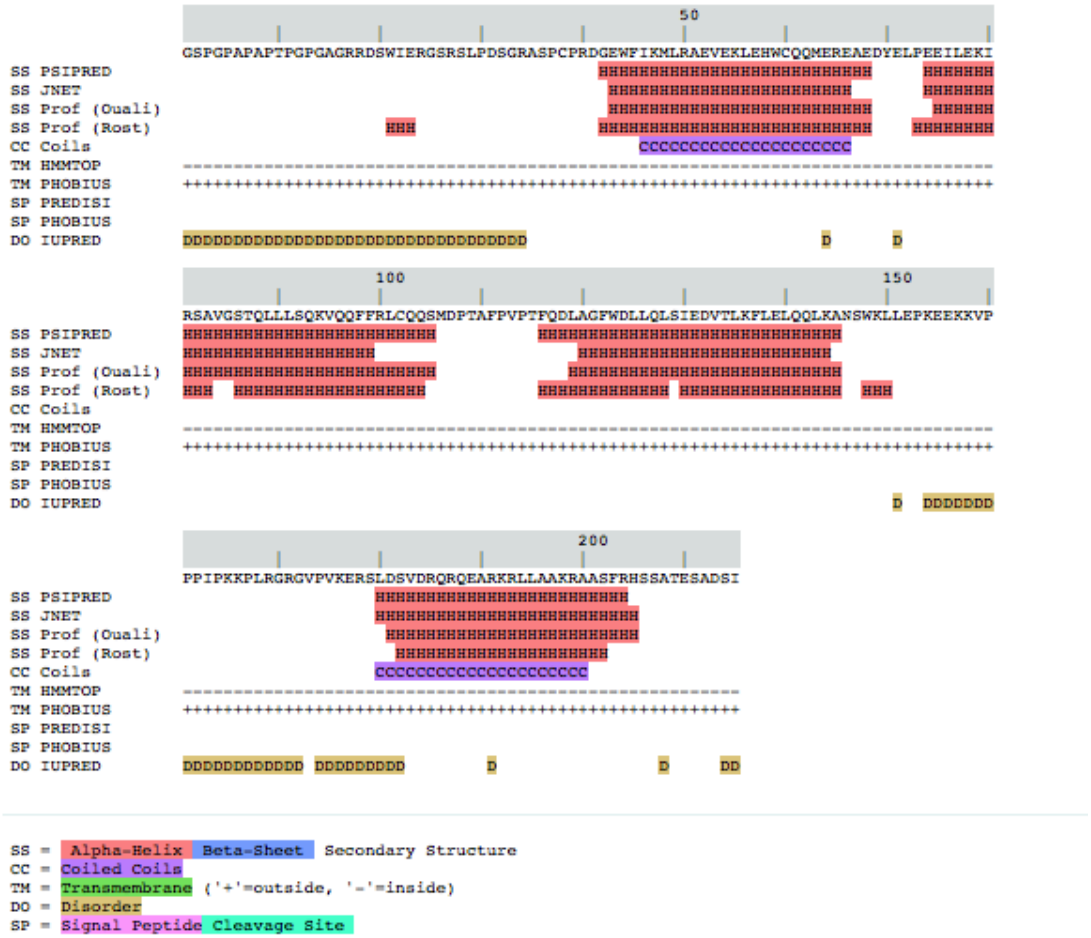


Figure 2.29 – Secondary structure prediction model for SAPAP3 domain 19. According to Quick2D prediction tool, the secondary structure comprising the sequence of domain 19 is mainly composed by alpha helix[156].

Together, these data showed that both the recombinant WT SAPAP domain 19 and its related mutants are monodisperse, stable and folded in accordance with both prediction analysis (Figure 2.29) and published data, and thus, had the appropriate production yields and quality attributes needed for the downstream studies.

Preliminary results of crystallization screenings for SAPAP3 domain 19 showed several putative crystals in two of the conditions tested (Figure 2.25 and 2.26), and attempts are now being carried out to confirm the nature of the putative crystals.

Despite we have only explored the C-terminus of SAPAP3 through the production of domain 19, the ESPRIT technology paved the way and provided the necessary tools to study other regions of SAPAP3. The recombinant production of domains already identified can be further improved by the refinement of their boundaries using strategies such as the use of limited proteolysis coupled to mass spectrometry[157, 158]; using different protein fusion proteins and

tags[159]; or even the inspection of different promoter systems, host strains, culture parameters[160, 161], cell lysis conditions[162] and the co-expression with counterparts[157, 158].

In summary, we have succeeded in identifying soluble domains of SAPAP3 covering a wide region of the protein sequence, and selected the best expressing soluble clone, corresponding to C-terminal region to proceed with its characterization. We have established a protocol for expression and purification of pure and folded recombinant WT SAPAP3 domain 19 and its mutants K910R and R770L, which were used, as will be described in the next chapter, for the interactomic characterization of SAPAP3 domain 19.

3. Interactome of WT SAPAP3 domain 19 and SAPAP3 mutants K910R and R770L

3.1 Introduction

Proteins do not function as single entities but rather as a dynamic network of team components that need to communicate in a space and time dependent way[163]. The cell can be viewed as a factory that comprises a complex network of interlocking assembly lines, each of which is composed by a set of large protein engines[164]. Nearly every cellular process occurs because proteins communicate with each other and exchange information[165]. The ability to interact between each other is highly dependent on their three-dimensional structure and on the presence of specific motifs[166]. This is a dynamic process, in which many interactions are transient, and others occur only in certain cellular contexts or at particular time in development[167]. The identification of protein interaction networks is essential since protein function is highly dependent on the interaction with nearest partners. In turn, protein-protein interactions (PPIs) are becoming increasingly attractive for their potential in providing new insights on disease phenotypes associated with genetic variations, including splice variants, allelic variants and point mutations[168, 169]. Furthermore, changes in PPIs associated with sequence variants contribute to the identification of novel drug targets and aids in the understanding of the disease molecular mechanisms.

Several strategies have been developed allowing for the high-throughput study of PPIs, and in general they comprise both genetic (like yeast two-hybrid, protein-fragment complementation assay, and mammalian protein–protein interaction trap) and biochemical approaches (such as affinity purification and protein microarrays). Genetic methods indirectly resolve PPIs based on the reconstitution of reporters upon interaction, whereas biochemical methods explore interactors by directly working with proteins in order to determine the composition of complexes[170].

Among the genetic assays, the classical yeast two-hybrid system (Figure 3.1-A) has become the most widely used method to assess protein-protein interactions[171]. In this assay, the proteins to be tested for interaction - the bait and the prey - are expressed as fusions with a DNA-binding domain (BD) and a transcriptional activating domain (AD); upon co-expression in yeast, interaction of the two hybrid proteins reconstitutes a transcription factor activity, which drives the expression of a reporter gene[172]. The main advantages of this method includes speed and simplicity, low cost and a broad application, ranging from the exploration of binary PPIs to large-scale interaction mapping[173]. However, it has also been associated with high rates of false-positives[173]. Additionally, forced subcellular localization of bait and prey proteins in the yeast nucleus may prohibit certain interactions, including interactions involving integral membrane proteins[174]. Moreover, proteins that require post-translational

modifications are unlikely to interact under Y2H conditions[175]. To date, several variants of the Y2H have emerged, most of them aiming to improve the throughput and to study PPIs in their natural environment.

Protein-fragment complementation assay is another genetic assay used for PPIs detection; it uses the direct physical interaction of two proteins to reconstitute a reporter protein from two genetically attached non-functional fragments (Figure 3.1-B)[176]. After reconstitution, the protein can activate a reporter that is dependent on refolding of two protein fragments and it does not require subsequent transcription or other amplification[177]. An advantage of this method is that it can be applied to specific organisms of interest[177].

By another way, the mammalian protein–protein interaction trap (MAPPIT) is based on reconstitution of a mammalian signal transduction pathway (the Janus kinase-signal transducer and activator of transcription cascade) that is initiated by interaction of the bait with the prey[171]. After the receptor (prey) has been stimulated with ligand (bait), their interaction drives a signaling cascade activation, resulting in STAT (signal transducers and activated of transcription)-dependent reporter gene activation (Figure 3.1-C)[171]. A clear advantage of MAPPIT is that it operates in intact mammalian cells; proteins that require post-translational modifications to exert their function can interact normally with their partners[176]. However, the scalability of MAPPIT is much lower than Y2H, because the complexity of handling and the cost of working with mammalian cells[176].

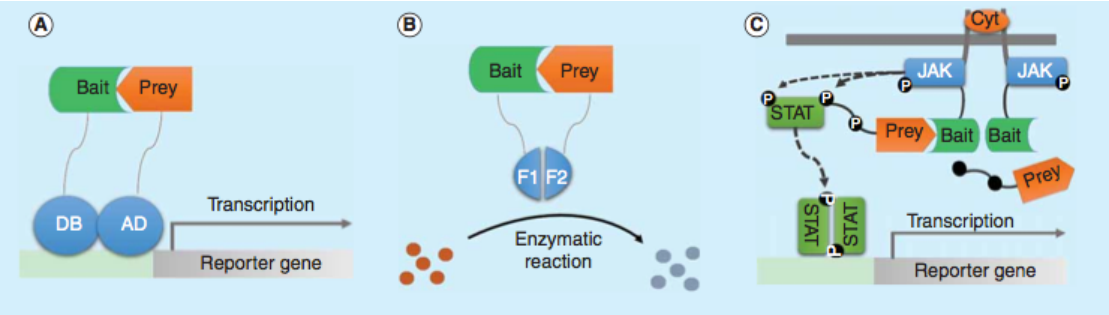


Figure 3.1 – Genetic methods for PPIs detection. (A) Yeast two hybrid: bait and prey are fused to the DNA-binding domain (DB) and transcriptional activating domain (AD) of an yeast transcription factor, respectively, and co-expressed in yeast cells. The interaction of bait and prey will form a transcription factor driving the expression of a reporter gene. (B) Protein-fragment complementation: the bait and prey proteins are fused to non-functional halves of a protein. Upon interaction of the bait and prey, the reconstitution of a reporter protein will perform enzymatic activity and transform the substrate (orange circles) to a detectable end product (gray circles). (C) Mammalian protein–protein interaction trap: bait proteins are coupled to a signaling-deficient cytokine receptor. Prey proteins are tethered to another receptor-moiety harboring STAT-recruitment sites. Upon interaction between bait and pray, a functional receptor is reconstituted and can be activated by a cytokine ligand. This ligand-binding leads to cross-phosphorylation of JAKs, which in turn phosphorylate the prey receptor-fragment, making the receptor accessible to STAT docking (P represents the phosphorylated sites). Recruited STATs are subsequently phosphorylated by JAKs that dimerize and enter the nucleus to activate reporter gene expression. Adapted from Feng *et al.*[176].

Several different biochemical strategies are also available to study the formation of protein complexes including, protein microarrays, Luminescence-based Mammalian IntERactome (LUMIER) and affinity chromatography.

Protein microarrays (Figure 3.2-A) is a biochemical technique, in which a library of complete or functional domains of recombinant proteins is printed onto a slide, and a tagged bait protein is then passed over the slide allowing it to specifically interact with a relevant bait protein, before detection by tag-specific immunoblotting[174]. It represents a powerful methodology with a wide range of research and clinical applications, in addition to the search of PPIs, it is useful for the identification and study of specific molecules and biomarkers, as well as in the analysis of post-translational modifications[171]. Although improvements in the technology, increasing its utility, accessibility and sensitivity, is still required[171].

LUMIER was developed for high-throughput PPI analysis in cultured mammalian cells and combines two-hybrid and biochemical strategies (Figure 3.2-B)[178]. In this approach, bait proteins are genetically fused to an affinity tag (usually FLAG) for purification, and prey proteins are attached to Renilla luciferase enzyme for detection[171]. The affinity tag insures purification of the protein complex of interest, and Renilla luciferase enzyme subsequently

reports the presence of the interaction partner[171]. It should however be noted that although overexpressing of proteins in LUMIER may be beneficial in assaying weak or transient interactions, it may also yield some false-positive results[171].

In affinity chromatography (AP), the bait protein contains an affinity tag (such as FLAG, Strep, Myc, hemagglutinin, protein A, His6- tag, calmodulin-binding protein, GFP and maltose-binding protein[179]) that is used to capture the bait protein from a cell lysate, capturing with it any interacting partners - the constituents of the resulting complex can then be identified by MS approaches[166]. To overcome the high false-positive rates of this approach, a derivative form of AP-MS called tandem affinity purification (TAP) has been developed[166]. In TAP, the target protein is fused with two sequential affinity tags - the protein A having high affinity for IgG and a calmodulin binding peptide having a high affinity for calmodulin - that are separated by a cleavage site of tobacco etch virus protease (Figure 3.2-C). In this approach two rounds of purification are performed: in the first the TAP-tagged protein is purified using a protein A affinity support followed by incubation with tobacco etch virus protease which releases the target protein; in the second purification, the protein complex is further purified via the calmodulin binding peptide of the TAP tag in the presence of calcium and released by use of calcium-chelating agents, such as ethylene glycol tetraacetic acid[166]. Finally, the isolated complex is identified and characterized by MS. The two rounds of purification enhance the specificity and sensitivity of this method, however, the increased purity and specificity create increased cost, and transient PPIs may be lost during the purification steps[176]. Additionally, identification of low abundance binding partners requires a relatively large amount of starting material[173].

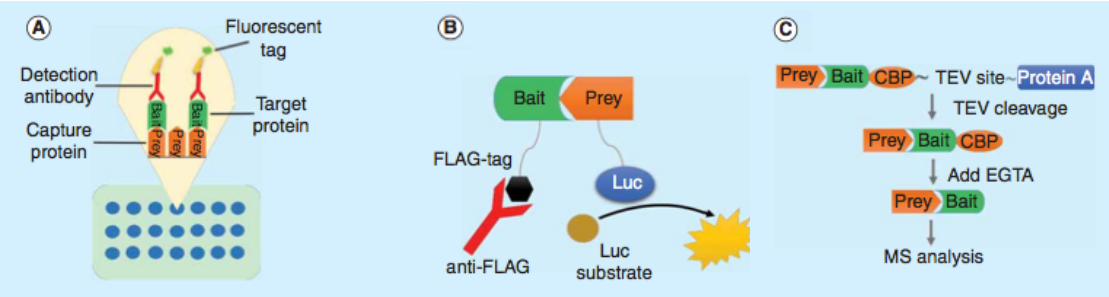


Figure 3.2 – Biochemical methods for PPIs detection. (A) Protein microarrays: prey proteins are immobilized on the array surface and incubated with bait; detection is carried out by means of a labeled secondary antibody. (B) LUMIER: the bait is fused to Renilla luciferase (Luc), the prey is tagged with FLAG (black), which are co-expressed in mammalian cells. Cell lysates are treated with anti-FLAG antibodies for immunoprecipitation. Subsequent light emission upon addition of the luciferase substrate associates with interaction between the pray and bait. (C) TAP-MS: a two-step purification is used: the first uses Protein A, which binds to IgG beads. The Protein A is subsequently removed by TEV protease. The second purification uses the affinity of a CBP for calmodulin, which is removed by EGTA, the interaction protein complexes are further identified by MS analysis. Adapted from Feng *et al.*[176].

Most of the affinity purification strategies depend on the isolation of material for subsequent MS-based protein identification. Sample preparation generally includes proteins denaturation, with a subsequent proteolytic digestion to generate small peptides. The implemented platform for AP-MS analysis is made by liquid chromatography (LC)-MS/MS, where the peptides are separated with an organic gradient that elutes the sample directly into the mass spectrometer coupled with a nano-electrospray source for ionization. Data are generally acquired by standardized methods for data-dependent acquisition (DDA) that collect the maximum number of informative mass spectra within a short time[179]. Matching raw MS data to peptide sequences can be performed using a number of proteomics software packages that include suitable scoring for ranking the probability of correct peptide identification and protein inference[179].

While characterization of complex interaction networks requires only protein identification, when one needs to obtain a dynamic view of the interactome, quantitative analysis is required. Mass spectrometry has allowed the development of highly multiplex quantification assays commonly divided in label-based and label-free methods. Stable isotope labelling methods include: stable isotope labelling by amino acids in cell culture (SILAC), isobaric tags for relative and absolute quantification (iTRAQ) and isotope coded affinity tags (ICAT). To differentially label proteins with stable isotopes and compare their relative abundance metabolic or chemical labelling can be used. In metabolic labelling, stable isotopes are incorporated by introducing amino acids containing stable isotopes in the growth media of

living cells, as is the case of SILAC method. In chemical modification-based approaches, like iTRAQ and ICAT, heavy and light chemical agents are introduced at the protein/peptide level. Prior the MS analysis, the samples from different experimental conditions are mixed and the relative quantification can be performed based on the expected shift in mass induced by the incorporated tags[165, 173].

Alternatively, label-free quantification methods have attracted interest based on recent improvements in high-throughput and automation of LC-MS instruments and the development of novel algorithms dealing with LC-MS data[180]. Label-free techniques measure protein abundance based on (i) the analysis of integrated chromatography peak areas of peptides, or (ii) the number of MS/MS spectra (spectral counting) in a LC-MS/MS run[180]. Spectral counting relies on the observation that more abundant proteins will be selected more often for fragmentation in a data dependent acquisition experiment and will thus produce more MS/MS spectra[180]. To provide reliable information, counting methods require an instrument that has the ability to sample a LC peak at a high frequency, therefore the faster is the instrument, the higher quality is the spectral counting data[181]. However, with this method the identification dynamic range is limited, and low-level ions may be missed in preference to higher intensity ions, limiting the application of spectral counting to moderately to highly abundant proteins, or to proteins whose abundance varies significantly between the samples[181]. Thus, it is not always appropriate for many dynamically-regulated interactions[181]. In addition to the spectral counting-based label-free quantification methods, the chromatographic peak area or ion intensity (ion count) for a given peptide at a specific retention time in an LC-MS run can be used to quantify proteins in a sample because this measure is linearly correlated to the concentration of that peptide in the sample. This method strongly relies on a good LC resolution and reproducibility, and should be performed on MS instruments with a high mass accuracy and resolution for the correct assignment of m/z values[180]. In this sense, sequential window acquisition of all theoretical fragment ion spectra (SWATH) has been developed[181], where a fragment ion library is first generated via data-dependent acquisition, which, along with the retention time of the peaks in the liquid chromatography, acts as a reference for the target extraction of fragment ions that are subsequently generated by SWATH[182]. In the target extraction, all precursor ions that continually enter the mass spectrometer are theoretically fragmented and measured in consecutive acquisition windows of defined mass width (SWATH) that progress over the full m/z range and retention time[181] (Figure 3.3). Thus, in SWATH, a combinatorial approach of data dependent acquisition with targeted extraction of fragment ions is used to quantify

proteins.

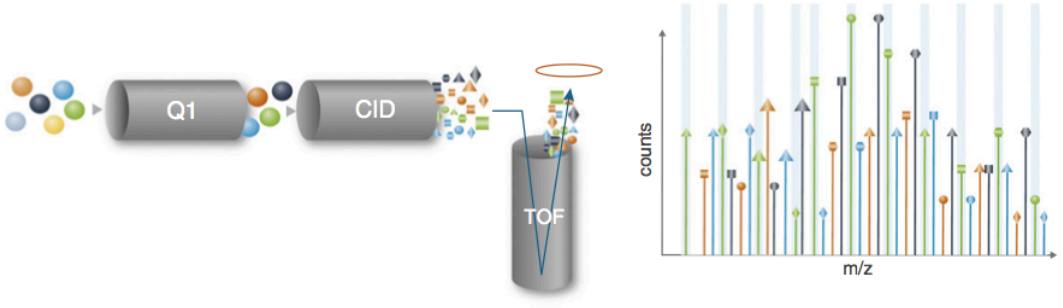


Figure 3.3 – Schematic representation of SWATH. The first mass selection quadrupole (Q1) isolates and transmits ions in a defined window that contain a number of different parent ions. The ions are then fragmented in the collision cell (CID) by collision-induced fragmentation and transmitted to the TOF analyser where all product ions are measured. The isolation window is stepped through the entire mass range of interest: the number of windows, their width, the mass range to be covered and also the accumulation time all contribute to the total cycle time. Adapted from Tate *et al.*[181].

In this chapter it will be presented how affinity purification combined with SWATH-MS method allowed the quantification of changes in protein-protein interactions between WT SAPAP3 domain 19 and SAPAP3 disease-associated mutations K910R and R770L. With this approach we expect to provide novel insights on the protein interaction network of SAPAP3 (specifically on its C-terminal domain), and if the disease associated mutations affect these interactions. By making use of this high-throughput proteomic strategy we hope to contribute to novel insights into the pathological causes of OCD and schizophrenia.

3.2 Materials and Methods

3.2.1 Isolation of synaptoneurosomes

Whole striatum from four 18 ± 1 day pregnant adult female Wistar rats were dissected and processed separately. After dissection the striatum tissues were put in 8 mL of purification medium (10 mM HEPES-Tris pH 7.4, 0.32 M sucrose and 0.1 M EGTA) and homogenized using a loose-fitting pestle. After this step, 7 mL of purification medium was added to the samples and homogenized again using a tight-fitting pestle. The samples were centrifuged at 800 *g* for 3 minutes at 4 °C for erythrocytes sedimentation. The supernatants were sequentially filtered through membranes of 150 µm, 50 µm and 8 µm, and then centrifuged at 10 000 *g* for 15 minutes at 4 °C for synaptoneurosome sedimentation. The pellets were resuspended in 1 mL of purification buffer containing a cocktail of proteases and phosphatase inhibitors and stored at – 80 °C until used.

3.2.2 Affinity purification assay

Synaptoneurosomes lysis

Frozen synaptoneurosomes were thawed on ice and centrifuged at 10 000 *g* for 15 minutes at 4 °C for synaptoneurosome sedimentation. The pellets were resuspended in 1 mL of lysis buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 % Triton X-100; 0.5 mM DTT, 5 mM $MgCl_2$ containing a cocktail of proteases (cOmplete, EDTA-free, Roche Diagnostics) and phosphatase (PhosStop, Roche Diagnostics) inhibitors and kept on ice for 15 minutes. The samples were centrifuged at 10 000 *g* for 15 minutes at 4 °C and the supernatants collected. Protein concentration was determined using the BCA protein determination assay (Thermo Fisher Scientific) according to the manufacturer's instructions.

SAPAP3 domain 19 immobilization in magnetic beads

His Mag Sepharose Ni^{2+} beads (GE Healthcare) were mixed thoroughly by vortexing and 100 µL of the homogeneous slurry were dispensed for each experimental condition in a 1.5 mL Eppendorf tube. The tubes were placed in a magnetic rack and the storage solution removed. The beads were then resuspended in 500 µL of equilibration buffer (20 mM sodium phosphate pH 7.4, 150 mM NaCl and 10 % glycerol), and 500 µL of the recombinant SAPAP3 domain 19 WT, K910R or R770L (0.04 µg/µL) were separately incubated with magnetic beads for 30 minutes using a benchtop shaker. The tubes were then placed in a magnetic rack and

the supernatants removed. The protein loaded magnetic beads were washed for three times with 500 µL of equilibration buffer to remove unbound material.

SAPAP3 domain 19 incubation with synaptoneurosome extracts

The magnetic beads with immobilized protein (WT SAPAP3 domain 19, K910 and R770L mutants) were incubated with synaptoneurosome extracts (180 µg/condition) during 3 hours at 4 °C, using naked magnetic beads as negative control. After that, samples were washed three times with wash buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 % Triton X-100; 0.5 mM DTT, 5 mM $MgCl_2$) to remove nonspecific interacting proteins. GFP-MBP at 1.8 µg/µL was added to each sample to act as internal standard in the mass spectrometry analysis. Samples were denatured with 30 µL of 2x Laemmli sample buffer (65.8 mM Tris-HCl pH 6.8, 2.1 % SDS, 26.3 % (w/v) glycerol, 0.01 % bromophenol blue) at 95 °C for 15 minutes.

3.2.3 Gel separation, gel band processing and peptide extraction (Short-GeLC)

Eight microliters of each denatured sample were pooled to generate the proteome library (protein composition of each condition), and 22 µL of each sample were used separately for SWATH-MS analysis to extract quantitative information. Samples were then alkylated with acrylamide and subjected to in-gel digestion by using the short-GeLC approach[183]. Briefly, the entire sample was loaded in a "4–20 % TGX Stain-Free Gel" (Bio-Rad) and subjected to a partially electrophoretic separation: 15 min at 110 V to allow the samples to enter into the gel. After SDS-PAGE, proteins were visualized with Colloidal Coomassie Blue G-250 (Bio-Rad). For this, the gels were washed with distilled water and immersed in staining solution (10 % (v/v) of 85 % solution of phosphoric acid, 10 % (w/v) ammonium sulphate and 20 % methanol) to which Colloidal Coomassie Blue G-250 powder was added with a filter to allow the formation of colloidal particles under agitation. After staining, each gel lane was sliced into 9 bands of equal size, and bands from each lane were grouped three by three, sliced into smaller pieces and destained using the destaining solution (50 mM ammonium bicarbonate and 30 % acetonitrile) following by a washing step with water, the gel pieces were then dehydrated on a Concentrador Plus/Vacufuge Plus (Eppendorf). Seventy microliters of trypsin (0.01 µg/µL solution in 10 mM ammonium bicarbonate) were added to the dried gel bands and left for 15 min, on ice, to rehydrate the gel pieces, followed by the addition of 30 µL of 10 mM ammonium bicarbonate and in-gel digestion performed overnight at room temperature in the dark.

After the digestion, the excess solution from gel pieces was collected to a low binding microcentrifuge tube (LoBind, Eppendorf) and peptides were extracted from the gel pieces by sequential addition of three solutions of acetonitrile (ACN) in 1 % formic acid (FA) (30 %, 50 %, and 98 % of ACN, respectively). After the addition of each solution, the tubes were shaken in the thermomixer (Eppendorf) at 1050 rpm for 15 min and the solution collected to the tube containing the previous fraction. The peptide mixtures were dried by rotary evaporation under vacuum (Concentrador Plus/Vacufuge Plus (Eppendorf)). The extracted peptides resulting from each band processing were combined and concentrated for quantitative analysis by acquisition in SWATH mode. Before performing the LC-MS/MS analysis the peptide mixture samples were subjected to solid phase extraction using OMIX tips with C18 stationary phase (Agilent Technologies) as recommended by the manufacturer. Eluates were dried by rotator evaporation, avoiding total evaporation.

Peptides mixtures were resuspended to 30 µL in a solution of 2 % ACN and 0.1 % FA followed by vortex, spin and sonication in water bath (2 min with pulses of 1 second – 1 s sonication followed by 1 s break pulse, at 20 % intensity, in a Sonics VibraCell 750 watts sonicator (Sonics&Materials). In order to remove insoluble material, the peptide mixtures were finally centrifuged for 5 min at 14 000 *g* and collected into the proper vials for LC-MS/MS analysis.

3.2.4 SWATH Acquisition

Samples were analysed on a Triple TOF^{TM} 5600 System (ABSciex) in two phases: data-dependent acquisition was followed by SWATH (Sequential Windowed data independent Acquisition of the Total High-resolution Mass Spectra) acquisition on each individual samples. Peptides were resolved by liquid chromatography (nanoLC Ultra 2D, Eksigent) on a ChromXP^{TM} C18 CL reverse phase column (300 µm ID × 15 cm length, 3 µm particles, 120 Å pore size, Eksigent) at 5 µL/min. Peptides were eluted into the mass spectrometer with an acetonitrile linear gradient in 0.1 % FA (5 % to 35 % ACN for 45 min), using an electrospray ionization source (DuoSpray^{TM} Source, ABSciex). For DDA experiments the mass spectrometer was set to scanning full spectra (350-1250 m/z) for 250 ms, followed by up to 100 MS/MS scans (100–1500 m/z from a dynamic accumulation time – minimum 30 ms for precursor above the intensity threshold of 1000 – in order to maintain a cycle time of 3.3 s). Candidate ions with a charge state between +2 and +5 and counts above a minimum threshold of 10 counts per second were isolated for fragmentation and one MS/MS spectra

was collected before adding those ions to the exclusion list for 25 seconds (mass spectrometer operated by Analyst® TF 1.6, ABSciex®). Rolling collision was used with a collision energy spread of 5.

3.2.5 Spectral library construction and SWATH data file processing

Peptide identification was performed with Protein Pilot software (v5.0, ABSciex®). Search parameters used were the following: SwissProt database, against a database composed by *Rattus Norvegicus* (release at January 2015), SAPAP3, GFP and MBP sequences, and using acrylamide alkylated cysteines as fixed modification. An independent False Discovery Rate (FDR) analysis using the target-decoy approach provided with Protein Pilot software was used to assess the quality of the identifications and positive identifications were considered when identified proteins and peptides reached a 5 % local FDR[184, 185].

To improve the sample coverage, an additional DDA experiment was done for each pool using an exclusion list of the peptides previously identified. For SWATH-MS based experiments, the mass spectrometer was operated in a looped product ion mode[186] and the same chromatographic conditions used as in the DDA run described above. The SWATH-MS setup was designed specifically for the samples to be analysed using the SWATH Variable Window Calculator_V1.0 (ABSciex) (Supplementary Table 6.1), to perform that a pool of all samples was analysed in DDA mode to adapt the SWATH windows. A set of 60 windows of variable width (containing 1 m/z for the window overlap) was constructed covering the precursor mass range of 350-1250 m/z. A 100 ms survey scan (350-1500 m/z) was acquired at the beginning of each cycle for instrument calibration and SWATH MS/MS spectra were collected from 100–1500 m/z for 50 ms resulting in a cycle time of 3.15 s from the precursors ranging from 350 to 1250 m/z. The collision energy for each window was determined according to the calculation for a charge +2 ion centred upon the window with a collision energy spread of 15. A specific library of precursor masses and fragment ions was created by combining all files from the DDA experiments, and used for subsequent SWATH processing.

Libraries were obtained using Protein Pilot™ software (v5.0, ABSciex) with the same parameters as described above. Data processing was performed using SWATH™ processing plug-in for PeakView™ (v2.0.01, ABSciex); briefly, peptides were selected automatically from the library using the following criteria: (i) the unique peptides for a specific targeted protein were ranked by the intensity of the precursor ion from the DDA analysis as estimated by the

ProteinPilot™ software, and (ii) peptides that contained biological modifications and/or were shared between different protein entries/isoforms were excluded from selection. Up to 15 peptides were chosen per protein, and SWATH™ quantitation was attempted for all proteins in library file that were identified below 5 % local FDR from ProteinPilot™ searches. In SWATH™ acquisition data, peptides were confirmed by finding and scoring peak groups that were a set of fragment ions for that peptide. Target fragment ions, up to 5, were automatically selected and peak groups were scored following the criteria described in Lambert *et al*.[168]. Peak group confidence threshold was determined based on a FDR analysis using the target-decoy approach and 1 % extraction FDR threshold was used for all the analyses. Peptide that met the 1 % FDR threshold in at least three biological replicates were retained, and the peak areas of the target fragment ions of those peptides were extracted across the experiments using an extracted-ion chromatogram (XIC) window of 3 minutes. The levels of the proteins were estimated by summing all the transitions from all the peptides for a given protein (an adaptation from Collins *et al.*[187] and normalized to the internal standard (GFP+MBP).

3.2.6 Data analysis

To quantify the interactomic changes between WT SAPAP3 domain 19 and K910R and R770L mutants, analysis of the relative protein interactors' quantification was processed using the Perseus 1.5.1.6 Software (Max Planck Institute of Biochemistry). After data $\log_2$ normalization, Student's t-test was used to test the differences between groups, and a false discovery rate (FDR or q-value) < 0.05 was considered to be significant. Gene ontology analyses were performed with GeneMANIA and Gorilla tools.

3 Results

3.3.1 Differential interactome between WT SAPAP3 domain 19 and SAPAP3 disease-associated mutants K910R and R770L

To characterize novel interactor partners of SAPAP3 domain 19, and to quantify changes in protein-protein interactions associated with point mutations K910R (obsessive-compulsive disorder) and R770L (schizophrenia), an interactomic strategy was implemented by using affinity-purification combined with SWATH-MS analysis (Figure 3.4).

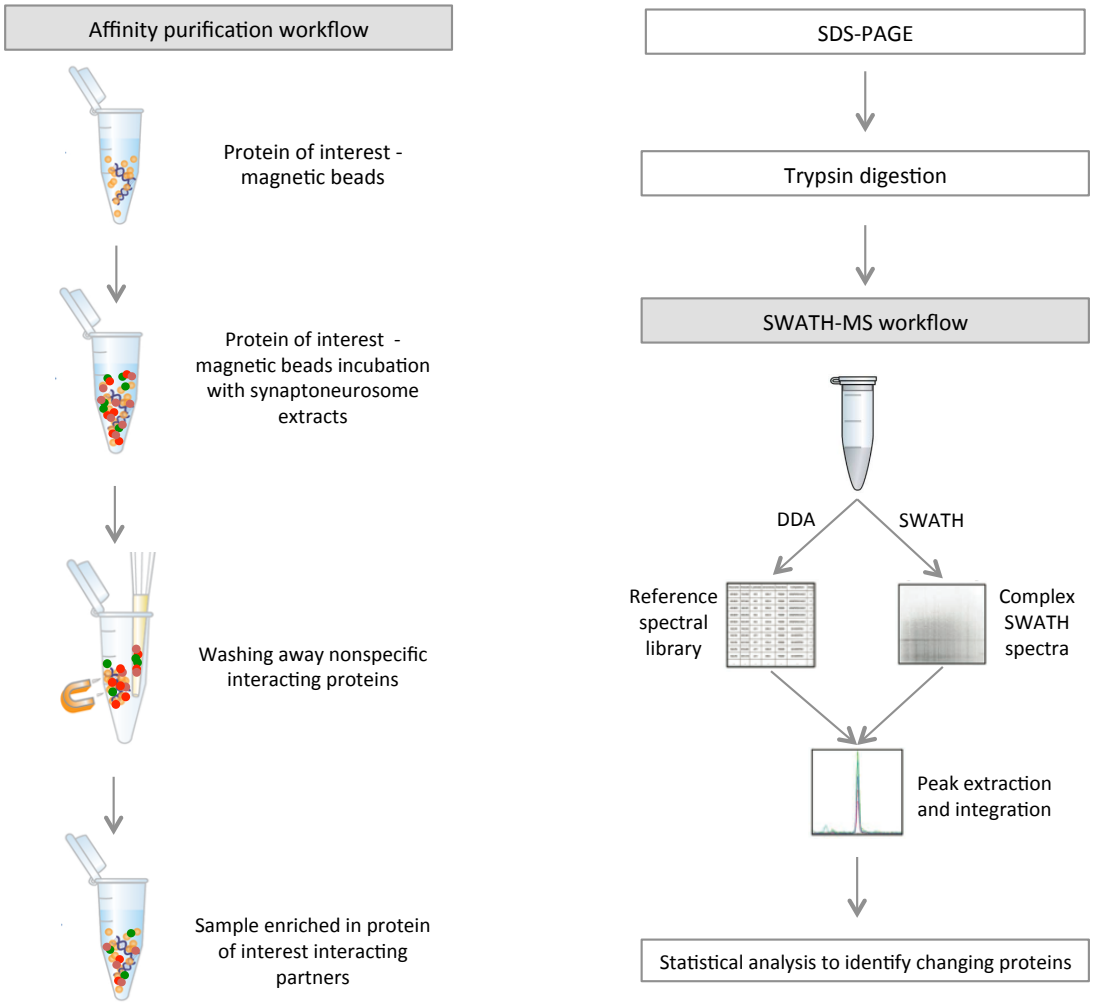


Figure 3.4 – AP-SWATH workflow. In the interactomic workflow established, first, the affinity purification involved the immobilization of histidine tagged protein (WT SAPAP3 domain 19 or the K910R and R770L mutants) onto hexahistidine affinity magnetic beads; the protein coated beads were then incubated with extracts of synaptoneurosomes and the non-specific interacting proteins washed away; naked beads were used as negative control. Each sample was denaturated and partially separated by SDS-PAGE. The proteins were trypsin digested, extracted from gel and analysed by SWATH-MS. Each condition was injected separately in a DDA mode to build the spectral library and in a SWATH mode. The individual SWATH runs were then matched against the spectral library, and specific peptide ions were extracted and integrated for quantification. Finally, the extracted data was statistically analyzed to identify changing proteins.

For the affinity purification, the recombinant WT SAPAP3 domain 19 and the K910R and R770L mutants were immobilized on Ni-NTA magnetic nickel beads and incubated with extracts of striatum synaptoneurosomes in quadruplicate (Figure 3.4).

Synaptoneurosomes (Figure 3.5) are structures enriched in released presynaptic structures (synaptosomes) with attached sealed postsynaptic entities (neurosomes)[188, 189], retaining the molecular machinery used in neuronal signalling[190], and where SAPAP3 is naturally localized. Thus, striatum synaptoneurosome extracts represents the ideal source of protein for isolation of SAPAP3 interacting partners.

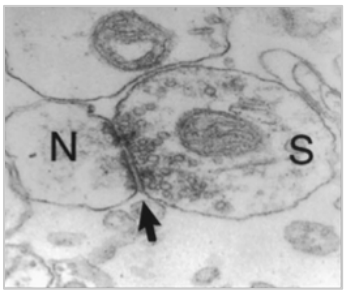


Figure 3.5 – Electron micrograph of synaptoneurosome preparation. Closed-up view of electron micrograph of a typical synaptoneurosome revealing its presynaptic (synaptosome, S) and postsynaptic (neurosome, N) components (20000 x). Adapted from Johnson *et al.*[191].

After synaptoneurosome extracts incubation, a washing step was performed to remove the non-specific binding partners. The recombinant WT SAPAP3 domain 19, and its mutants, binding with the interactor partners were then recovered; partially separated by electrophoresis, followed by in gel trypsin digested, and analysis by SWATH-MS. The SWATH-MS workflow involved several steps as shown in Figure 3.4. Each sample was injected separately in a data-dependent acquisition mode and in SWATH mode. In DDA mode, a spectral library was generated, which was required in order to identify the peptides acquired in SWATH runs. The SWATH acquisition mode involved sampling the m/z 350-1250 range in 60 steps at widths of 25 m/z to ensure complete sample coverage. The individual SWATH runs were then matched against the spectral library, and specific peptide ions were extracted to enable quantification of the area under the curve between samples. Finally, the extracted data was analyzed, allowing not only the identification of SAPAP3 domain 19 putative interactors, but also to trace changing interaction profiles under the mutant conditions.

In this project, SWATH strategy was used to monitoring the interaction changes between the WT SAPAP3 domain 19 and two mutants, one associated with OCD (K910R) and other with schizophrenia (R770L). This strategy allowed the identification of 787 proteins (Supplementary Table 6.2), from which 502 proteins were quantified based on a confidence threshold determined on a FDR analysis, where peptides within 1 % FDR threshold (in at least three of the four biological replicates) were retained. A complete quantitative data matrix of the 502 proteins identified was obtained for 4 biological replicates for each condition with SWATH, represented as a heat map in Figure 3.6.

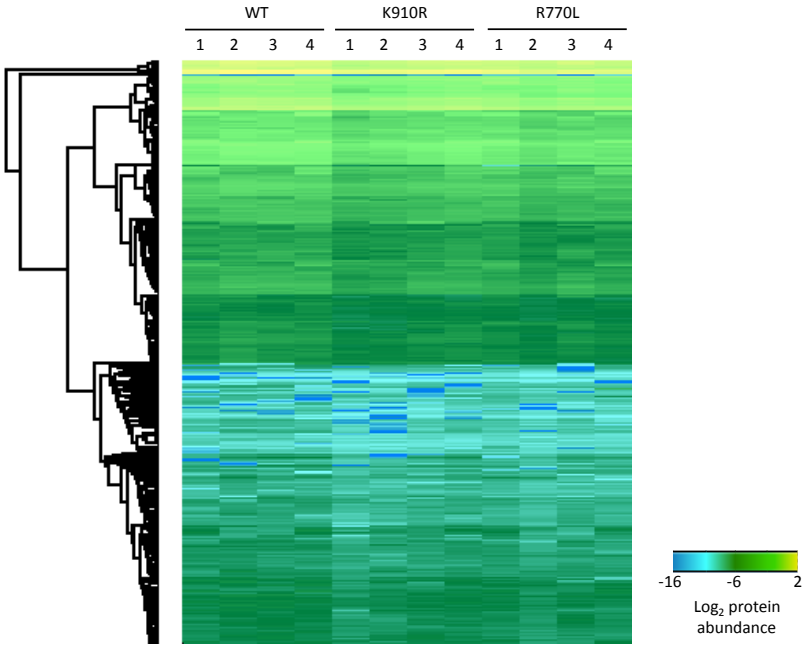


Figure 3.6 – Quantitative data matrix from SWATH results. Quantitative data for 502 proteins obtained by SWATH analysis of 12 samples including four SAPAP3 domain 19 WT (WT), four SAPAP3 domain 19 K910R mutant (K910R) and four SAPAP3 domain 19 R770L mutant (R770L). The proteins were clustered in the vertical direction hierarchically using Euclidian distance. This data matrix constituted the basis for all downstream quantitative comparisons. Blue represents low levels of protein, green intermediate levels, and yellow high levels of protein.

Since we were interested in monitoring the interactome changes in pairs of samples, we used a Student's t-test, considering a false discovery rate (FDR or q-value) < 0.05 after $\log_2$ normalization, to determine the statistical significance between the data from WT SAPAP3 domain 19 vs. mutant K910R and from WT SAPAP3 domain 19 vs. mutant R770L groups.

The comparison between the data from WT SAPAP3 domain 19 vs. mutant K910R mutant has showed significant differences in the intensity of 43 proteins (Figure 3.7-A and Supplementary Table 6.3). In turn, the analysis revealed that the protein amounts for 124

proteins intensities were significantly and statistically different between WT SAPAP3 domain 19 vs. mutant R770L mutant (Figure 3.7-B and Supplementary Table 6.4).

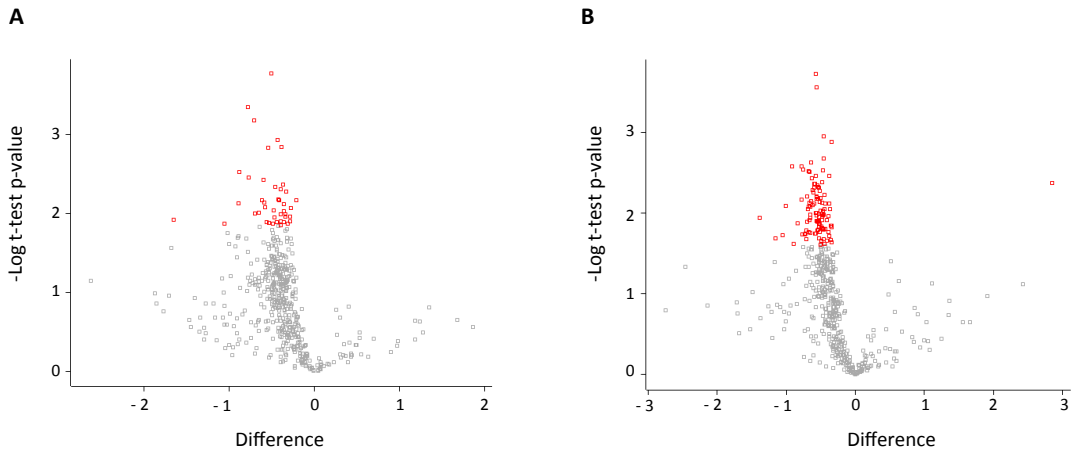


Figure 3.7 – Volcano plot of putative protein interactors found using AP-SWATH-MS. The volcano plot displays the relationship between fold-change and significance between (A) SAPAP3 domain 19 WT vs. mutant K910R, and (B) SAPAP3 domain 19 WT vs. mutant R770L. The y-axis is the negative log2 of p-values (a higher value indicates greater significance) and the x-axis is the difference in intensities between the conditions under study. Interactors identified as significant are labelled in red (FDR t-test < 0.05).

Comparing the identity of those proteins in both analysis it was observed that 24 proteins were shared between the two groups (Figure 3.8).

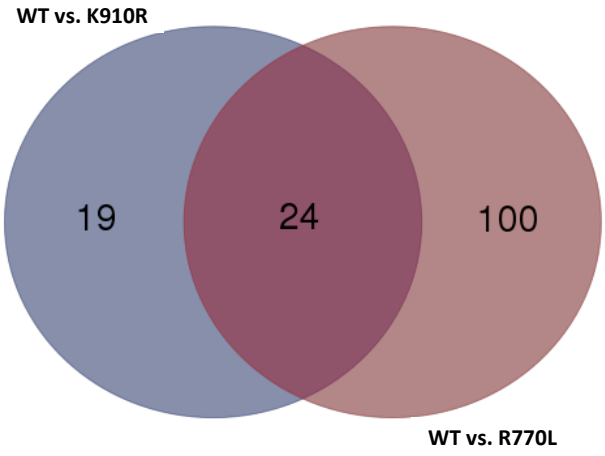


Figure 3.8 – Venn diagram illustrating the number of shared and unique significantly different binding partners. A total of 43 and 124 proteins were found to the significantly different between SAPAP3 WT vs. K910R and WT vs. R770L, respectively. There are 24 common proteins between the two groups.

To associate multiple proteins to a specific experimental condition, principal component analysis (PCA) was used. PCA of the differential interactomes found with Student's t-test for WT vs. K910R and WT vs. R770L revealed a clear separation of the groups as can be visualized by the scores plots of Figure 3.9.

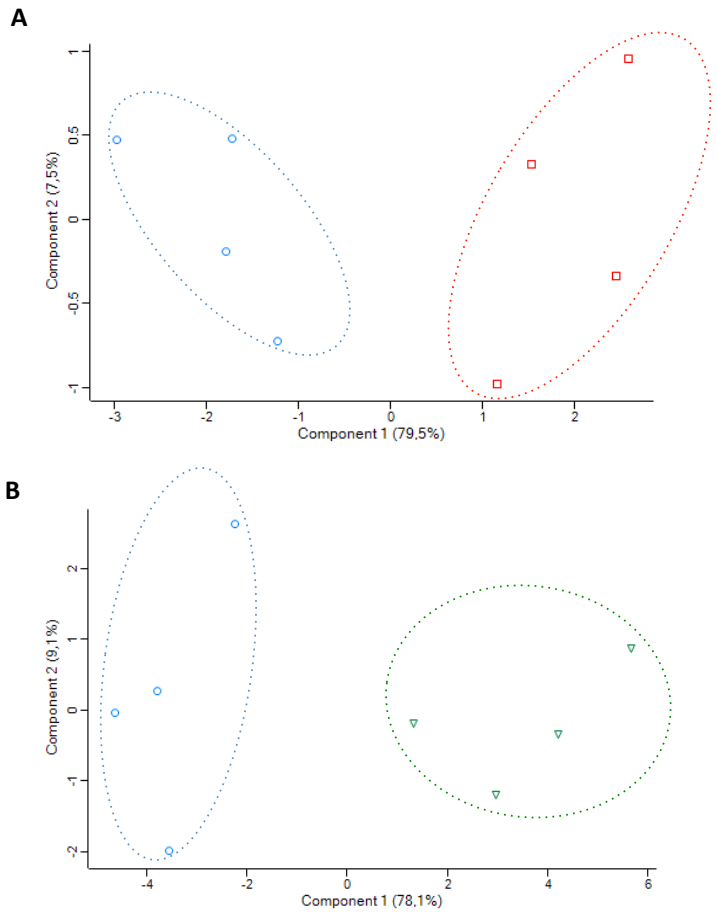


Figure 3.9 – Principal component analysis of differential interactomes. PCA was performed using in (A) the 43 significant different variables found for WT vs. K910R; and in (B) the 124 statistically different variables found for WT vs. R770L. The scores plot for each case shows the formation of two main groups along the first component (component 1). Blue: WT SAPAP3; red: K910R mutant and green: R770L mutant.

In order to check if the proteins whose binding was affected by each mutation were already known to interact with each other and/or with SAPAP3, we analysed the two protein lists by GeneMANIA Cytoscape app[192], a tool that allows the construction of a composite gene-gene functional interaction network from a gene list, which can include the genes most related to the original list, and functional annotations from Gene Ontology[192]. GeneMANIA searches many large, publicly available biological datasets of organism-specific weighted networks to construct the resulting composite network, containing over 500 million interactions for 8

organisms: *A. thaliana*, *C. elegans*, *D. melanogaster*, *D. rerio*, *H. sapiens*, *M. musculus*, *R. Norvegicus*, and *S. cerevisiae*. The searched databases are collections of annotated interaction networks, organized into groups such as co-expression, genetic interaction, pathway, predicted, co-localization, physical interaction and shared protein domains, and the user can select any combination of these as the basis of the composite network.

In order to investigate the known protein-protein interactions and to investigate the enriched functional annotations, the 43 proteins found to be significantly changed for K910R mutant when compared to the WT form, and the 124 proteins significantly altered in R770L mutant vs. WT were also analysed by GeneMANIA tool, using the *H. sapiens* database, displaying known, validated physical interactions only, and without enrichment of known closely related interaction partners not present in the gene list query. For SAPAP3 domain 19 vs. K910R group, 16 genes out of the 43 listed showed 16 known physical interactions, as can be viewed in the network of Figure 3.10.

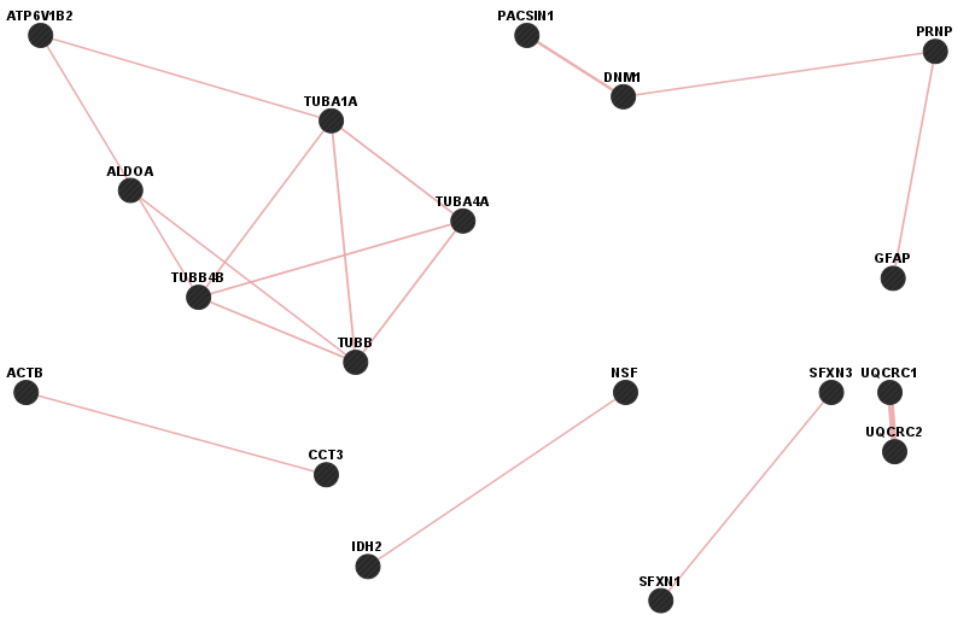


Figure 3.10 – Network view of known physical interactions for the statistically significant proteins found for WT SAPAP3 domain 19 vs. K910R. The gene list was analysed through GeneMANIA to search for known physical interactions between the putative interactors. The black circles (nodes) represent genes and the pink lines (edges) the interaction between the genes.

The results showed the presence of an intricate network of interactions among 6 proteins, a linear interaction network between 4 proteins and 4 binary interactions, which might indicate that some of these networks could possibly translate into the formation of protein

complexes, and that, due to the nature of the research approach, SAPAP3 domain 19 might be interacting with only a fraction of the proteins detected.

The GeneMANIA app also displays Gene Ontology (GO) terms enriched among the genes in the network, which could be particularly helpful to interpret the biological meaning of the protein interactions affected by K910R mutation. The GeneMANIA algorithm presents the GO categories and respective Q-values from a FDR corrected hypergeometric test for enrichment, along with coverage ratios for the number of annotated genes in the displayed network vs. the number of genes with that annotation in the genome. For the list of proteins affected by K910R mutation, the most enriched GO terms identified through GeneMANIA algorithm are listed in Table 3.1.

For this group of proteins, although the GO association confidence given by the quality score (Q-value), was not very significant due to the low number of proteins in the query network, the results point towards the OCD related mutation predominantly influencing changes in the interaction of SAPAP3 with proteins or protein complexes involved in cellular respiration, protein folding and several other mitochondrial processes.

Table 3.1 – GeneMANIA functional enrichment analysis of the statistically significant proteins found for WT SAPAP3 domain 19 vs. K910R.

GO annotation	Q-value	Occurrence
Cellular respiration	3.75E-7	9 (140)
Respiratory electron transport chain	4.28E-7	8 (103)
Electron transport chain	4.28E-7	8 (104)
'de novo' posttranslational protein folding	1.06E-6	6 (38)
Energy derivation by oxidation of organic compounds	1.67E-6	10 (285)
'de novo' protein folding	1.67E-6	6 (43)
Oxidative phosphorylation	1.30E-5	6 (62)
Protein folding	2.42E-5	7 (124)
Mitochondrial respiratory chain	4.22E-4	5 (59)
Respiratory chain	4.88E-4	5 (62)
Mitochondrial inner membrane	6.17E-4	7 (208)
Mitochondrial membrane part	6.91E-4	6 (131)
Organelle inner membrane	7.86E-4	7 (221)
Microtubule	1.26E-3	6 (149)
Mitochondrial electron transport, NADH to ubiquinone	2.40E-3	4 (41)
NADH dehydrogenase activity	2.40E-3	4 (42)
NADH dehydrogenase (quinone) activity	2.40E-3	4 (42)
NADH dehydrogenase (ubiquinone) activity	2.40E-3	4 (42)
Respiratory chain complex I	2.72E-3	4 (45)
NADH dehydrogenase complex	2.72E-3	4 (45)

By evidencing each of these annotations in the network (Figure 3.11), we can see in yellow the genes involved in protein folding, and in blue genes involved in cellular respiration; the remaining genes presented in the figure did not belong to any other overrepresented GO annotation with the exception of the TUBB gene, which fit in the microtubule annotation together with TUBA1A, TUBA4A, TUBB4B and CCT3 genes.

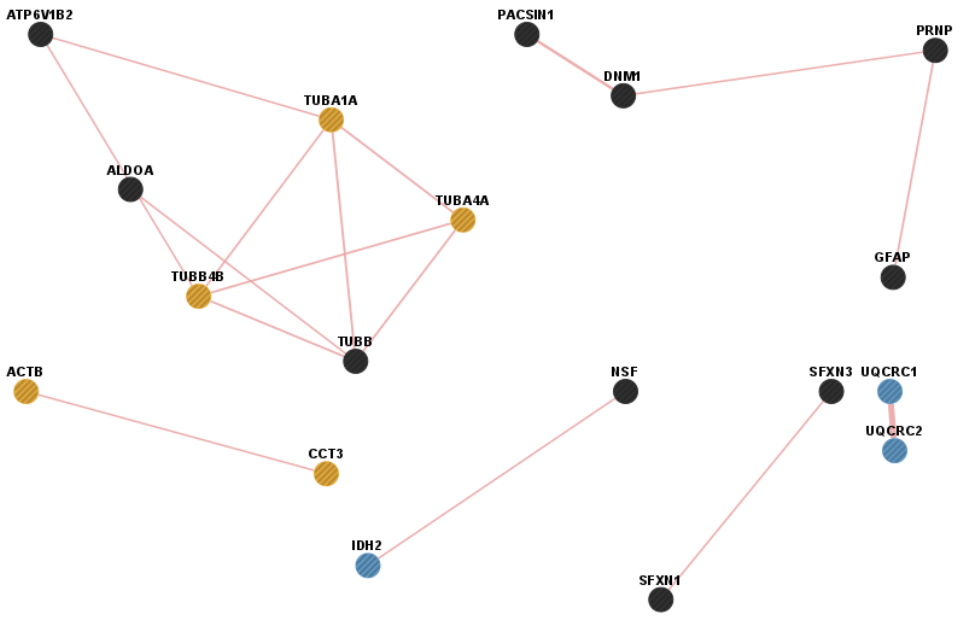


Figure 3.11 – Network view of known physical interactions for the statistically significant proteins found for WT SAPAP3 domain 19 vs. K910R with GO annotations. The yellow circles represent genes involved in protein folding, and the blue circles genes involved in cellular respiration.

This low number of interacting partners varying from the WT sequence to the K910R was somehow expected, as we were addressing a conservative mutation, mutating a positively charged residue by another positively charged residue.

When considering the 124 proteins from the WT SAPAP3 domain 19 vs. R770L group, GeneMANIA results showed that out of the 124 proteins submitted, 92 proteins present 192 known physical interactions, with an intricate network of physical interactions (Figure 3.12).

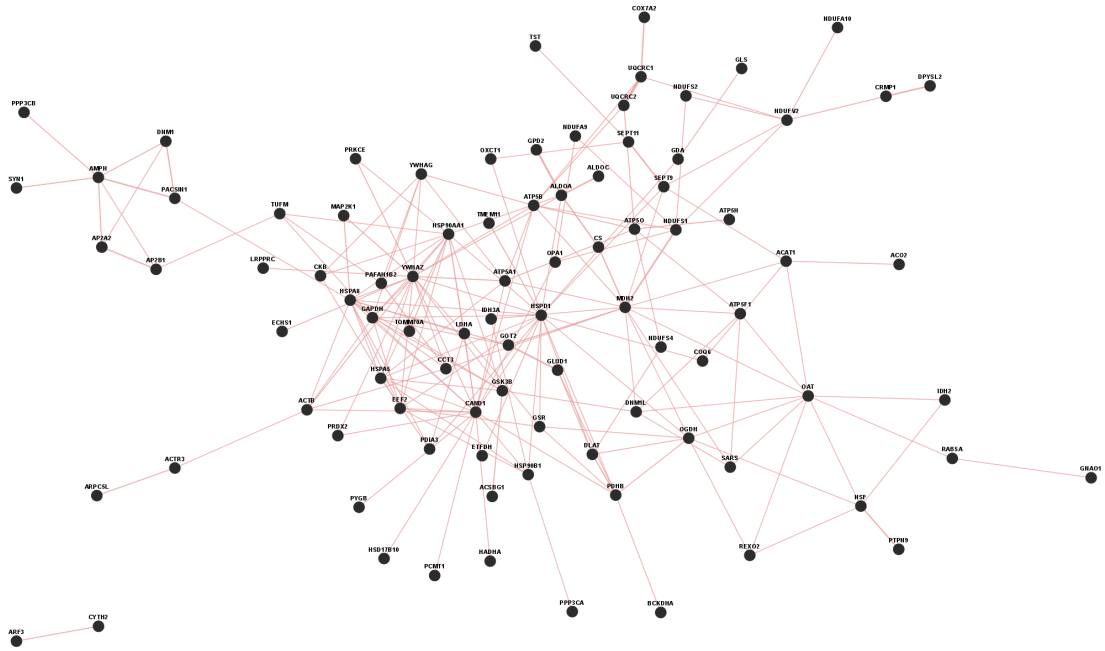


Figure 3.12 – Network view of known physical interactions for the statistically significant proteins found for WT SAPAP3 domain 19 vs. R770L. The gene list was analysed through GeneMANIA to search for known physical interactions between the putative interactors. The black circles (nodes) represent genes and the pink lines (edges) the interaction between the genes.

Gene ontology terms enriched among the genes in the network displayed by GeneMANIA are listed in Table 3.2. The main terms correspond to genes from the mitochondrial matrix and cellular respiration, which were highlighted in the network of Figure 3.13; once again pointing to the possible involvement of SAPAP3 in regulatory metabolic processes, but in this with much better confidence (Q-values), due to the larger size of the query gene list.

Table 3.2 - GeneMANIA functional enrichment analysis of the statistically significant proteins found for WT SAPAP3 domain 19 vs. R770L.

GO annotation	Q-value	Occurrence
Mitochondrial matrix	2.95E-28	32 (257)
Cellular respiration	6.27E-24	24 (140)
Energy derivation by oxidation of organic compounds	2.14E-20	27 (285)
Mitochondrial membrane	7.32E-19	26 (294)
Mitochondrial inner membrane	1.20E-18	23 (208)
Organelle inner membrane	4.10E-18	23 (221)
Respiratory electron transport chain	2.96E-16	17 (103)
Electron transport chain	3.08E-16	17 (104)
Mitochondrial membrane part	5.70E-16	18 (131)
Aerobic respiration	8.36E-10	9 (33)
Oxidoreductase complex	2.82E-8	10 (68)
Carboxylic acid catabolic process	4.74E-8	13 (161)
Organic acid catabolic process	4.74E-8	13 (161)
Small molecule catabolic process	4.75E-8	14 (199)
Tricarboxylic acid cycle	4.96E-8	7 (21)
Oxidative phosphorylation	2.35E-7	9 (62)
Cellular amino acid catabolic process	3.19E-7	10 (90)
ATP metabolic process	4.14E-7	12 (158)
Purine nucleoside monophosphate metabolic process	1.38E-6	12 (177)
Purine ribonucleoside monophosphate metabolic process	1.38E-6	12 (177)

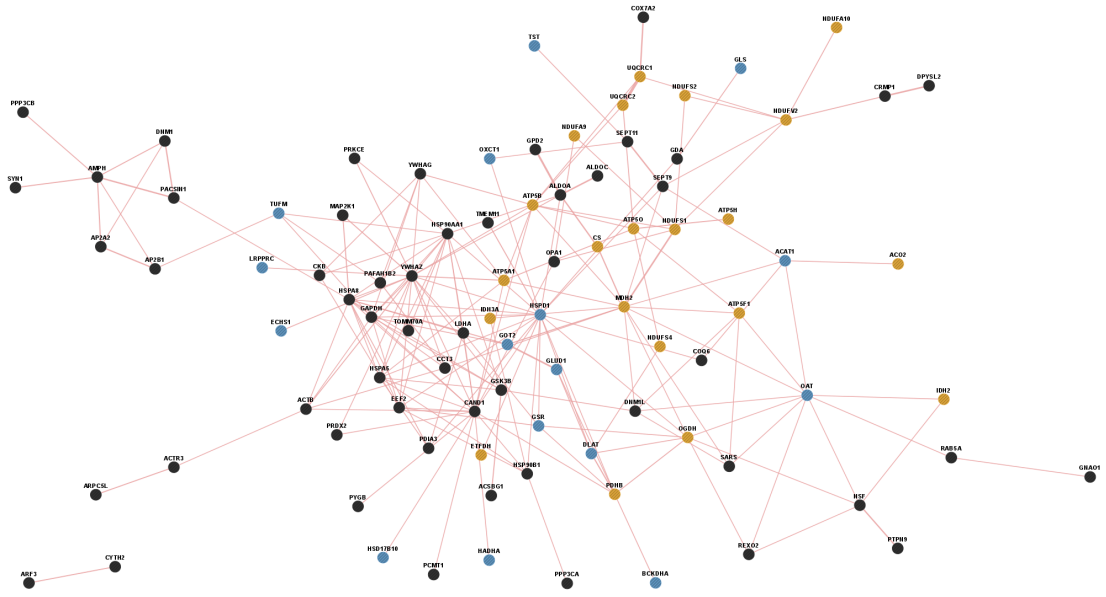


Figure 3.13 – Network view of known physical interactions for the statistically significant proteins found for WT SAPAP3 domain 19 vs. R770L with GO annotations. The yellow circles represent genes involved in cellular respiration, and the blue circles genes related with mitochondria matrix.

Taken together, these results suggest the putative association of SAPAP3 with mitochondrial proteins, which could be perturbed by SAPAP3 disease-associated mutations.

3.3.2 Identification of SAPAP3 domain 19 interaction network

With the previous results pointing towards a putative physiological interaction of SAPAP and mitochondrial proteins, we focused our attention on the whole list of interacting partners of the WT SAPAP3 domain 19, in order to confirm such association and possibly raising other clues about the physiological role of SAPAP3 at the postsynaptic density.

When analysing the list of validated proteins of the DDA analysis for the WT SAPAP3 domain 19, we were able to identify 594 proteins (Supplementary Table 6.5), based on a confidence threshold of FDR < 1 %.

Considering the identity of the 594 proteins identified as putative interactors of WT SAPAP3 domain 19, we have performed a similar GeneMANIA analysis as done with the list of the differentially interacting proteins between the wild-type and mutant forms of SAPAP3 domain 19. The analysis results presented in Figure 3.14 have shown that 541 out of the 594 proteins submitted were shown to be involved in known 4158 physical interactions with other proteins present in the query list.

The presence of clusters of known physical interactions in the network suggests that putative partners may share some common functional pathways, and that in addition to the identification of protein interactions, protein complexes can be identified as well.

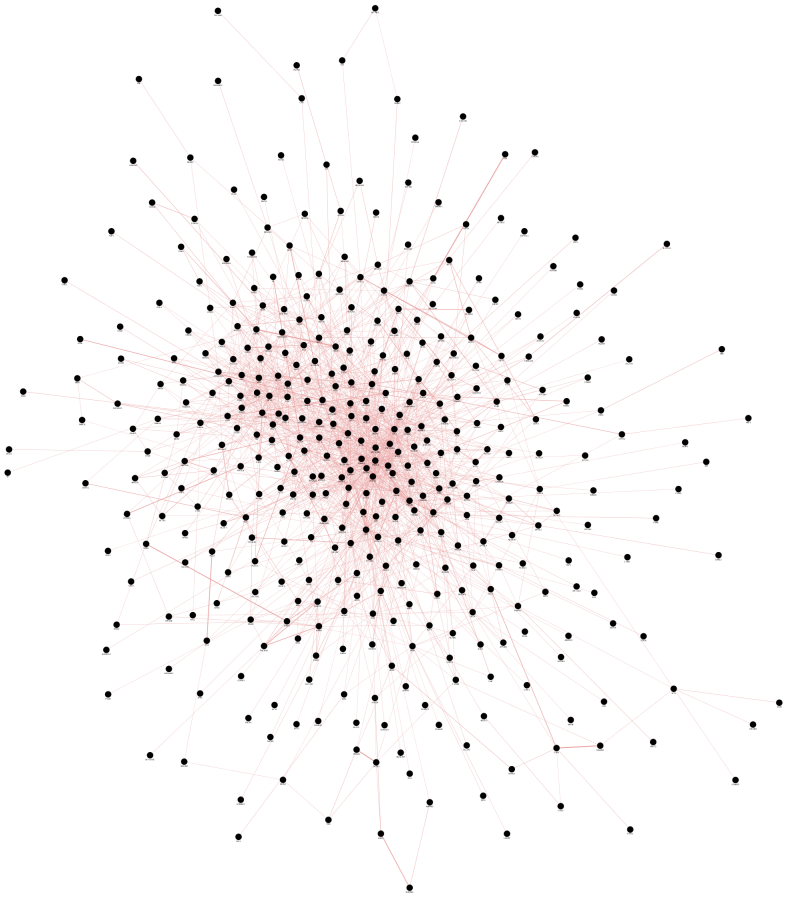


Figure 3.14 – Network view of known physical interactions within the putative interactors found for SAPAP3 domain 19. The SAPAP3 domain 19 putative interactors found in the present work were analysed through GeneMANIA Cytoscape App to search for known physical interactions between the putative interactors. The black circles (nodes) are genes and the pink lines (edges) the interaction between the genes.

The Gene Ontology (GO) analysis for the putative interactors of SAPAP3 domain 19 (the most enriched GO terms identified through GeneMANIA algorithm are listed in Table 3.1), confirmed the strong overrepresentation of GO terms involved in metabolic processes, cellular respiration and mitochondrial components. As expected, and as the query list involved a significantly higher number of proteins, the confidence values are significantly enhanced.

Table 3.3 – GeneMANIA functional enrichment analysis of the SAPAP3 domain 19 putative interactors.

GO term	Q-value	Occurrences
Mitochondrial matrix	1.5E-65	81 (257)
Cellular respiration	1.4E-43	51 (140)
Energy derivation by oxidation of organic compounds	1.1E-38	62 (285)
Mitochondrial inner membrane	1.7E-36	53 (208)
Organelle inner membrane	4.1E-35	53(221)
MItochondrial membrane	8.1E-35	59 (294)
Respiratory electron transport chain	1.3E-29	36 (103)
Electron transport chain	1.7E-29	36 (104)
Aerobic respiration	6.8E-20	19 (33)
Proton transport	1.9E-19	25 (75)
Hydrogen transport	3.6E-19	25 (77)
Monovalent inorganic cation transport	1.1E-18	41 (272)
Carboxylic acid catabolic process	7.1E-17	31 (161)
Organic acid catabolic process	7.1E-17	31 (161)
Mitochondrial membrane part	2.3E-16	28 (131)
Small molecule catabolic process	4.8E-16	33 (199)
Hydrogen ion transmembrane transport	1.5E-15	18 (44)
Mitochondrion organization	2.5E-15	34 (225)
ATPase activity	1.3E-11	28 (197)
Glucose metabolic process	4.0E-11	25 (161)

To verify if the overrepresentation of putative partners involved in mitochondrial and metabolic processes occurred as a consequence of a specific binding of these group of proteins to SAPAP3 domain 19, or if we were obtaining a biased binding towards the Ni^{2+} magnetic beads, we have compared the results obtained for the WT SAPAP3 domain19 with the naked beads control. For this, a GO cellular component comparative enrichment analysis with the Gorilla tool[193] (**G**ene **O**ntology en**RI**chment ana**L**ysis and visua**L**iz**A**tion tool) was performed for the SAPAP3 domain 19 condition, using the negative control as background. In this case, it was used a different gene ontology tool, since GeneMANIA did not allow this kind of comparative analysis. The output file showed an overrepresentation of proteins localized in the mitochondrial part for SAPAP3 domain 19 condition (Figure 3.15), when compared to the negative control, showing that this group of proteins was specifically binding to SAPAP3 domain 19.

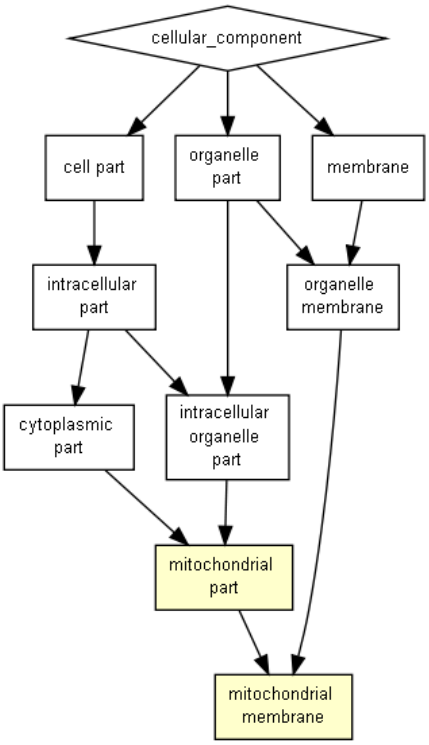


Figure 3.15 – GO cellular component enrichment analysis with GOrilla toll of SAPAP3 domain 19 condition, using the negative control condition as background. Each box is a GO term, which are linked by directed edges representing parent-to-child relationships. Yellow boxes present a p-value between 10^{-3} to 10^{-5}.

Taken together, these results suggest a new, unreported to our knowledge role for SAPAP3, apart from bridging PSD-95 and Shank, which will be further discussed in the following section.

3.4 Discussion

The present chapter aimed to identify novel SAPAP3 domain 19 interacting proteins and quantify the interacting changes as a result of SAPAP3 mutants K910R and R770L. For this purpose, affinity purification samples resulted from the incubation of the recombinant proteins with synaptoneurosome extracts were analysed by a high-throughput, sensitive and unbiased method called SWATH-MS.

As a result, 502 distinct proteins were identified and relative quantified for all the experimental conditions. Among these group of proteins, 43 proteins were identified as being differentially present between the wild type and K910R mutant conditions, and 124 proteins between the wild type and R770L mutant conditions. The number of proteins statistically different found for WT/R770L was higher than for WT/K910R, which was expectable since R770L mutant involved a modification of a positively charged amino acid (arginine) for a hydrophobic amino acid (leucine), while in the case of K910R mutant a positively charged amino acid (lysine) was replaced by another positively charge amino acid (arginine). To understand which groups of proteins were differentially affected between WT vs. OCD (K910R) mutant and WT vs. schizophrenia (R770L) mutant a GO functional enrichment analysis was performed in this group of differential interacting proteins. The results clearly reveal a preponderance of proteins related with mitochondria, metabolism and cellular respiration (Table 3.1 and Table 3.2).

Considering the whole list of putative interacting partners of WT SAPAP3 domain 19, it was created an interactive map with the already known physical protein-protein interactions, where 541 proteins present 4158 known physical interactions. (Figure 3.14). The presence of several known physical interactions between the members of the network indicates that putative partners may share some common functional pathways, and that in addition to the identification of protein interactions, protein complexes could probably be identified as well. In this analysis the SAPAP3 gene was included in the obtained network, to check if known physical interactors of SAPAP3 were detected in our analysis. Focusing on SAPAP3 (*DLGAP3*) it was possible to identify five already known interactors – PSD-95 (*DLG4*), SAP-97 (*DLG1*), Tyrosine-protein kinase FYN (*FYN*), Cullin-3 (*CUL3*) and Cullin-associated NEDD8 - dissociated protein 1 (*CAND1*) – were detected as reported partners of our bait protein (Figure 3.16).

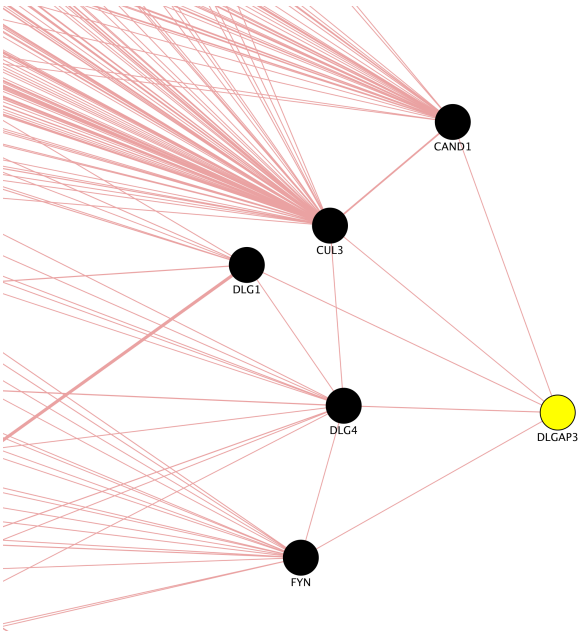


Figure 3.16 – Closed-up view of SAPAP3 (*Dlgap3*) known physical interactions between the putative interactors of the network. The image presents the already known physical interactions between SAPAP3 (*Dlgap3*, yellow circle) and PSD-95 (*Dlg4*), SAP-97 (*Dlg1*), Tyrosine-protein kinase Fyn (*Fyn*), Cullin-3 (*Cul3*) and Cullin-associated NEDD8-dissociated protein 1 (*Cand1*) (black circles).

The interaction of SAPAP3 with PSD-95 (*DLG4*), a member of membrane-associated guanylate kinase, was first identified in two independent yeast two-hybrid experiments[98, 101]; as already mentioned in this document. This interaction is mediated by the GK domain of PSD-95 and the N terminal region of SAPAP containing multiple 14 amino acids repeats. The sequence of the recombinant SAPAP3 domain 19 used in the present study did not comprise the described interaction region of SAPAP3, leading us to question whether the interaction between these two proteins could also be mediated by a different region of SAPAP3 sequence. To validate this interaction, we have co-expressed SAPAP3 domain 19 with FLAG tagged SH3-GK domain of PSD-95. After purification of SAPAP3 domain 19 by Ni^{2+} NTA affinity chromatography, we have performed an anti-FLAG Western blot (Figure 3.17), which confirmed the co-purification of PSD-95 SH3-GK domain, and the interaction with SAPAP3 C-terminal domain.

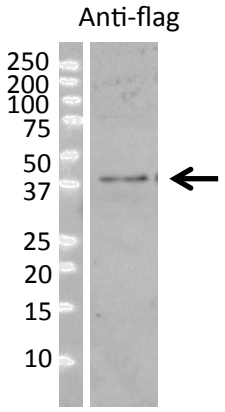


Figure 3.17 – Western blot analysis with anti-flag antibody after SAPAP3 domain 19 Ni^{2+}NTA purification when co-expressed SH3-GK domain of PSD-95 tagged with flag. The expected molecular weight for SH3-GK domain of PSD-95 is indicated with the black arrow.

These two results suggested that in addition to the interaction through the N-terminus region containing the multiple 14 amino acids repeats of SAPAP; PSD-95 could also bind to SAPAP through the C-terminus region. However, this information should be further confirmed using other techniques, such as surface plasmon resonance or isothermal titration calorimetry.

In one of the mentioned yeast two-hybrid works, where PSD-95 was identified as an interactor of SAPAP[98], the authors also examined whether SAPAP interacts with other member of membrane-associated guanylate kinase, and in these experiments SAP-97 (*DLG1*) was identified as an interactor. Our results point towards an interaction occurring via the C-terminus region of SAPAP, or mediated through an intermediate partner.

The observed interactions of SAPAP3 with Cullin-3 (CUL3) and with Cullin-associated NEDD8-dissociated protein 1 (CAND1) were previously reported in a proteomic screening[194], using CUL3 and CAND1 as bait proteins. CUL3, which activity is regulated by CAND1, is a component of the cullin-RING ubiquitin ligase, involved in diverse cellular processes by targeting numerous proteins[195]. In the mentioned proteomic study, the dynamics of cullin-RING ubiquitin ligase network was explored, and the interaction with SAPAP3 was found using CUL3 and CAND1 as bait proteins, however these interactions were not further explored due to lack of statistical significance, but the detection of both proteins as SAPAP3 C-terminal domain interactors in the current study, further reinforces the initial results and render such interactions amenable to be further validated.

Finally, the association between SAPAP3 and FYN, or more specifically its interaction with the SH3 domain of FYN, was observed in a peptide array target screening, where the main goal consisted in identification of human protein-protein interactions mediated by SH3 domains[196]. FYN, a tyrosine protein kinase was reported to be involved in the regulation of NMDA receptors[197], but its interaction with SAPAP3 was not further explored in the referred study. However, this evidence points towards an affinity of SAPAP3 for SH3 domains, which could not only explain the SAPAP3-FYN interaction observed in our study, but also the SAPAP3 interaction with PSD-95, in the way that it could also be mediated by the SH3 domain of PSD-95 and not via the GK domain as previously reported. Interestingly, SHANK partner was not detected as SAPAP3 domain 19 partner, which undeniably makes sense, as SAPAP3 C-terminal (QTRL) interaction sequence with SHANK is not present in the SAPAP3 domain 19. The remaining known established interactors for SAPAP3 namely FAK and PYK2 were not detected. However, taking into account that SAPAP3 interaction sequences with these protein interactors are not known; our results clearly suggest that these proteins may not interact with the region covered by domain 19 of SAPAP3.

When a gene ontology functional enrichment analysis was performed with the whole list of proteins identified as putative interactors of WT SAPAP3 domain 19, in similarity to the results obtained for the proteins that interact differentially with the SAPAP3 mutants, an overrepresentation of GO terms involved the mitochondria, such as metabolic processes, cellular respiration and mitochondrial components (Table 3.3) were recognized, suggesting a possible association of SAPAP3 with synaptic mitochondria, whose dysregulation is known to underlie several brain dysfunctions[198-200].

In fact, several lines of evidence indicate that mitochondrial abnormalities could be an important component in the neurobiology of psychiatric disorders, such as bipolar disorder, depression, autism, and schizophrenia[198]. The association of mitochondria with psychiatric disorders was evidenced by findings from genetic, post-mortem brain, brain-imaging and biomarker studies in humans with psychiatric disorders and rodent models of such disorders[201].

The brain has high-energy requirements, where approximately 20 % of the oxygen and 25 % of the glucose consumed by the human body is used for brain functions[202]. Glucose enters cells trough specific glucose transporters (GLUTs) and is phosphorylated by hexokinase (HK) to produce glucose-6-phosphate, which can be processed through different metabolic pathways[203]. In neurons, this pathway is predominantly oxidative where a series of metabolic steps take place to the full oxidation of glucose, in the mitochondria, resulting in

the production of approximately 30-36 molecules of ATP per glucose molecule[203]. Within the neurons, synapses use a large proportion of this energy, which is required for reestablishment of ion gradients that underlie synaptic and action potential signals; for the transport and posttranslational modifications of channels and receptors in the pre- and postsynaptic compartments; and for the uptake and recycling of neurotransmitters[202].

As a synaptic energy supplier, mitochondria constantly move along axons and dendrites, dividing and fusing in response to synaptic changes[201]. The mitochondrial transport across the cytoskeleton microtubules is primarily done by motor proteins, such as dynein and KIF5 and is regulated by several cellular components as shown in Figure 3.18[204].

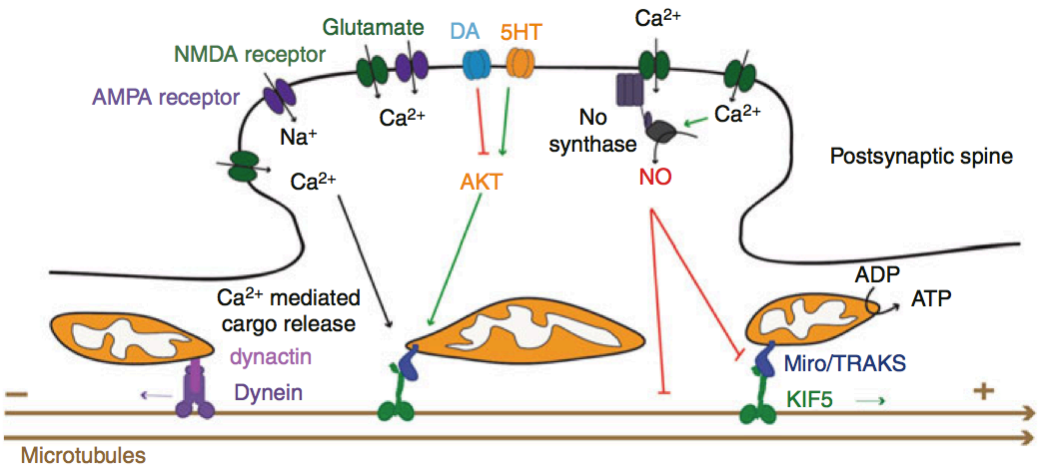


Figure 3.18 – Transport of mitochondria in the synapse. Mitochondria are primarily transported along microtubules within the synapse. Microtubules have plus and minus ends, showing mixed polarity in dendrites, whereas in axons are aligned with the plus ends directed distally. KIF5 motors move mitochondria in the plus end direction via adaptor proteins, such as Miro and TRAK. Transport in the minus direction is performed by dynein motors with adaptors such as dynactin. During excitatory transmission, glutamate activates NMDA receptors postsynaptically, which mediated the influx of Ca^{2+} into the postsynaptic site, facilitating the release and stopping of mitochondria from microtubules through its interactions with the Ca^{2+} sensitive EF hand domain of Miro. Furthermore, the NMDA receptors activate neuronal nitric oxide synthase (nNOS), where NO leads to disruption of the cytoskeleton and disrupted trafficking. Serotonin (5HT) and Dopamine (DA) receptors are also present in the postsynaptic region causing mitochondrial starting and stopping, respectively, via inverse actions on the AKT signalling pathway. Adapted from Atkin *et al.*[204]

Besides, its crucial role in the generation of energy, mitochondria are also involved in modulation of cellular calcium levels; production of free radicals; and regulation of apoptosis[201].

Calcium (Ca^{2+}) contributes to the regulation of both neurotransmission and neuronal plasticity in the brain[201]. Basal intracellular Ca^{2+} levels are extremely low compared with

those of the extracellular space and endoplasmic reticulum (ER), and several mechanisms that regulate the restoration and maintenance of these levels are mediated by mitochondria. The influx of Ca^{2+} to the mitochondrial matrix is increased in conditions of high cytosolic Ca^{2+} levels and low ATP/ADP ratios, and conversely lowered in situations of low cytosolic Ca^{2+} levels and high ATP/ADP ratios[201]. Additionally, mitochondria can function at some synapses as sequester of cytosolic Ca^{2+} and a source of cytosolic Ca^{2+} through persistent release from the mitochondrial matrix, contributing to the time course and amplitude of neurotransmitter release from the presynaptic site[201].

Mitochondria are the main intracellular sources of reactive oxygen species (ROS), and therefore are the prime target of cell oxidative damage[205]. Despite their damage potential, ROS are important molecules that act on signaling functions in physiological processes, including synaptic plasticity, learning and memory[205]. On the other hand, mitochondria possesses various antioxidant defenses, such as coenzyme Q10 (ubiquinone), creatine, nicotinamide, and glutathione, for ROS balance to interrupt or minimize oxidative processes[205].

Mitochondria are also associated with apoptosis, which comprises a complex network of physiologic biochemical pathways of programmed cell death, since most neurons are eliminated soon after synaptic contacts have been made between neurons and their targets[206]. One way to activate apoptosis involves mitochondria; however, in spite of occurring in physiological conditions, apoptosis also occurs in a range of pathological conditions including psychiatric disorders[206].

Consequently, taken together mitochondrial dysfunction not only impairs energy production but also affects other key brain cellular processes.

Over the past years, several studies suggest that oxidative stress could play an important role in the etiology of OCD. These studies observed increased lipid peroxidation[207]; decreased levels of vitamin E[208], catalase[209], glutathione peroxidase[209], and selenium[209]; increased superoxide dismutase[207, 209]; and changes in overall oxidative status[208, 210]. The cause of the oxidant-antioxidant imbalance observed in OCD is not known, but mitochondrial dysfunctions could be in the origin of this imbalance.

For schizophrenia, a disturbance of energy metabolism has also been suggested as part of the etiopathogenesis of schizophrenia, which is indicative of mitochondrial dysfunction[211]. Patients with schizophrenia have been associated with impaired glucose metabolism showed by lower metabolic rates, lower blood flow and altered expression of glycolytic enzymes[212, 213]. Another study combining transcriptomic, proteomic and metabolomics

approaches strongly suggests that mitochondrial dysfunction and oxidative stress are involved in the pathophysiology of schizophrenia[214]. A reduction in the number of mitochondria in the striatum tissue was observed in schizophrenia patients, suggesting a decreased capacity to respond to energy requirements[215]. Several evidences suggest also an impaired oxidative phosphorylation in schizophrenia[216], lowered activity of some tricarboxylic acid cycle components were reported in the dorsolateral prefrontal cortex in brain tissue of schizophrenia[217]. Gene expression profiles, also reveals alterations of specific metabolic pathways in schizophrenia[218]. Finally, studies with animal models showed changes in the respiratory chain of all brain regions evaluated[219], and increased oxidative stress due to mitochondrial dysfunction[220]. Concluding, the disruption in any component of the mitochondria pathways holds the potential to change its functioning and metabolism, compromising brain function[221].

To our knowledge, the present study is the first one showing an association between SAPAP3 and mitochondrial components. Although the specific role of this interaction is still unknown, a series of functional studies, could provide further clues about this association. In line with this hypothesis, it should be interesting to evaluate the mitochondrial dynamics in SAPAP3 knockout animals, using for example high-resolution imaging of mitochondria in live cells. It could be also interesting to evaluate the mitochondrial shape, the abundance of mitochondrial complexes and the mitochondrial respiration in the knockout animals when compared to controls. Interestingly, to reinforce the association between SAPAP3 and mitochondria, we found dynein in our network of interactor partners, which is a mitochondria associated motor protein that participates in the transport of mitochondria along the dendrites. As mentioned in the first section of this document, the interaction between dynein and SAPAP1 has already been described[112, 113], however none of those works stated the involvement of SAPAP with mitochondria through its interaction with dynein. Consequently, taking into account our data, we hypothesize that SAPAP3 may also be involved in the positioning of mitochondria within the dendrites. In fact, the mitochondrial transport within the dendrite is important not only as a source of energy, but also for neuroplasticity processes including neural differentiation, neurite outgrowth, neurotransmitters package and release, as well as dendritic remodeling[211], which disruption could contribute to the pathogenesis of psychiatric disorders such as OCD and schizophrenia.

4. Concluding remarks

SAPAP3, a scaffolding protein that is found in excitatory, glutamate-responsive synapses and is highly expressed only in the striatum region of the brain, has been target of substantial interest in OCD and more recently in schizophrenia. This protein is implicated in the molecular organization of synapses and neuronal cell signalling, linking ion channels to the sub-synaptic cytoskeleton. Because this type of proteins work as complexes to regulate diverse biological processes, discovery of SAPAP3 protein-protein interactions and exploration of differences between the PPIs network in abnormal conditions may provide useful information toward understanding the complexity and the pathogenesis of the related disorders.

In the present work, it was developed a toolbox for the biochemical, structural and interactomic characterization of SAPAP3. First, through the use of ESPRIT we have successfully identified soluble domains of SAPAP3 covering a wide region of the protein sequence, and selected the best expressing soluble clone, corresponding to C-terminal region to proceed with its characterization. Despite we have only explored the C-terminus of SAPAP3 through the production of domain 19, the ESPRIT technology paved the way and provided the necessary tools to study other regions of the protein.

The protein quality was confirmed by analytical approaches, and the structural characterization was initiated by a preliminary crystallization trial, which yielded putative protein crystals, whose quality will soon be assessed.

To obtain further insights into the function of this specific domain of SAPAP3, an interactomic analysis was performed where novel putative interactor partners were identified, with some of these partners present in significantly different levels, as a consequence of SAPAP3 OCD and schizophrenia associated point mutations. Additionally, five previously described interacting partners were also detected in the network, validating our technical approach. Gene Ontology functional enrichment analysis, using the putative interactors, showed a significant overrepresentation of proteins associated with mitochondria and related metabolism. To our knowledge, this is the first study presenting an association between SAPAP3 and mitochondria components. Although the specific role of this interaction is still unknown, we hypothesize the potential involvement of SAPAP3 in the positioning of mitochondria within the dendrites. Nevertheless, a series of functional studies need to be performed in the near future to validate such association and to provide further clues about its functional significance.

Concluding, the present work opens the way to study what we believe to be a novel role for SAPAP3, through the identified association with mitochondria related components, which

may also help elucidating the functional relevance of the protein mutants in the pathophysiology of OCD and schizophrenia.

5. References

1. Weissman MM, Bland RC, Canino GJ, *et al.* The cross national epidemiology of obsessive compulsive disorder. The Cross National Collaborative Group. The Journal of clinical psychiatry. 1994;55 Suppl:5-10.

2. Karno M, Golding JM, Sorenson SB, *et al.* The epidemiology of obsessive-compulsive disorder in five US communities. Archives of general psychiatry. 1988;45(12):1094-1099.

3. Ruscio AM, Stein DJ, Chiu WT, *et al.* The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. Molecular psychiatry. 2010;15(1):53-63.

4. Miguel EC, Leckman JF, Rauch S, *et al.* Obsessive-compulsive disorder phenotypes: implications for genetic studies. Molecular psychiatry. 2005;10(3):258-275.

5. American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders : DSM-5. Washington, D.C.: American Psychiatric Association 2013.

6. Seibell PJ, Hollander E. Management of obsessive-compulsive disorder. F1000prime reports. 2014;6:68.

7. Goodman WK, Grice DE, Lapidus KA, *et al.* Obsessive-compulsive disorder. The Psychiatric clinics of North America. 2014;37(3):257-267.

8. Hyman SE. Neuroscience: obsessed with grooming. Nature. 2007;448(7156):871-872.

9. Abramowitz JS, Taylor S, McKay D. Obsessive-compulsive disorder. Lancet. 2009;374(9688):491-499.

10. Haynes WI, Mallet L. High-frequency stimulation of deep brain structures in obsessive-compulsive disorder: the search for a valid circuit. The European journal of neuroscience. 2010;32(7):1118-1127.

11. Geller DA. Obsessive-compulsive and spectrum disorders in children and adolescents. The Psychiatric clinics of North America. 2006;29(2):353-370.

12. Grant JE, Mancebo MC, Pinto A, *et al.* Late-onset obsessive compulsive disorder: clinical characteristics and psychiatric comorbidity. Psychiatry research. 2007;152(1):21-27.

13. Hollander E, Greenwald S, Neville D, *et al.* Uncomplicated and comorbid obsessive-compulsive disorder in an epidemiologic sample. Depression and anxiety. 1996;4(3):111-119.

14. Stein DJ. Obsessive-compulsive disorder. Lancet. 2002;360(9330):397-405.

15. Browne HA, Gair SL, Scharf JM, *et al.* Genetics of obsessive-compulsive disorder and related disorders. The Psychiatric clinics of North America. 2014;37(3):319-335.

16. McKay D, Abramowitz JS, Calamari JE, *et al.* A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. Clinical psychology review. 2004;24(3):283-313.

17. Goodman WK, Price LH, Rasmussen SA, *et al.* The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Archives of general psychiatry. 1989;46(11):1006-1011.

18. Mataix-Cols D, Rauch SL, Manzo PA, *et al.* Use of factor-analyzed symptom dimensions to predict outcome with serotonin reuptake inhibitors and placebo in the treatment of obsessive-compulsive disorder. The American journal of psychiatry. 1999;156(9):1409-1416.

19. Girishchandra BG, Khanna S. Phenomenology of obsessive compulsive disorder: a factor analytic approach. Indian journal of psychiatry. 2001;43(4):306-316.

20. Cavallini MC, Di Bella D, Siliprandi F, *et al.* Exploratory factor analysis of obsessive-compulsive patients and association with 5-HTTLPR polymorphism. American journal of medical genetics. 2002;114(3):347-353.

21. Feinstein SB, Fallon BA, Petkova E, *et al.* Item-by-item factor analysis of the Yale-Brown Obsessive Compulsive Scale Symptom Checklist. The Journal of neuropsychiatry and clinical neurosciences. 2003;15(2):187-193.

22. Kim SJ, Lee HS, Kim CH. Obsessive-compulsive disorder, factor-analyzed symptom dimensions and serotonin transporter polymorphism. Neuropsychobiology. 2005;52(4):176-182.

23. Delorme R, Bille A, Betancur C, *et al.* Exploratory analysis of obsessive compulsive symptom dimensions in children and adolescents: a prospective follow-up study. BMC psychiatry. 2006;6:1.

24. McKay D, Piacentini J, Greisberg S, *et al.* The structure of childhood obsessions and compulsions: dimensions in an outpatient sample. Behaviour research and therapy. 2006;44(1):137-146.

25. Cullen B, Brown CH, Riddle MA, *et al.* Factor analysis of the Yale-Brown Obsessive Compulsive Scale in a family study of obsessive-compulsive disorder. Depression and anxiety. 2007;24(2):130-138.

26. Hasler G, Pinto A, Greenberg BD, *et al.* Familiality of factor analysis-derived YBOCS dimensions in OCD-affected sibling pairs from the OCD Collaborative Genetics Study. Biological psychiatry. 2007;61(5):617-625.

27. Pinto A, Eisen JL, Mancebo MC, *et al.* Taboo thoughts and doubt/checking: a refinement of the factor structure for obsessive-compulsive disorder symptoms. Psychiatry research. 2007;151(3):255-258.

28. Stein DJ, Andersen EW, Overo KF. Response of symptom dimensions in obsessive-compulsive disorder to treatment with citalopram or placebo. Revista brasileira de psiquiatria. 2007;29(4):303-307.

29. Stewart SE, Rosario MC, Brown TA, *et al.* Principal components analysis of obsessive-compulsive disorder symptoms in children and adolescents. Biological psychiatry. 2007;61(3):285-291.

30. Pauls DL, Abramovitch A, Rauch SL, *et al.* Obsessive-compulsive disorder: an integrative genetic and neurobiological perspective. Nature reviews Neuroscience. 2014;15(6):410-424.

31. Kobak KA, Greist JH, Jefferson JW, *et al.* Behavioral versus pharmacological treatments of obsessive compulsive disorder: a meta-analysis. Psychopharmacology. 1998;136(3):205-216.

32. Schruers K, Koning K, Luermans J, *et al.* Obsessive-compulsive disorder: a critical review of therapeutic perspectives. Acta psychiatrica Scandinavica. 2005;111(4):261-271.

33. Bokor G, Anderson PD. Obsessive-compulsive disorder. Journal of pharmacy practice. 2014;27(2):116-130.

34. Franklin ME, Abramowitz JS, Kozak MJ, *et al.* Effectiveness of exposure and ritual prevention for obsessive-compulsive disorder: randomized compared with nonrandomized samples. Journal of consulting and clinical psychology. 2000;68(4):594-602.

35. Foa EB. Cognitive behavioral therapy of obsessive-compulsive disorder. Dialogues in clinical neuroscience. 2010;12(2):199-207.

36. Abramowitz JS. The psychological treatment of obsessive-compulsive disorder. Canadian journal of psychiatry Revue canadienne de psychiatrie. 2006;51(7):407-416.

37. Math SB, Janardhan Reddy YC. Issues in the pharmacological treatment of obsessive-compulsive disorder. International journal of clinical practice. 2007;61(7):1188-1197.

38. Kellner M. Drug treatment of obsessive-compulsive disorder. Dialogues in clinical neuroscience. 2010;12(2):187-197.

39. Abudy A, Juven-Wetzler A, Zohar J. Pharmacological management of treatment-resistant obsessive-compulsive disorder. CNS drugs. 2011;25(7):585-596.

40. Zohar J, Judge R. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. OCD Paroxetine Study Investigators. The British journal of psychiatry : the journal of mental science. 1996;169(4):468-474.

41. Aouizerate B, Guehl D, Cuny E, *et al.* Pathophysiology of obsessive-compulsive disorder: a necessary link between phenomenology, neuropsychology, imagery and physiology. Progress in neurobiology. 2004;72(3):195-221.

42. Marazziti D, Consoli G. Treatment strategies for obsessive-compulsive disorder. Expert opinion on pharmacotherapy. 2010;11(3):331-343.

43. Kordon A, Zurowski B, Wahl K, *et al.* [Evidence-based pharmacotherapy and other somatic treatment approaches for obsessive-compulsive disorder: state of the art]. Der Nervenarzt. 2011;82(3):319-320, 322-314.

44. de Koning PP, Figee M, van den Munckhof P, *et al.* Current status of deep brain stimulation for obsessive-compulsive disorder: a clinical review of different targets. Current psychiatry reports. 2011;13(4):274-282.

45. Ting JT, Feng G. Neurobiology of obsessive-compulsive disorder: insights into neural circuitry dysfunction through mouse genetics. Current opinion in neurobiology. 2011;21(6):842-848.

46. Pittenger C, Bloch MH, Williams K. Glutamate abnormalities in obsessive compulsive disorder: neurobiology, pathophysiology, and treatment. Pharmacology & therapeutics. 2011;132(3):314-332.

47. Ting JT, Feng G. Glutamatergic Synaptic Dysfunction and Obsessive-Compulsive Disorder. Current chemical genomics. 2008;2:62-75.

48. Brem S, Hauser TU, Iannaccone R, *et al.* Neuroimaging of cognitive brain function in paediatric obsessive compulsive disorder: a review of literature and preliminary meta-analysis. Journal of neural transmission. 2012;119(11):1425-1448.

49. van den Heuvel OA, van der Werf YD, Verhoef KM, *et al.* Frontal-striatal abnormalities underlying behaviours in the compulsive-impulsive spectrum. Journal of the neurological sciences. 2010;289(1-2):55-59.

50. Kalra SK, Swedo SE. Children with obsessive-compulsive disorder: are they just "little adults"? The Journal of clinical investigation. 2009;119(4):737-746.

51. Menzies L, Chamberlain SR, Laird AR, *et al.* Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. Neuroscience and biobehavioral reviews. 2008;32(3):525-549.

52. Baxter LR, Jr., Phelps ME, Mazziotta JC, *et al.* Local cerebral glucose metabolic rates in obsessive-compulsive disorder. A comparison with rates in unipolar depression and in normal controls. Archives of general psychiatry. 1987;44(3):211-218.

53. Nordahl TE, Benkelfat C, Semple WE, *et al.* Cerebral glucose metabolic rates in obsessive compulsive disorder. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 1989;2(1):23-28.

54. Swedo SE, Schapiro MB, Grady CL, *et al.* Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder. Archives of general psychiatry. 1989;46(6):518-523.

55. Adler CM, McDonough-Ryan P, Sax KW, *et al.* fMRI of neuronal activation with symptom provocation in unmedicated patients with obsessive compulsive disorder. Journal of psychiatric research. 2000;34(4-5):317-324.

56. Breiter HC, Rauch SL, Kwong KK, *et al.* Functional magnetic resonance imaging of symptom provocation in obsessive-compulsive disorder. Archives of general psychiatry. 1996;53(7):595-606.

57. Rauch SL, Jenike MA, Alpert NM, *et al.* Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. Archives of general psychiatry. 1994;51(1):62-70.

58. Baxter LR, Jr., Schwartz JM, Bergman KS, *et al.* Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. Archives of general psychiatry. 1992;49(9):681-689.

59. Benkelfat C, Nordahl TE, Semple WE, *et al.* Local cerebral glucose metabolic rates in obsessive-compulsive disorder. Patients treated with clomipramine. Archives of general psychiatry. 1990;47(9):840-848.

60. Saxena S, Brody AL, Ho ML, *et al.* Differential cerebral metabolic changes with paroxetine treatment of obsessive-compulsive disorder vs major depression. Archives of general psychiatry. 2002;59(3):250-261.

61. Schwartz JM, Stoessel PW, Baxter LR, Jr., *et al.* Systematic changes in cerebral glucose metabolic rate after successful behavior modification treatment of obsessive-compulsive disorder. Archives of general psychiatry. 1996;53(2):109-113.

62. Dougherty DD, Baer L, Cosgrove GR, *et al.* Prospective long-term follow-up of 44 patients who received cingulotomy for treatment-refractory obsessive-compulsive disorder. The American journal of psychiatry. 2002;159(2):269-275.

63. Ahmari SE, Spellman T, Douglass NL, *et al.* Repeated cortico-striatal stimulation generates persistent OCD-like behavior. Science. 2013;340(6137):1234-1239.

64. Burguiere E, Monteiro P, Feng G, *et al.* Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. Science. 2013;340(6137):1243-1246.

65. van Grootheest DS, Cath DC, Beekman AT, *et al.* Twin studies on obsessive-compulsive disorder: a review. Twin research and human genetics : the official journal of the International Society for Twin Studies. 2005;8(5):450-458.

66. Pauls DL, Alsobrook JP, 2nd, Goodman W, *et al.* A family study of obsessive-compulsive disorder. The American journal of psychiatry. 1995;152(1):76-84.

67. Nestadt G, Samuels J, Riddle M, *et al.* A family study of obsessive-compulsive disorder. Archives of general psychiatry. 2000;57(4):358-363.

68. Nakatani E, Krebs G, Micali N, *et al.* Children with very early onset obsessive-compulsive disorder: clinical features and treatment outcome. Journal of child psychology and psychiatry, and allied disciplines. 2011;52(12):1261-1268.

69. Goodman WK, Lydiard RB. Recognition and treatment of obsessive-compulsive disorder. The Journal of clinical psychiatry. 2007;68(12):e30.

70. Nestadt G, Grados M, Samuels JF. Genetics of obsessive-compulsive disorder. The Psychiatric clinics of North America. 2010;33(1):141-158.

71. Flaisher-Grinberg S, Klavir O, Joel D. The role of 5-HT2A and 5-HT2C receptors in the signal attenuation rat model of obsessive-compulsive disorder. The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum. 2008;11(6):811-825.

72. Tizabi Y, Louis VA, Taylor CT, *et al.* Effect of nicotine on quinpirole-induced checking behavior in rats: implications for obsessive-compulsive disorder. Biological psychiatry. 2002;51(2):164-171.

73. Perani D, Garibotto V, Gorini A, *et al.* In vivo PET study of 5HT(2A) serotonin and D(2) dopamine dysfunction in drug-naive obsessive-compulsive disorder. NeuroImage. 2008;42(1):306-314.

74. Pauls DL. The genetics of obsessive-compulsive disorder: a review. Dialogues in clinical neuroscience. 2010;12(2):149-163.

75. Chakrabarty K, Bhattacharyya S, Christopher R, *et al.* Glutamatergic dysfunction in OCD. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 2005;30(9):1735-1740.

76. Coric V, Taskiran S, Pittenger C, *et al.* Riluzole augmentation in treatment-resistant obsessive-compulsive disorder: an open-label trial. Biological psychiatry. 2005;58(5):424-428.

77. Aboujaoude E, Barry JJ, Gamel N. Memantine augmentation in treatment-resistant obsessive-compulsive disorder: an open-label trial. Journal of clinical psychopharmacology. 2009;29(1):51-55.

78. Kushner MG, Kim SW, Donahue C, *et al.* D-cycloserine augmented exposure therapy for obsessive-compulsive disorder. Biological psychiatry. 2007;62(8):835-838.

79. Hanna GL, Veenstra-VanderWeele J, Cox NJ, *et al.* Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. American journal of medical genetics. 2002;114(5):541-552.

80. Arnold PD, Sicard T, Burroughs E, *et al.* Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder. Archives of general psychiatry. 2006;63(7):769-776.

81. Dickel DE, Veenstra-VanderWeele J, Cox NJ, *et al.* Association testing of the positional and functional candidate gene SLC1A1/EAAC1 in early-onset obsessive-compulsive disorder. Archives of general psychiatry. 2006;63(7):778-785.

82. Arnold PD, Rosenberg DR, Mundo E, *et al.* Association of a glutamate (NMDA) subunit receptor gene (GRIN2B) with obsessive-compulsive disorder: a preliminary study. Psychopharmacology. 2004;174(4):530-538.

83. Delorme R, Krebs MO, Chabane N, *et al.* Frequency and transmission of glutamate receptors GRIK2 and GRIK3 polymorphisms in patients with obsessive compulsive disorder. Neuroreport. 2004;15(4):699-702.

84. Sampaio AS, Fagerness J, Crane J, *et al.* Association between polymorphisms in GRIK2 gene and obsessive-compulsive disorder: a family-based study. CNS neuroscience & therapeutics. 2011;17(3):141-147.

85. Shmelkov SV, Hormigo A, Jing D, *et al.* Slitrk5 deficiency impairs corticostriatal circuitry and leads to obsessive-compulsive-like behaviors in mice. Nature medicine. 2010;16(5):598-602, 591p following 602.

86. Scimemi A, Tian H, Diamond JS. Neuronal transporters regulate glutamate clearance, NMDA receptor activation, and synaptic plasticity in the hippocampus. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2009;29(46):14581-14595.

87. Rothstein JD, Martin L, Levey AI, *et al.* Localization of neuronal and glial glutamate transporters. Neuron. 1994;13(3):713-725.

88. Aoyama K, Suh SW, Hamby AM, *et al.* Neuronal glutathione deficiency and age-dependent neurodegeneration in the EAAC1 deficient mouse. Nature neuroscience. 2006;9(1):119-126.

89. Welch JM, Lu J, Rodriguiz RM, *et al.* Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. Nature. 2007;448(7156):894-900.

90. Ting JT, Peca J, Feng G. Functional consequences of mutations in postsynaptic scaffolding proteins and relevance to psychiatric disorders. Annual review of neuroscience. 2012;35:49-71.

91. Center GSL. Learn Genetics Available: http://learn.genetics.utah.edu. Accessed 26 August 2015.

92. Sheng M, Hoogenraad CC. The postsynaptic architecture of excitatory synapses: a more quantitative view. Annual review of biochemistry. 2007;76:823-847.

93. Feng W, Zhang M. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. Nature reviews Neuroscience. 2009;10(2):87-99.

94. Carlin RK, Grab DJ, Cohen RS, *et al.* Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. The Journal of cell biology. 1980;86(3):831-845.

95. Walsh MJ, Kuruc N. The postsynaptic density: constituent and associated proteins characterized by electrophoresis, immunoblotting, and peptide sequencing. Journal of neurochemistry. 1992;59(2):667-678.

96. Verpelli C, Schmeisser MJ, Sala C, *et al.* Scaffold proteins at the postsynaptic density. Advances in experimental medicine and biology. 2012;970:29-61.

97. de Bartolomeis A, Fiore G. Postsynaptic density scaffolding proteins at excitatory synapse and disorders of synaptic plasticity: implications for human behavior pathologies. International review of neurobiology. 2004;59:221-254.

98. Takeuchi M, Hata Y, Hirao K, *et al.* SAPAPs. A family of PSD-95/SAP90-associated proteins localized at postsynaptic density. The Journal of biological chemistry. 1997;272(18):11943-11951.

99. Tong J, Yang H, Eom SH, *et al.* Structure of the GH1 domain of guanylate kinase-associated protein from Rattus norvegicus. Biochemical and biophysical research communications. 2014;452(1):130-135.

100. Sievers F, Wilm A, Dineen D, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

101. Kim E, Naisbitt S, Hsueh YP, *et al.* GKAP, a novel synaptic protein that interacts with the guanylate kinase-like domain of the PSD-95/SAP90 family of channel clustering molecules. The Journal of cell biology. 1997;136(3):669-678.

102. Welch JM, Wang D, Feng G. Differential mRNA expression and protein localization of the SAP90/PSD-95-associated proteins (SAPAPs) in the nervous system of the mouse. The Journal of comparative neurology. 2004;472(1):24-39.

103. Hung AY, Sung CC, Brito IL, *et al.* Degradation of postsynaptic scaffold GKAP and regulation of dendritic spine morphology by the TRIM3 ubiquitin ligase in rat hippocampal neurons. PloS one. 2010;5(3):e9842.

104. Chen M, Wan Y, Ade K, *et al.* Sapap3 deletion anomalously activates short-term endocannabinoid-mediated synaptic plasticity. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2011;31(26):9563-9573.

105. Wan Y, Feng G, Calakos N. Sapap3 deletion causes mGluR5-dependent silencing of AMPAR synapses. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2011;31(46):16685-16691.

106. Naisbitt S, Kim E, Tu JC, *et al.* Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. Neuron. 1999;23(3):569-582.

107. Boeckers TM, Winter C, Smalla KH, *et al.* Proline-rich synapse-associated proteins ProSAP1 and ProSAP2 interact with synaptic proteins of the SAPAP/GKAP family. Biochemical and biophysical research communications. 1999;264(1):247-252.

108. Sheng M, Kim E. The Shank family of scaffold proteins. Journal of cell science. 2000;113 (Pt 11):1851-1856.

109. Dosemeci A, Jaffe H. Regulation of phosphorylation at the postsynaptic density during different activity states of Ca2+/calmodulin-dependent protein kinase II. Biochemical and biophysical research communications. 2010;391(1):78-84.

110. Hirao K, Hata Y, Ide N, *et al.* A novel multiple PDZ domain-containing molecule interacting with N-methyl-D-aspartate receptors and neuronal cell adhesion proteins. The Journal of biological chemistry. 1998;273(33):21105-21110.

111. Kawabe H, Hata Y, Takeuchi M, *et al.* nArgBP2, a novel neural member of ponsin/ArgBP2/vinexin family that interacts with synapse-associated protein 90/postsynaptic density-95-associated protein (SAPAP). The Journal of biological chemistry. 1999;274(43):30914-30918.

112. Hirao K, Hata Y, Deguchi M, *et al.* Association of synapse-associated protein 90/ postsynaptic density-95-associated protein (SAPAP) with neurofilaments. Genes to cells : devoted to molecular & cellular mechanisms. 2000;5(3):203-210.

113. Haraguchi K, Satoh K, Yanai H, *et al.* The hDLG-associated protein DAP interacts with dynein light chain and neuronal nitric oxide synthase. Genes to cells : devoted to molecular & cellular mechanisms. 2000;5(11):905-911.

114. Naisbitt S, Valtschanoff J, Allison DW, *et al.* Interaction of the postsynaptic density-95/guanylate kinase domain-associated protein complex with a light chain of myosin-V and dynein. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2000;20(12):4524-4534.

115. Mungrue IN, Bredt DS. nNOS at a glance: implications for brain and brawn. Journal of cell science. 2004;117(Pt 13):2627-2629.

116. Hirokawa N, Niwa S, Tanaka Y. Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. Neuron. 2010;68(4):610-638.

117. Moutin E, Raynaud F, Fagni L, *et al.* GKAP-DLC2 interaction organizes the postsynaptic scaffold complex to enhance synaptic NMDA receptor activity. Journal of cell science. 2012;125(Pt 8):2030-2040.

118. Kitano J, Yamazaki Y, Kimura K, *et al.* Tamalin is a scaffold protein that interacts with multiple neuronal proteins in distinct modes of protein-protein association. The Journal of biological chemistry. 2003;278(17):14762-14768.

119. Bongiorno-Borbone L, Kadare G, Benfenati F, *et al.* FAK and PYK2 interact with SAP90/PSD-95-Associated Protein-3. Biochemical and biophysical research communications. 2005;337(2):641-646.

120. Xiong WC, Mei L. Roles of FAK family kinases in nervous system. Frontiers in bioscience : a journal and virtual library. 2003;8:s676-682.

121. Huang Y, Lu W, Ali DW, *et al.* CAKbeta/Pyk2 kinase is a signaling link for induction of long-term potentiation in CA1 hippocampus. Neuron. 2001;29(2):485-496.

122. Rauch SL, Carlezon WA, Jr. Neuroscience. Illuminating the neural circuitry of compulsive behaviors. Science. 2013;340(6137):1174-1175.

123. Wan Y, Ade KK, Caffall Z, *et al.* Circuit-selective striatal synaptic dysfunction in the Sapap3 knockout mouse model of obsessive-compulsive disorder. Biological psychiatry. 2014;75(8):623-630.

124. Bienvenu OJ, Wang Y, Shugart YY, *et al.* Sapap3 and pathological grooming in humans: Results from the OCD collaborative genetics study. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2009;150B(5):710-720.

125. Zuchner S, Wendland JR, Ashley-Koch AE, *et al.* Multiple rare SAPAP3 missense variants in trichotillomania and OCD. Molecular psychiatry. 2009;14(1):6-9.

126. Boardman L, van der Merwe L, Lochner C, *et al.* Investigating SAPAP3 variants in the etiology of obsessive-compulsive disorder and trichotillomania in the South African white population. Comprehensive psychiatry. 2011;52(2):181-187.

127. Li JM, Lu CL, Cheng MC, *et al.* Exonic resequencing of the DLGAP3 gene as a candidate gene for schizophrenia. Psychiatry research. 2013;208(1):84-87.

128. Krueger DD, Osterweil EK, Chen SP, *et al.* Cognitive dysfunction and prefrontal synaptic abnormalities in a mouse model of fragile X syndrome. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(6):2587-2592.

129. Mesholam-Gately RI, Giuliano AJ, Goff KP, *et al.* Neurocognition in first-episode schizophrenia: a meta-analytic review. Neuropsychology. 2009;23(3):315-336.

130. Insel TR. Rethinking schizophrenia. Nature. 2010;468(7321):187-193.

131. McGlashan TH, Hoffman RE. Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Archives of general psychiatry. 2000;57(7):637-648.

132. Stephan KE, Baldeweg T, Friston KJ. Synaptic plasticity and dysconnection in schizophrenia. Biological psychiatry. 2006;59(10):929-939.

133. Haller CS, Padmanabhan JL, Lizano P, *et al.* Recent advances in understanding schizophrenia. F1000prime reports. 2014;6:57.

134. Balu DT, Li Y, Puhl MD, *et al.* Multiple risk pathways for schizophrenia converge in serine racemase knockout mice, a mouse model of NMDA receptor hypofunction. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(26):E2400-2409.

135. Timms AE, Dorschner MO, Wechsler J, *et al.* Support for the N-methyl-D-aspartate receptor hypofunction hypothesis of schizophrenia from exome sequencing in multiplex families. JAMA Psychiatry. 2013;70(6):582-590.

136. Kidd SA, Lachiewicz A, Barbouth D, *et al.* Fragile X syndrome: a review of associated medical problems. Pediatrics. 2014;134(5):995-1005.

137. Yumerefendi H, Desravines DC, Hart DJ. Library-based methods for identification of soluble expression constructs. Methods. 2011;55(1):38-43.

138. Yumerefendi H, Tarendeau F, Mas PJ, *et al.* ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. Journal of structural biology. 2010;172(1):66-74.

139. Ward JJ, Sodhi JS, McGuffin LJ, *et al.* Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. Journal of molecular biology. 2004;337(3):635-645.

140. An Y, Meresse P, Mas PJ, *et al.* CoESPRIT: a library-based construct screening method for identification and expression of soluble protein complexes. PloS one. 2011;6(2):e16261.

141. Dyson MR, Perera RL, Shadbolt SP, *et al.* Identification of soluble protein fragments by gene fragmentation and genetic selection. Nucleic acids research. 2008;36(9):e51.

142. Hart DJ, Tarendeau F. Combinatorial library approaches for improving soluble protein expression in Escherichia coli. Acta crystallographica Section D, Biological crystallography. 2006;62(Pt 1):19-26.

143. Ostermeier M, Lutz S. The creation of ITCHY hybrid protein libraries. Methods in molecular biology. 2003;231:129-141.

144. Reich S, Puckey LH, Cheetham CL, *et al.* Combinatorial Domain Hunting: An effective approach for the identification of soluble protein domains adaptable to high-throughput applications. Protein science : a publication of the Protein Society. 2006;15(10):2356-2365.

145. Kawasaki M, Inagaki F. Random PCR-based screening for soluble domains using green fluorescent protein. Biochemical and biophysical research communications. 2001;280(3):842-844.

146. Thorstenson YR, Hunicke-Smith SP, Oefner PJ, *et al.* An automated hydrodynamic process for controlled, unbiased DNA shearing. Genome research. 1998;8(8):848-855.

147. Pedelacq JD, Piltch E, Liong EC, *et al.* Engineering soluble proteins for structural genomics. Nature biotechnology. 2002;20(9):927-932.

148. Maxwell KL, Mittermaier AK, Forman-Kay JD, *et al.* A simple in vivo assay for increased protein solubility. Protein science : a publication of the Protein Society. 1999;8(9):1908-1911.

149. Waugh DS. Making the most of affinity tags. Trends in biotechnology. 2005;23(6):316-320.

150. Cabantous S, Waldo GS. In vivo and in vitro protein solubility assays using split GFP. Nature methods. 2006;3(10):845-854.

151. Listwan P, Terwilliger TC, Waldo GS. Automated, high-throughput platform for protein solubility screening using a split-GFP system. Journal of structural and functional genomics. 2009;10(1):47-55.

152. Unger T, Jacobovitch Y, Dantes A, *et al.* Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. Journal of structural biology. 2010;172(1):34-44.

153. Sun G, Anderson VE. Prevention of artifactual protein oxidation generated during sodium dodecyl sulfate-gel electrophoresis. Electrophoresis. 2004;25(7-8):959-965.

154. Greenfield NJ. Analysis of the kinetics of folding of proteins and peptides using circular dichroism. Nature protocols. 2006;1(6):2891-2899.

155. An Y, Yumerefendi H, Mas PJ, *et al.* ORF-selector ESPRIT: a second generation library screen for soluble protein expression employing precise open reading frame selection. Journal of structural biology. 2011;175(2):189-197.

156. Biegert A, Mayer C, Remmert M, *et al.* The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic acids research. 2006;34(Web Server issue):W335-339.

157. Christ D, Winter G. Identification of protein domains by shotgun proteolysis. Journal of molecular biology. 2006;358(2):364-371.

158. Gao X, Bain K, Bonanno JB, *et al.* High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. Journal of structural and functional genomics. 2005;6(2-3):129-134.

159. Bird LE. High throughput construction and small scale expression screening of multi-tag vectors in Escherichia coli. Methods. 2011;55(1):29-37.

160. Francis DM, Page R. Strategies to optimize protein expression in E. coli. Curr Protoc Protein Sci. 2010;Chapter 5:Unit 5 24 21-29.

161. Gopal GJ, Kumar A. Strategies for the production of recombinant protein in Escherichia coli. Protein J. 2013;32(6):419-425.

162. Structural Genomics C, China Structural Genomics C, Northeast Structural Genomics C, *et al.* Protein production and purification. Nature methods. 2008;5(2):135-146.

163. Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell. 1998;92(3):291-294.

164. Kann MG. Protein interactions and disease: computational approaches to uncover the etiology of diseases. Briefings in bioinformatics. 2007;8(5):333-346.

165. Ngounou Wetie AG, Sokolowska I, Woods AG, *et al.* Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cellular and molecular life sciences : CMLS. 2014;71(2):205-228.

166. Marcilla M, Albar JP. Quantitative proteomics: A strategic ally to map protein interaction networks. IUBMB life. 2013;65(1):9-16.

167. Bonetta L. Protein-protein interactions: Interactome under construction. Nature. 2010;468(7325):851-854.

168. Lambert JP, Ivosev G, Couzens AL, *et al.* Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. Nature methods. 2013;10(12):1239-1245.

169. Chen JY, Youn E, Mooney SD. Connecting protein interaction data, mutations, and disease using bioinformatics. Methods in molecular biology. 2009;541:449-461.

170. Lievens S, Eyckerman S, Lemmens I, *et al.* Large-scale protein interactome mapping: strategies and opportunities. Expert review of proteomics. 2010;7(5):679-690.

171. Petschnigg J, Snider J, Stagljar I. Interactive proteomics research technologies: recent applications and advances. Curr Opin Biotechnol. 2011;22(1):50-58.

172. Cagney G, Uetz P. High-throughput screening for protein-protein interactions using yeast two-hybrid arrays. Curr Protoc Protein Sci. 2001;Chapter 19:Unit 19 16.

173. Berggard T, Linse S, James P. Methods for the detection and analysis of protein-protein interactions. Proteomics. 2007;7(16):2833-2842.

174. Carneiro DG, Clarke T, Davies CC, *et al.* Identifying novel protein interactions: Proteomic methods, optimisation approaches and data analysis pipelines. Methods. 2015.

175. Rao VS, Srinivas K, Sujini GN, *et al.* Protein-protein interaction detection: methods and analysis. Int J Proteomics. 2014;2014:147648.

176. Feng S, Zhou L, Huang C, *et al.* Interactomics: toward protein function and regulation. Expert review of proteomics. 2015;12(1):37-60.

177. Stynen B, Tournu H, Tavernier J, *et al.* Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. Microbiol Mol Biol Rev. 2012;76(2):331-382.

178. Barrios-Rodiles M, Brown KR, Ozdamar B, *et al.* High-throughput mapping of a dynamic signaling network in mammalian cells. Science. 2005;307(5715):1621-1625.

179. Morris JH, Knudsen GM, Verschueren E, *et al.* Affinity purification-mass spectrometry and network analysis to understand protein-protein interactions. Nature protocols. 2014;9(11):2539-2554.

180. Van Oudenhove L, Devreese B. A review on recent developments in mass spectrometry instrumentation and quantitative tools advancing bacterial proteomics. Appl Microbiol Biotechnol. 2013;97(11):4749-4762.

181. Tate S, Larsen B, Bonner R, *et al.* Label-free quantitative proteomics trends for protein-protein interactions. Journal of proteomics. 2013;81:91-101.

182. Basak T, Bhat A, Malakar D, *et al.* In-depth comparative proteomic analysis of yeast proteome using iTRAQ and SWATH based MS. Mol Biosyst. 2015;11(8):2135-2143.

183. Anjo SI, Santa C, Manadas B. Short GeLC-SWATH: a fast and reliable quantitative approach for proteomic screenings. Proteomics. 2015;15(4):757-762.

184. Tang WH, Shilov IV, Seymour SL. Nonlinear fitting method for determining local false discovery rates from decoy database searches. Journal of proteome research. 2008;7(9):3661-3667.

185. Sennels L, Bukowski-Wills JC, Rappsilber J. Improved results in proteomics by use of local and peptide-class specific false discovery rates. BMC bioinformatics. 2009;10:179.

186. Gillet LC, Navarro P, Tate S, *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Molecular & cellular proteomics : MCP. 2012;11(6):O111 016717.

187. Collins BC, Gillet LC, Rosenberger G, *et al.* Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. Nature methods. 2013;10(12):1246-1253.

188. Hollingsworth EB, McNeal ET, Burton JL, *et al.* Biochemical characterization of a filtered synaptoneurosome preparation from guinea pig cerebral cortex: cyclic adenosine 3':5'-monophosphate-generating systems, receptors, and enzymes. The Journal of neuroscience : the official journal of the Society for Neuroscience. 1985;5(8):2240-2253.

189. Titulaer MN, Ghijsen WE. Synaptoneurosomes. A preparation for studying subhippocampal GABAA receptor activity. Methods in molecular biology. 1997;72:49-59.

190. Westmark PR, Westmark CJ, Jeevananthan A, *et al.* Preparation of synaptoneurosomes from mouse cortex using a discontinuous percoll-sucrose density gradient. Journal of visualized experiments : JoVE. 2011(55).

191. Johnson MW, Chotiner JK, Watson JB. Isolation and characterization of synaptoneurosomes from single rat hippocampal slices. Journal of neuroscience methods. 1997;77(2):151-156.

192. Montojo J, Zuberi K, Rodriguez H, *et al.* GeneMANIA: Fast gene network construction and function prediction for Cytoscape. F1000Res. 2014;3:153.

193. Eden E, Navon R, Steinfeld I, *et al.* GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC bioinformatics. 2009;10:48.

194. Bennett EJ, Rush J, Gygi SP, *et al.* Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. Cell. 2010;143(6):951-965.

195. Kim SH, Kim HJ, Kim S, *et al.* Drosophila Cand1 regulates Cullin3-dependent E3 ligases by affecting the neddylation of Cullin3 and by controlling the stability of Cullin3 and adaptor protein. Developmental biology. 2010;346(2):247-257.

196. Wu C, Ma MH, Brown KR, *et al.* Systematic identification of SH3 domain-mediated human protein-protein interactions by peptide array target screening. Proteomics. 2007;7(11):1775-1785.

197. Trepanier CH, Jackson MF, MacDonald JF. Regulation of NMDA receptors by the tyrosine kinase Fyn. FEBS J. 2012;279(1):12-19.

198. Morris G, Berk M. The many roads to mitochondrial dysfunction in neuroimmune and neuropsychiatric disorders. BMC Med. 2015;13:68.

199. Lord LD, Expert P, Huckins JF, *et al.* Cerebral energy metabolism and the brain's functional network architecture: an integrative review. J Cereb Blood Flow Metab. 2013;33(9):1347-1354.

200. Toker L, Agam G. Mitochondrial dysfunction in psychiatric morbidity: current evidence and therapeutic prospects. Neuropsychiatr Dis Treat. 2015;11:2441-2447.

201. Manji H, Kato T, Di Prospero NA, *et al.* Impaired mitochondrial function in psychiatric disorders. Nature reviews Neuroscience. 2012;13(5):293-307.

202. Belanger M, Allaman I, Magistretti PJ. Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. Cell Metab. 2011;14(6):724-738.

203. Magistretti PJ, Allaman I. A cellular perspective on brain energy metabolism and functional imaging. Neuron. 2015;86(4):883-901.

204. Atkin T, MacAskill A, Kittler J. Neuronal Mitochondrial Transport and Dysfunction. In Reeve AK, Krishnan KJ, Duchen MR, *et al.*, (Eds). Mitochondrial Dysfunction in Neurodegenerative Disorders: Springer London 2012:157-173.

205. Streck EL, Goncalves CL, Furlanetto CB, *et al.* Mitochondria and the central nervous system: searching for a pathophysiological basis of psychiatric disorders. Revista brasileira de psiquiatria. 2014;36(2):156-167.

206. Hroudova J, Fisar Z. Connectivity between mitochondrial functions and psychiatric disorders. Psychiatry Clin Neurosci. 2011;65(2):130-141.

207. Kuloglu M, Atmaca M, Tezcan E, *et al.* Antioxidant enzyme activities and malondialdehyde levels in patients with obsessive-compulsive disorder. Neuropsychobiology. 2002;46(1):27-32.

208. Ersan S, Bakir S, Erdal Ersan E, *et al.* Examination of free radical metabolism and antioxidant defence system elements in patients with obsessive-compulsive disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2006;30(6):1039-1042.

209. Ozdemir E, Cetinkaya S, Ersan S, *et al.* Serum selenium and plasma malondialdehyde levels and antioxidant enzyme activities in patients with obsessive-compulsive disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2009;33(1):62-65.

210. Behl A, Swami G, Sircar SS, *et al.* Relationship of possible stress-related biochemical markers to oxidative/antioxidative status in obsessive-compulsive disorder. Neuropsychobiology. 2010;61(4):210-214.

211. Goncalves VF, Andreazza AC, Kennedy JL. Mitochondrial dysfunction in schizophrenia: an evolutionary perspective. Hum Genet. 2015;134(1):13-21.

212. Hill K, Mann L, Laws KR, *et al.* Hypofrontality in schizophrenia: a meta-analysis of functional imaging studies. Acta psychiatrica Scandinavica. 2004;110(4):243-256.

213. Martins-de-Souza D, Maccarrone G, Wobrock T, *et al.* Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. Journal of psychiatric research. 2010;44(16):1176-1189.

214. Prabakaran S, Swatton JE, Ryan MM, *et al.* Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. Molecular psychiatry. 2004;9(7):684-697, 643.

215. Kung L, Roberts RC. Mitochondrial pathology in human schizophrenic striatum: a postmortem ultrastructural study. Synapse. 1999;31(1):67-75.

216. Martins-de-Souza D, Gattaz WF, Schmitt A, *et al.* Alterations in oligodendrocyte proteins, calcium homeostasis and new potential markers in schizophrenia anterior temporal lobe are revealed by shotgun proteome analysis. Journal of neural transmission. 2009;116(3):275-289.

217. Bubber P, Hartounian V, Gibson GE, *et al.* Abnormalities in the tricarboxylic acid (TCA) cycle in the brains of schizophrenia patients. Eur Neuropsychopharmacol. 2011;21(3):254-260.

218. Middleton FA, Mirnics K, Pierri JN, *et al.* Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2002;22(7):2718-2729.

219. de Oliveira L, Fraga DB, De Luca RD, *et al.* Behavioral changes and mitochondrial dysfunction in a rat model of schizophrenia induced by ketamine. Metab Brain Dis. 2011;26(1):69-77.

220. Frey BN, Valvassori SS, Gomes KM, *et al.* Increased oxidative stress in submitochondrial particles after chronic amphetamine exposure. Brain Res. 2006;1097(1):224-229.

221. Martins-de-Souza D. Proteomics as a tool for understanding schizophrenia. Clin Psychopharmacol Neurosci. 2011;9(3):95-101.

6. Supplementary data

```
   1  ATGCGTGGCT ATCATGGTGA CCGTGGCTCG CACCCGCGTC CGGCCCGTTT TGCAGACCAG CAGCACATGG ATGTCGGCCC GGCAGCGCGC GCCCCGTATC
      TACGCACCGA TAGTACCACT GGCACCGAGC GTGGGCGCAG GCCGGGCAAA ACGTCTGGTC GTCGTGTACC TACAGCCGGG CCGTCGCGCG CGGGGCATAG
 101  TGCTGGGTTC GCGTGAAGCA TTTAGCACCG AACCGCGCTT CTGTGCACCG CGCGCTGGTC TGGGCCATAT TTCCCCGGAA GGCCCGCTGT CGCTGAGCGA
      ACGACCCAAG CGCACTTCGT AAATCGTGGC TTGGCGCGAA GACACGTGGC GCGCGACCAG ACCCGGTATA AAGGGGCCTT CCGGGCGACA GCGACTCGCT
 201  AGGTCCGTCA GTGGGCCCGG AAGGCGGTCC GGCCGGTGCC GGCGTTGGCG GTGGCAGCTC TACCTTTCCG CGTATGTATC CGGGTCAGGG TCCGTTTGAC
      TCCAGGCAGT CACCCGGGCC TTCCGCCAGG CCGGCCACGG CCGCAACCGC CACCGTCGAG ATGGAAAGGC GCATACATAG GCCCAGTCCC AGGCAAACTG
 301  ACCTGCGAAG ATTGTGTTGG TCATCCGCAG GGTAAAGGCG CCCCGCGTCT GCCGCCGACC CTGCTGGACC AGTTTGAAAA GCAACTGCCG GTTCAGCAAG
      TGGACGCTTC TAACACAACC AGTAGGCGTC CCATTTCCGC GGGGCGCAGA CGGCGGCTGG GACGACCTGG TCAAACTTTT CGTTGACGGC CAAGTCGTTC
 401  ATGGTTTCCA CACGCTGCCG TATCAGCGTG GTCCGGCAGG TGCTGGTCCG GGTCCGGCTC CGGGCACCGG TACGGCGCCG GAACCGCGTT CTGAAAGTCC
      TACCAAAGGT GTGCGACGGC ATAGTCGCAC CAGGCCGTCC ACGACCAGGC CCAGGCCGAG GCCCGTGGCC ATGCCGCGGC CTTGGCGCAA GACTTTCAGG
 501  GTCCCGTATC CGCCATCTGG TGCACAGCGT TCAGAAACTG TTTGCCAAGT CACATTCGCT GGAAGCACCG GGCAAGCGCG ACTACAACGG TCCGAAAGCG
      CAGGGCATAG GCGGTAGACC ACGTGTCGCA AGTCTTTGAC AAACGGTTCA GTGTAAGCGA CCTTCGTGGC CCGTTCGCGC TGATGTTGCC AGGCTTTCGC
 601  GAAGGCCGTG GTGGCTCTGG TGGCGATAGT TATCCGGGTC CGGGTTCCGG TGGCCCGCAC ACCTCACATC ACCACCATCA TCACCACCAC CACCACCATC
      CTTCCGGCAC CACCGAGACC ACCGCTATCA ATAGGCCCAG GCCCAAGGCC ACCGGGCGTG TGGAGTGTAG TGGTGGTAGT AGTGGTGGTG GTGGTGGTAG
 701  ACCAGTCTCG TCATGGTAAA CGCAGCAAAT CTAAGGACCG TAAAGGCGAT GGTCGCCACC AGGCGAAGAG TACCGGTTGG TGGAGTTCCG ATGACAATCT
      TGGTCAGAGC AGTACCATTT GCGTCGTTTA GATTCCTGGC ATTTCCGCTA CCAGCGGTGG TCCGCTTCTC ATGGCCAACC ACCTCAAGGC TACTGTTAGA
 801  GGACAGTGAT TCCGGTTTTC TGGCCGGTGG CCGCCCGCCG GGTGAACCGG GTGGCCCGTT CTGCCTGGAA GGCCCGGATG GCTCTTACCG TGATCTGAGT
      CCTGTCACTA AGGCCAAAAG ACCGGCCACC GGCGGGCGGC CCACTTGGCC CACCGGGCAA GACGGACCTT CCGGGCCTAC CGAGAATGGC ACTAGACTCA
 901  TTTAAAGGTC GCAGCGGTGG CTCTGAAGGC CGTTGCCTGG CGTGTACCGG CATGTCAATG TCGCTGGACG GCCAGTCCGT TAAACGCTCA GCTTGGCATA
      AAATTTCCAG CGTCGCCACC GAGACTTCCG GCAACGGACC GCACATGGCC GTACAGTTAC AGCGACCTGC CGGTCAGGCA ATTTGCGAGT CGAACCGTAT
1001  CGATGATGGT CAGCCAGGGT CGTGATGGTT ATCCGGGTGC CGGTCCGGGT AAAGGTCTGC TGGGCCCGGA AACCAAAGCG AAGGCCCGCA CGTATCACTA
      GCTACTACCA GTCGGTCCCA GCACTACCAA TAGGCCCACG GCCAGGCCCA TTTCCAGACG ACCCGGGCCT TTGGTTTCGC TTCCGGGCGT GCATAGTGAT
1101  CCTGCAGGTG CCGCAAGATG ACTGGGGTGG CTACCCGACC GGTGGCAAAG ATGGCGAAAT TCCGTGTCGT CGCATGCGTA GCGGTTCTTA TATCAAAGCG
      GGACGTCCAC GGCGTTCTAC TGACCCCACC GATGGGCTGG CCACCGTTTC TACCGCTTTA AGGCACAGCA GCGTACGCAT CGCCAAGAAT ATAGTTTCGC
1201  ATGGGCGATG AAGAAAGTGG TGACTCCGAT GGCTCACCGA AAACGAGCCC GAAGGCAGTT GCTCGTCGCT TCACCACGCG TCGCTCATCG AGCGTCGATC
      TACCCGCTAC TTCTTTCACC ACTGAGGCTA CCGAGTGGCT TTTGCTCGGG CTTCCGTCAA CGAGCAGCGA AGTGGTGCGC AGCGAGTAGC TCGCAGCTAG
1301  AGGCACGCAT TAACTGCTGT GTGCCGCCGC GTATTCATCC GCGCTCTAGT ATCCCGGGTT ACAGTCGTTC CCTGACCACG GGCCAGCTGA GCGATGAACT
      TCCGTGCGTA ATTGACGACA CACGGCGGCG CATAAGTAGG CGCGAGATCA TAGGGCCCAA TGTCAGCAAG GGACTGGTGC CCGGTCGACT CGCTACTTGA
1401  GAATCAGCAA CTGGAAGCCG TCTGCGGTTC CGTGTTCGGC GAACTGGAAT CACAGGCCGT GGACGCCCTG GATCTGCCGG GTTGTTTTCG TATGCGCTCA
      CTTAGTCGTT GACCTTCGGC AGACGCCAAG GCACAAGCCG CTTGACCTTA GTGTCCGGCA CCTGCGGGAC CTAGACGGCC CAACAAAAGC ATACGCGAGT
1501  CACTCGTATC TGCGCGCAAT TCAGGCTGGC TGCAGCCAAG ATGACGATTG TCTGCCGCTG CTGGCGACCC CGGCCGCAGT TAGCGGTCGC CCGGGCTCCT
      GTGAGCATAG ACGCGCGTTA AGTCCGACCG ACGTCGGTTC TACTGCTAAC AGACGGCGAC GACCGCTGGG GCCGGCGTCA ATCGCCAGCG GGCCCGAGGA
1601  CATTTAACTT CCGTAAAGCA CCGCCGCCGA TCCCGCCGGG CTCGCAGGCA CCGCCGCGTA TTAGCATCAC CGCTCAATCG AGCACGGATT CGGCGCATGA
      GTAAATTGAA GGCATTTCGT GGCGGCGGCT AGGGCGGCCC GAGCGTCCGT GGCGGCGCAT AATCGTAGTG GCGAGTTAGC TCGTGCCTAA GCCGCGTACT
1701  ATCTTTTACC GCAGCTGAAG GTCCGGCCCG TCGCTGCTCT AGTGCAGATG GTCTGGATGG TCCGGCGATG GGCGCACGCA CGCTGGAACT GGCTCCGGTG
      TAGAAAATGG CGTCGACTTC CAGGCCGGGC AGCGACGAGA TCACGTCTAC CAGACCTACC AGGCCGCTAC CCGCGTGCGT GCGACCTTGA CCGAGGCCAC
1801  CCGCCGCGTG CGTCTCCGAA GCCGCCGACC CTGATTATCA AAACGATTCC GGGTCGTGAA GAACTGCGCT CTCTGGCGCG TCAGCGCAAA TGGCGCCCGA
      GGCGGCGCAC GCAGAGGCTT CGGCGGCTGG GACTAATAGT TTTGCTAAGG CCCAGCACTT CTTGACGCGA GAGACCGCGC AGTCGCGTTT ACCGCGGGCT
1901  GTATTGGCGT GCAAGTTGAA ACCATCAGCG ACTCTGATAC GGAAAACCGT TCGCGTCGCG AATTTCACAG CATCGGTGTC CAGGTGGAAG AAGATAAGCG
      CATAACCGCA CGTTCAACTT TGGTAGTCGC TGAGACTATG CCTTTTGGCA AGCGCAGCGC TTAAAGTGTC GTAGCCACAG GTCCACCTTC TTCTATTCGC
2001  TCGCGCACGT TTTAAACGCA GTAATTCCGT TACCGCCGGC GTCCAGGCAG ACCTGGAACT GGAAGGTCTG GCAGGCCTGG CTACCGTGGC CACCGAAGAT
      AGCGCGTGCA AAATTTGCGT CATTAAGGCA ATGGCGGCCG CAGGTCCGTC TGGACCTTGA CCTTCCAGAC CGTCCGGACC GATGGCACCG GTGGCTTCTA
2101  AAAGCGCTGC AGTTTGGTCG TTCTTTCCAA CGCCATGCTA GTGAACCGCA GCCGGGTCCG CGTGCCCCGA CCTACAGCGT GTTCCGTACC GTTCACACGC
      TTTCGCGACG TCAAACCAGC AAGAAAGGTT GCGGTACGAT CACTTGGCGT CGGCCCAGGC GCACGGGGCT GGATGTCGCA CAAGGCATGG CAAGTGTGCG
2201  AGGGTCAATG GGCCTATCGC GAAGGCTACC CGCTGCCGTA TGAACCGCCG GCAACCGATG GTTCTCCGGG TCCGGCCCCG GCACCGACGC CGGGTCCGGG
      TCCCAGTTAC CCGGATAGCG CTTCCGATGG GCGACGGCAT ACTTGGCGGC CGTTGGCTAC CAAGAGGCCC AGGCCGGGGC CGTGGCTGCG GCCCAGGCCC
2301  TGCGGGCCGT CGCGATAGTT GGATTGAACG TGGTTCACGC TCGCTGCCGG ACTCTGGCCG TGCCAGTCCG TGCCCGCGTG ATGGTGAATG GTTTATCAAA
      ACGCCCGGCA GCGCTATCAA CCTAACTTGC ACCAAGTGCG AGCGACGGCC TGAGACCGGC ACGGTCAGGC ACGGGCGCAC TACCACTTAC CAAATAGTTT
2401  ATGCTGCGCG CAGAAGTGGA AAAGCTGGAA CATTGGTGTC AGCAAATGGA ACGTGAAGCT GAAGATTATG AACTGCCGGA AGAAATTCTG GAAAAAATCC
      TACGACGCGC GTCTTCACCT TTTCGACCTT GTAACCACAG TCGTTTACCT TGCACTTCGA CTTCTAATAC TTGACGGCCT TCTTTAAGAC CTTTTTTAGG
2501  GCTCCGCAGT CGGTTCAACC CAGCTGCTGC TGTCGCAAAA AGTGCAGCAA TTTTTCCGTC TGTGTCAGCA AAGCATGGAC CCGACCGCTT TTCCGGTTCC
      CGAGGCGTCA GCCAAGTTGG GTCGACGACG ACAGCGTTTT TCACGTCGTT AAAAAGGCAG ACACAGTCGT TTCGTACCTG GGCTGGCGAA AAGGCCAAGG
2601  GACCTTCCAG GACCTGGCCG GTTTCTGGGA TCTGCTGCAA CTGAGCATTG AAGATGTCAC CCTGAAATTT CTGGAACTGC AGCAACTGAA AGCCAATAGC
      CTGGAAGGTC CTGGACCGGC CAAAGACCCT AGACGACGTT GACTCGTAAC TTCTACAGTG GGACTTTAAA GACCTTGACG TCGTTGACTT TCGGTTATCG
2701  TGGAAGCTGC TGGAACCGAA AGAAGAAAAG AAAGTGCCGC CGCCGATCCC GAAAAAGCCG CTGCGTGGTC GCGGCGTTCC GGTCAAAGAA CGTTCGCTGG
      ACCTTCGACG ACCTTGGCTT TCTTCTTTTC TTTCACGGCG GCGGCTAGGG CTTTTTCGGC GACGCACCAG CGCCGCAAGG CCAGTTTCTT GCAAGCGACC
2801  ACAGCGTTGA TCGTCAGCGC CAAGAAGCGC GTAAACGTCT GCTGGCGGCA AAACGTGCAG CGAGCTTTCG CCATTCCTCA GCAACCGAAT CAGCGGACTC
      TGTCGCAACT AGCAGTCGCG GTTCTTCGCG CATTTGCAGA CGACCGCCGT TTTGCACGTC GCTCGAAAGC GGTAAGGAGT CGTTGGCTTA GTCGCCTGAG
2901  CATTGAAATC TACATCCCGG AAGCCCAAAC CCGTCTGTGA
      GTAACTTTAG ATGTAGGGCC TTCGGGTTTG GGCAGACACT
```

Supplementary Figure 6.1 – cDNA encoding for SAPAP3 with codons optimized for *E. coli*.

Table 6.1 – SWATH-MS method

Window	m/z range	Width (Da)	Window	m/z range	Width (Da)
1	349.5-363.1	13.6	**31**	613.7-622.8	9.1
2	362.1-378.4	16.3	**32**	621.8-631.3	9.5
3	377.4-392.8	15.4	**33**	630.3-639.4	9.1
4	391.8-405.4	13.6	**34**	638.4-647.5	9.1
5	404.4-417.1	12.7	**35**	646.5-655.6	9.1
6	416.1-427.4	11.3	**36**	654.6-663.7	9.1
7	426.4-437.3	10.9	**37**	662.7-672.7	10
8	436.3-446.8	10.5	**38**	671.7-681.7	10
9	445.8-456.2	10.4	**39**	680.7-690.7	10
10	455.2-465.2	10	**40**	689.7-700.2	10.5
11	464.2-474.2	10	**41**	699.2-709.6	10.4
12	473.2-482.8	9.6	**42**	708.6-719.5	10.9
13	481.8-490.4	8.6	**43**	718.5-729	10.5
14	489.4-498.1	8.7	**44**	728-738.4	10.4
15	497.1-505.3	8.2	**45**	737.4-748.3	10.9
16	504.3-512.5	8.2	**46**	747.3-758.2	10.9
17	511.5-519.7	8.2	**47**	757.2-768.6	11.4
18	518.7-526.4	7.7	**48**	767.6-780.3	12.7
19	525.4-533.2	7.8	**49**	779.3-792.4	13.1
20	532.2-540.4	8.2	**50**	791.4-806.8	15.4
21	539.4-547.1	7.7	**51**	805.8-822.1	16.3
22	546.1-554.4	8.3	**52**	821.1-838.8	17.7
23	553.4-561.6	8.2	**53**	837.8-858.1	20.3
24	560.6-568.8	8.2	**54**	857.1-879.3	22.2
25	567.8-576.4	8.6	**55**	878.3-903.1	24.8
26	575.4-584.1	8.7	**56**	902.1-931.9	29.8
27	583.1-591.7	8.6	**57**	930.9-968.8	37.9
28	590.7-599.4	8.7	**58**	967.8-1016.5	48.7
29	598.4-607	8.6	**59**	1015.5-1091.2	75.7
30	606-614.7	8.7	**60**	1090.2-1249.6	159.4

Table 6.2 – List of the proteins identified in the present study that constitute the spectral library for SWTAH-MS analysis

Uniprot Accession Number	Protein name	No identified peptides w/ 95 % confidence
P61765	Syntaxin-binding protein 1	65
P21575	Dynamin-1	62
Q9ER34	Aconitate hydratase, mitochondrial	64
P05708	Hexokinase-1	50
P06687	Sodium/potassium-transporting ATPase subunit alpha-3	51
P09951	Synapsin-1	60
Q62910	Synaptojanin-1	47
P62944	AP-2 complex subunit beta	43
P15999	ATP synthase subunit alpha, mitochondrial	51
P11980	Pyruvate kinase PKM	45
P05065	Fructose-bisphosphate aldolase A	50
Q2TA68	Dynamin-like 120 kDa protein, mitochondrial	38
P18484	AP-2 complex subunit alpha-2	40
Q5SGE0	Leucine-rich PPR motif-containing protein, mitochondrial	39
P16086	Spectrin alpha chain, non-erythrocytic 1	34
Q5XIN6	LETM1 and EF-hand domain-containing protein 1, mitochondrial	38
P11505	Plasma membrane calcium-transporting ATPase 1	33
P47942	Dihydropyrimidinase-related protein 2	42
Q9QUL6	Vesicle-fusing ATPase	37
P53534	Glycogen phosphorylase, brain form (Fragment)	38
P13233	2',3'-cyclic-nucleotide 3'-phosphodiesterase	42
P10860	Glutamate dehydrogenase 1, mitochondrial	40
Q5XI78	2-oxoglutarate dehydrogenase, mitochondrial	36
O35303	Dynamin-1-like protein	34
P63039	60 kDa heat shock protein, mitochondrial	35
P50554	4-aminobutyrate aminotransferase, mitochondrial	36
P04797	Glyceraldehyde-3-phosphate dehydrogenase	38
P13264	Glutaminase kidney isoform, mitochondrial	37
P0AEY0	Maltose-binding periplasmic protein	39
Q66HF1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	32
P52873	Pyruvate carboxylase, mitochondrial	31
P10719	ATP synthase subunit beta, mitochondrial	38
P63018	Heat shock cognate 71 kDa protein	30
Q6P6V0	Glucose-6-phosphate isomerase	31
Q5XHZ0	Heat shock protein 75 kDa, mitochondrial	30
000001		33
Q62950	Dihydropyrimidinase-related protein 1	29
P06685	Sodium/potassium-transporting ATPase subunit alpha-1	28
Q6P6R2	Dihydrolipoyl dehydrogenase, mitochondrial	28
Q9WVC0	Septin-7	27
P02688	Myelin basic protein	30
P35571	Glycerol-3-phosphate dehydrogenase, mitochondrial	24
Q63537	Synapsin-2	30
P14408	Fumarate hydratase, mitochondrial	26
Q05962	ADP/ATP translocase 1	29

P97685	Neurofascin	24
P62815	V-type proton ATPase subunit B, brain isoform	25
P48721	Stress-70 protein, mitochondrial	24
P25809	Creatine kinase U-type, mitochondrial	26
P42212	Green fluorescent protein	32
P09117	Fructose-bisphosphate aldolase C	25
P56574	Isocitrate dehydrogenase [NADP], mitochondrial	23
Q3KRE8	Tubulin beta-2B chain	25
P85834	Elongation factor Tu, mitochondrial	22
P13596	Neural cell adhesion molecule 1	23
P19511	ATP synthase F(0) complex subunit B1, mitochondrial	23
P32551	Cytochrome b-c1 complex subunit 2, mitochondrial	23
O08838	Amphiphysin	23
Q641Y2	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial	23
Q9JHU0	Dihydropyrimidinase-related protein 5	22
P68370	Tubulin alpha-1A chain	26
P11506	Plasma membrane calcium-transporting ATPase 2	21
Q63198	Contactin-1	20
P47709	Rabphilin-3A	21
Q68FX0	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	27
P51650	Succinate-semialdehyde dehydrogenase, mitochondrial	23
P05197	Elongation factor 2	21
P07335	Creatine kinase B-type	20
P15146	Microtubule-associated protein 2	19
P61265	Syntaxin-1B	21
P60711	Actin, cytoplasmic 1	24
Q63560	Microtubule-associated protein 6	20
Q5BK63	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial	20
Q02253	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	21
P21707	Synaptotagmin-1	19
Q5XIM9	T-complex protein 1 subunit beta	19
P97536	Cullin-associated NEDD8-dissociated protein 1	19
P63086	Mitogen-activated protein kinase 1	19
P09812	Glycogen phosphorylase, muscle form	18
P50137	Transketolase	18
Q68FY0	Cytochrome b-c1 complex subunit 1, mitochondrial	21
P31596	Excitatory amino acid transporter 2	21
O35077	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	20
Q9Z0W5	Protein kinase C and casein kinase substrate in neurons protein 1	18
P21396	Amine oxidase [flavin-containing] A	19
O08839	Myc box-dependent-interacting protein 1	18
P04636	Malate dehydrogenase, mitochondrial	19
P0C2X9	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	18
P17764	Acetyl-CoA acetyltransferase, mitochondrial	19
P56522	NADPH:adrenodoxin oxidoreductase, mitochondrial	18
O35179	Endophilin-A1	17
P62632	Elongation factor 1-alpha 2	20
P59215	Guanine nucleotide-binding protein G(o) subunit alpha	17

Q8R491	EH domain-containing protein 3	16
Q99NA5	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	20
Q62951	Dihydropyrimidinase-related protein 4 (Fragment)	16
Q64428	Trifunctional enzyme subunit alpha, mitochondrial	18
Q9JJM9	Septin-5	17
P09606	Glutamine synthetase	20
P06686	Sodium/potassium-transporting ATPase subunit alpha-2	16
O70351	3-hydroxyacyl-CoA dehydrogenase type-2	17
P11598	Protein disulfide-isomerase A3	18
P26284	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	16
Q05546	Tenascin-R	18
Q63028	Alpha-adducin	16
B3GNI6	Septin-11	17
Q5RKI0	WD repeat-containing protein 1	15
Q01986	Dual specificity mitogen-activated protein kinase kinase 1	16
P27139	Carbonic anhydrase 2	20
P35435	ATP synthase subunit gamma, mitochondrial	17
P00507	Aspartate aminotransferase, mitochondrial	16
Q9JHY2	Sideroflexin-3	16
P19332	Microtubule-associated protein tau	16
Q66HF8	Aldehyde dehydrogenase X, mitochondrial	17
Q3KR86	MIC	15
P46462	Transitional endoplasmic reticulum ATPase	16
P06761	78 kDa glucose-regulated protein	16
Q704S8	Carnitine O-acetyltransferase	15
P11507	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	16
P41565	Isocitrate dehydrogenase [NAD] subunit gamma 1, mitochondrial	16
Q9WU34	Neuronal-specific septin-3	17
Q02563	Synaptic vesicle glycoprotein 2A	16
P08461	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	15
Q4V7C7	Actin-related protein 3	15
P47860	ATP-dependent 6-phosphofructokinase, platelet type	17
Q9WTT6	Guanine deaminase	15
Q06647	ATP synthase subunit O, mitochondrial	17
P16638	ATP-citrate synthase	14
P19643	Amine oxidase [flavin-containing] B	15
P85845	Fascin	16
P84092	AP-2 complex subunit mu	15
Q05140	Clathrin coat assembly protein AP180	16
P85970	Actin-related protein 2/3 complex subunit 2	15
O70511	Ankyrin-3	14
P42123	L-lactate dehydrogenase B chain	14
P63329	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	15
P11442	Clathrin heavy chain 1	14
Q561S0	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial	15
P11884	Aldehyde dehydrogenase, mitochondrial	15
P13086	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	17

Q60587	Trifunctional enzyme subunit beta, mitochondrial	15
P34058	Heat shock protein HSP 90-beta	13
Q924S5	Lon protease homolog, mitochondrial	14
Q63633	Solute carrier family 12 member 5	14
P33124	Long-chain-fatty-acid--CoA ligase 6	13
P29147	D-beta-hydroxybutyrate dehydrogenase, mitochondrial	14
P31399	ATP synthase subunit d, mitochondrial	14
Q5FVI6	V-type proton ATPase subunit C 1	14
Q62717	Calcium-dependent secretion activator 1	12
P97700	Mitochondrial 2-oxoglutarate/malate carrier protein	14
Q6P502	T-complex protein 1 subunit gamma	15
P15650	Long-chain specific acyl-CoA dehydrogenase, mitochondrial	12
Q63065	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 1, mitochondrial	13
P67779	Prohibitin	14
Q64057	Alpha-aminoadipic semialdehyde dehydrogenase	14
Q499N5	Acyl-CoA synthetase family member 2, mitochondrial	15
P16036	Phosphate carrier protein, mitochondrial	15
P10888	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	15
P20788	Cytochrome b-c1 complex subunit Rieske, mitochondrial	15
Q62952	Dihydropyrimidinase-related protein 3	13
Q5FVQ8	NLR family member X1	13
P61107	Ras-related protein Rab-14	13
P07722	Myelin-associated glycoprotein	13
P12785	Fatty acid synthase	12
Q4V8B0	Oxidation resistance protein 1	12
P12007	Isovaleryl-CoA dehydrogenase, mitochondrial	13
Q3KRE0	ATPase family AAA domain-containing protein 3	14
P04177	Tyrosine 3-monooxygenase	12
Q63716	Peroxiredoxin-1	12
Q08877	Dynamin-3	12
Q5XIH7	Prohibitin-2	11
P14669	Annexin A3	13
Q9QYJ6	cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A	12
P56571	ES1 protein homolog, mitochondrial	12
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	14
P32851	Syntaxin-1A	12
Q6AY30	Saccharopine dehydrogenase-like oxidoreductase	13
P49432	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	12
P47858	ATP-dependent 6-phosphofructokinase, muscle type	12
Q64536	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 2, mitochondrial	12
Q5XIE6	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	12
Q5M7U6	Actin-related protein 2	12
Q5XIT1	Microtubule-associated protein RP/EB family member 3	12
Q01062	cGMP-dependent 3',5'-cyclic phosphodiesterase	12
Q920L2	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	12
P85969	Beta-soluble NSF attachment protein	12
Q9Z2F5	C-terminal-binding protein 1	12

P97546	Neuroplastin	11
Q9Z2L0	Voltage-dependent anion-selective channel protein 1	12
P04642	L-lactate dehydrogenase A chain	12
Q05683	Glutamate decarboxylase 2	11
Q9Z0V6	Thioredoxin-dependent peroxide reductase, mitochondrial	12
P14604	Enoyl-CoA hydratase, mitochondrial	14
P0C1X8	AP2-associated protein kinase 1	11
Q7TPB1	T-complex protein 1 subunit delta	11
Q91Y81	Septin-2	10
Q68FW7	Threonine--tRNA ligase, mitochondrial	10
Q794F9	4F2 cell-surface antigen heavy chain	11
P07340	Sodium/potassium-transporting ATPase subunit beta-1	13
Q99P39	Cysteine desulfurase, mitochondrial	12
Q09073	ADP/ATP translocase 2	11
Q8K4Y5	Leucine-rich glioma-inactivated protein 1	11
P54311	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	12
Q8VHF5	Citrate synthase, mitochondrial	10
P27605	Hypoxanthine-guanine phosphoribosyltransferase	11
Q4QRB4	Tubulin beta-3 chain	12
P46101	Dipeptidyl aminopeptidase-like protein 6	11
Q8CFD0	Sideroflexin-5	12
Q6IFW6	Keratin, type I cytoskeletal 10	11
B2GV06	Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	11
O88483	[Pyruvate dehydrogenase [acetyl-transferring]]-phosphatase 1, mitochondrial	9
P04182	Ornithine aminotransferase, mitochondrial	9
P19804	Nucleoside diphosphate kinase B	10
D4AAT7	ATP-dependent (S)-NAD(P)H-hydrate dehydratase	10
Q68FX9	NAD-dependent protein deacylase sirtuin-5, mitochondrial	10
Q9QZR6	Septin-9	10
P0DJJ3	SH3-containing GRB2-like protein 3-interacting protein 1	9
P60881	Synaptosomal-associated protein 25	10
P21708	Mitogen-activated protein kinase 3	10
Q6UPE1	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	10
P00388	NADPH--cytochrome P450 reductase	10
P85973	Purine nucleoside phosphorylase	10
P63102	14-3-3 protein zeta/delta	10
Q75Q39	Mitochondrial import receptor subunit TOM70	10
Q6PCU2	V-type proton ATPase subunit E 1	8
P11275	Calcium/calmodulin-dependent protein kinase type II subunit alpha	9
Q05695	Neural cell adhesion molecule L1	11
Q5EB81	NADH-cytochrome b5 reductase 1	9
P29411	GTP:AMP phosphotransferase AK3, mitochondrial	9
P52303	AP-1 complex subunit beta-1	9
P47819	Glial fibrillary acidic protein	10
P05712	Ras-related protein Rab-2A	11
Q62696	Disks large homolog 1	9
Q1HCL7	NAD kinase 2, mitochondrial	9
Q4V8F9	Hydroxysteroid dehydrogenase-like protein 2	9

P48500	Triosephosphate isomerase	10
P60203	Myelin proteolipid protein	13
Q08163	Adenylyl cyclase-associated protein 1	9
P31662	Sodium-dependent neutral amino acid transporter SLC6A17	8
P28480	T-complex protein 1 subunit alpha	7
P07895	Superoxide dismutase [Mn], mitochondrial	9
D3ZQG6	Tripartite motif-containing protein 2	9
P07633	Propionyl-CoA carboxylase beta chain, mitochondrial	9
P48768	Sodium/calcium exchanger 2	8
Q3MIE4	Synaptic vesicle membrane protein VAT-1 homolog	9
P23965	Enoyl-CoA delta isomerase 1, mitochondrial	8
Q9Z214	Homer protein homolog 1	8
Q8R431	Monoglyceride lipase	8
Q6Q0N1	Cytosolic non-specific dipeptidase	8
Q1WIM3	Cell adhesion molecule 3	9
P00406	Cytochrome c oxidase subunit 2	10
Q924N5	Long-chain-fatty-acid--CoA ligase ACSBG1	8
Q63564	Synaptic vesicle glycoprotein 2B	9
Q9JI66	Electrogenic sodium bicarbonate cotransporter 1	9
O88377	Phosphatidylinositol 5-phosphate 4-kinase type-2 beta	8
P41499	Tyrosine-protein phosphatase non-receptor type 11	9
P02091	Hemoglobin subunit beta-1	8
Q6PST4	Atlastin-1	8
Q5GFD9	Protein IMPACT	8
P51647	Retinal dehydrogenase 1	8
Q01205	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	9
P19234	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial	8
Q5XIF3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	7
P81155	Voltage-dependent anion-selective channel protein 2	9
P14882	Propionyl-CoA carboxylase alpha chain, mitochondrial	8
Q52KK3	Solute carrier family 25 member 51	7
P25286	V-type proton ATPase 116 kDa subunit a isoform 1	7
Q5M7W7	Probable proline--tRNA ligase, mitochondrial	8
Q9Z0J5	Thioredoxin reductase 2, mitochondrial	7
B2RYW9	Fumarylacetoacetate hydrolase domain-containing protein 2	7
Q6JE36	Protein NDRG1	8
Q5RJQ4	NAD-dependent protein deacetylase sirtuin-2	8
Q5XI79	NADH dehydrogenase [ubiquinone] complex I, assembly factor 7	7
Q6MGB5	Estradiol 17-beta-dehydrogenase 8	8
P97532	3-mercaptopyruvate sulfurtransferase	8
P35704	Peroxiredoxin-2	8
B0BNF1	Septin-8	8
Q9WU82	Catenin beta-1	8
P05370	Glucose-6-phosphate 1-dehydrogenase	6
Q91ZN1	Coronin-1A	7
Q9QX69	LanC-like protein 1	8
O35331	Pyridoxal kinase	7
Q4G064	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial	8

Q5I6B8	Phosphatidylinositol 4-phosphate 5-kinase type-1 gamma	8
P24942	Excitatory amino acid transporter 1	7
P84850	D-2-hydroxyglutarate dehydrogenase, mitochondrial	7
P60892	Ribose-phosphate pyrophosphokinase 1	6
Q68FQ0	T-complex protein 1 subunit epsilon	8
Q5M821	Protein phosphatase 1H	8
Q9Z270	Vesicle-associated membrane protein-associated protein A	8
Q5I0G4	Glycine--tRNA ligase (Fragment)	7
P31016	Disks large homolog 4	7
P01946	Hemoglobin subunit alpha-1/2	7
O08618	Phosphoribosyl pyrophosphate synthase-associated protein 2	7
O35796	Complement component 1 Q subcomponent-binding protein, mitochondrial	9
Q920F5	Malonyl-CoA decarboxylase, mitochondrial	8
Q5BJU7	Wiskott-Aldrich syndrome protein family member 1	7
Q704E8	ATP-binding cassette sub-family B member 7, mitochondrial	7
Q6AYS7	Aminoacylase-1A	7
P30835	ATP-dependent 6-phosphofructokinase, liver type	7
P11951	Cytochrome c oxidase subunit 6C-2	7
Q64568	Plasma membrane calcium-transporting ATPase 3	7
Q9JJ19	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	9
Q7TS56	Carbonyl reductase family member 4	7
P61206	ADP-ribosylation factor 3	7
Q6QIX3	Zinc transporter 3	7
P20651	Serine/threonine-protein phosphatase 2B catalytic subunit beta isoform	7
Q64591	2,4-dienoyl-CoA reductase, mitochondrial	7
Q9WVK7	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	6
O35180	Endophilin-A3	6
Q64542	Plasma membrane calcium-transporting ATPase 4	7
Q9QXQ0	Alpha-actinin-4	7
Q63345	Myelin-oligodendrocyte glycoprotein	6
O55012	Phosphatidylinositol-binding clathrin assembly protein	6
Q8K3P6	Calcium-binding mitochondrial carrier protein SCaMC-2	6
P49803	Regulator of G-protein signaling 7	6
P21913	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	6
Q63965	Sideroflexin-1	7
Q62847	Gamma-adducin	7
Q6PDU7	ATP synthase subunit g, mitochondrial	7
Q62915	Peripheral plasma membrane protein CASK	6
O88994	Mitochondrial amidoxime reducing component 2	6
P97686	Neuronal cell adhesion molecule	8
P12075	Cytochrome c oxidase subunit 5B, mitochondrial	8
P34926	Microtubule-associated protein 1A	6
P04764	Alpha-enolase	7
Q63362	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	6
Q3T1J1	Eukaryotic translation initiation factor 5A-1	6
Q5FVJ0	Protein RUFY3	7
Q9R063	Peroxiredoxin-5, mitochondrial	6
Q04400	Adenylate cyclase type 5	6

P24368	Peptidyl-prolyl cis-trans isomerase B	6
Q5XIF6	Tubulin alpha-4A chain	6
Q66HA8	Heat shock protein 105 kDa	6
P24329	Thiosulfate sulfurtransferase	5
Q62784	Type I inositol 3,4-bisphosphate 4-phosphatase	6
P69682	Adaptin ear-binding coat-associated protein 1	6
P47875	Cysteine and glycine-rich protein 1	6
P38650	Cytoplasmic dynein 1 heavy chain 1	6
P85515	Alpha-centractin	6
Q6P6Q2	Keratin, type II cytoskeletal 5	6
P82995	Heat shock protein HSP 90-alpha	5
P70619	Glutathione reductase (Fragment)	6
P32089	Tricarboxylate transport protein, mitochondrial	5
P27653	C-1-tetrahydrofolate synthase, cytoplasmic	5
Q5I0L3	Tyrosine--tRNA ligase, mitochondrial	6
Q505J9	ATPase family AAA domain-containing protein 1	5
Q4QQV3	Protein FAM162A	6
Q5PPN7	Coiled-coil domain-containing protein 51	6
O88588	Phosphofurin acidic cluster sorting protein 1	6
O88767	Protein DJ-1	6
Q9WTT2	Caseinolytic peptidase B protein homolog	7
P11960	2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial (Fragment)	6
P24473	Glutathione S-transferase kappa 1	5
Q6AYT7	Monoacylglycerol lipase ABHD12	5
Q64605	Receptor-type tyrosine-protein phosphatase S	5
Q62889	Neuroligin-3	6
Q5XI22	Acetyl-CoA acetyltransferase, cytosolic	6
Q5M7A7	CB1 cannabinoid receptor-interacting protein 1	5
Q07803	Elongation factor G, mitochondrial	5
P69897	Tubulin beta-5 chain	5
M0RC99	Ras-related protein Rab-5A	5
P62630	Elongation factor 1-alpha 1	5
Q2PQA9	Kinesin-1 heavy chain	5
Q6P799	Serine--tRNA ligase, cytoplasmic	5
Q5FVH2	Phospholipase D3	5
Q1WIM1	Cell adhesion molecule 4	6
Q05764	Beta-adducin	5
Q99PD4	Actin-related protein 2/3 complex subunit 1A	6
Q568Z9	Phytanoyl-CoA hydroxylase-interacting protein	4
P55161	Nck-associated protein 1	5
Q5RKI8	ATP-binding cassette sub-family B member 8, mitochondrial	5
Q6IMF3	Keratin, type II cytoskeletal 1	4
Q3B8Q0	Microtubule-associated protein RP/EB family member 2	5
Q5FVI3	Leucine-rich repeat-containing protein 57	5
Q5EB62	Solute carrier family 25 member 46	5
P31977	Ezrin	5
Q9Z250	Protein lin-7 homolog A	5
P62260	14-3-3 protein epsilon	6
P38718	Mitochondrial pyruvate carrier 2	5

D3ZFB6	Proline-rich transmembrane protein 2	7
Q62634	Vesicular glutamate transporter 1	5
Q99M64	Phosphatidylinositol 4-kinase type 2-alpha	5
Q6MG12	Uncharacterized protein C6orf136 homolog	4
P32232	Cystathionine beta-synthase	4
Q5I0K5	Mycophenolic acid acyl-glucuronide esterase, mitochondrial	4
Q62813	Limbic system-associated membrane protein	4
P0CC10	Leucine-rich repeat-containing protein 4B	4
Q8VHK2	Caskin-1	5
P19627	Guanine nucleotide-binding protein G(z) subunit alpha	4
Q64559	Cytosolic acyl coenzyme A thioester hydrolase	4
P63142	Potassium voltage-gated channel subfamily A member 2	4
Q68FT3	Pyridine nucleotide-disulfide oxidoreductase domain-containing protein 2	5
P60522	Gamma-aminobutyric acid receptor-associated protein-like 2	5
Q5M7W5	Microtubule-associated protein 4	4
P62882	Guanine nucleotide-binding protein subunit beta-5	5
P35434	ATP synthase subunit delta, mitochondrial	4
P08503	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	5
Q6AXT5	Ras-related protein Rab-21	4
P54313	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	4
O08662	Phosphatidylinositol 4-kinase alpha	5
P07936	Neuromodulin	4
Q63092	CaM kinase-like vesicle-associated protein	4
Q505J6	Mitochondrial glutamate carrier 2	4
P61983	14-3-3 protein gamma	4
P0C089	Phosphatidylglycerophosphatase and protein-tyrosine phosphatase 1	4
P63100	Calcineurin subunit B type 1	4
Q66HA6	ADP-ribosylation factor-like protein 8B	4
Q05982	Nucleoside diphosphate kinase A	5
Q9R1Z0	Voltage-dependent anion-selective channel protein 3	4
Q6AYG5	Ethylmalonyl-CoA decarboxylase	4
Q9WUS0	Adenylate kinase 4, mitochondrial	4
Q9R0I8	Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha	5
P35738	2-oxoisovalerate dehydrogenase subunit beta, mitochondrial	4
P18265	Glycogen synthase kinase-3 alpha	5
P62483	Voltage-gated potassium channel subunit beta-2	4
Q9Z1N4	3'(2'),5'-bisphosphate nucleotidase 1	5
P45592	Cofilin-1	4
P97576	GrpE protein homolog 1, mitochondrial	4
Q5XIM4	ATP synthase subunit s, mitochondrial	4
Q9QXY2	SRC kinase signaling inhibitor 1	4
Q9NQR8	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 4	4
Q6GQP4	Ras-related protein Rab-31	4
B0K020	CDGSH iron-sulfur domain-containing protein 1	4
Q4KLP0	Probable 2-oxoglutarate dehydrogenase E1 component DHKTD1, mitochondrial	4
Q5BJQ0	Chaperone activity of bc1 complex-like, mitochondrial	4
Q4V7D2	Protein rogdi homolog	3

P22062	Protein-L-isoaspartate(D-aspartate) O-methyltransferase	4
Q9WTP0	Band 4.1-like protein 1	4
P97849	Long-chain fatty acid transport protein 1	4
Q5XIT9	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	5
Q62844	Tyrosine-protein kinase Fyn	4
Q64548	Reticulon-1	5
O88339	Epsin-1	5
P52481	Adenylyl cyclase-associated protein 2	4
Q63377	Sodium/potassium-transporting ATPase subunit beta-3	4
P63331	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform	4
Q64560	Tripeptidyl-peptidase 2	3
P20069	Mitochondrial-processing peptidase subunit alpha	3
P62332	ADP-ribosylation factor 6	4
Q6AYQ8	Acylpyruvase FAHD1, mitochondrial	4
P13437	3-ketoacyl-CoA thiolase, mitochondrial	3
Q01728	Sodium/calcium exchanger 1	4
Q5RJL0	Ermin	3
Q63622	Disks large homolog 2	3
P62986	Ubiquitin-60S ribosomal protein L40	4
Q68FS4	Cytosol aminopeptidase	3
P04775	Sodium channel protein type 2 subunit alpha	4
O88778	Protein bassoon	4
O35964	Endophilin-A2	5
P63004	Platelet-activating factor acetylhydrolase IB subunit alpha	4
O35764	Neuronal pentraxin receptor	3
P31647	Sodium- and chloride-dependent GABA transporter 3	4
P0C5X8	Protein tweety homolog 1	4
P39052	Dynamin-2	3
P10760	Adenosylhomocysteinase	4
P13638	Sodium/potassium-transporting ATPase subunit beta-2	5
Q7TT47	Paraplegin	3
P97834	COP9 signalosome complex subunit 1	3
Q641Z6	EH domain-containing protein 1	4
Q9WVK3	Peroxisomal trans-2-enoyl-CoA reductase	3
Q63617	Hypoxia up-regulated protein 1	2
Q68FU7	Ubiquinone biosynthesis monooxygenase COQ6	3
O08776	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 3	3
Q63279	Keratin, type I cytoskeletal 19	4
Q6PEC0	Bis(5'-nucleosyl)-tetraphosphatase [asymmetrical]	3
Q5M9G9	Protein TBRG4	4
Q6AY19	AarF domain-containing protein kinase 4	3
O35095	Neurochondrin	3
P16617	Phosphoglycerate kinase 1	3
P04218	OX-2 membrane glycoprotein	4
B2RYT9	Translational activator of cytochrome c oxidase 1	3
P29419	ATP synthase subunit e, mitochondrial	3
D3ZAF6	ATP synthase subunit f, mitochondrial	3
P85968	6-phosphogluconate dehydrogenase, decarboxylating	3

Q6P9T8	Tubulin beta-4B chain	3
P68255	14-3-3 protein theta	3
P13852	Major prion protein	3
Q5M9I5	Cytochrome b-c1 complex subunit 6, mitochondrial	3
Q6AXW1	Glutaredoxin-2, mitochondrial	3
P11608	ATP synthase protein 8	3
O35094	Mitochondrial import inner membrane translocase subunit TIM44	4
P18088	Glutamate decarboxylase 1	3
Q03346	Mitochondrial-processing peptidase subunit beta	3
Q5U2X7	Mitochondrial import inner membrane translocase subunit Tim21	4
P41498	Low molecular weight phosphotyrosine protein phosphatase	3
B0BN86	Transmembrane protein 11, mitochondrial	3
Q9WU49	Calcium-regulated heat stable protein 1	3
P62744	AP-2 complex subunit sigma	3
P07825	Synaptophysin	3
O88917	Latrophilin-1	4
Q63270	Cytoplasmic aconitate hydratase	3
P23576	Gamma-aminobutyric acid receptor subunit alpha-2	3
P18266	Glycogen synthase kinase-3 beta	3
Q63372	Neurexin-1	4
P63045	Vesicle-associated membrane protein 2	3
P09216	Protein kinase C epsilon type	3
P35213	14-3-3 protein beta/alpha	3
O88954	MAGUK p55 subfamily member 3	2
Q66HR2	Microtubule-associated protein RP/EB family member 1	4
P36876	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	3
P68403	Protein kinase C beta type	4
A1L108	Actin-related protein 2/3 complex subunit 5-like protein	2
Q5BJP6	Ribosome-releasing factor 2, mitochondrial	3
Q9QY17	Protein kinase C and casein kinase substrate in neurons 2 protein	3
O08700	Vacuolar protein sorting-associated protein 45	2
Q6IRE4	Tumor susceptibility gene 101 protein	3
P35281	Ras-related protein Rab-10	2
Q63692	Hsp90 co-chaperone Cdc37	3
Q6P6Q9	Calcium uptake protein 1, mitochondrial	3
P26817	Beta-adrenergic receptor kinase 1	2
Q3ZB98	Breast carcinoma-amplified sequence 1 homolog (Fragment)	3
P23928	Alpha-crystallin B chain	3
Q5PPI4	Lysosome-associated membrane glycoprotein 5	3
P0C546	Mitochondrial coenzyme A transporter SLC25A42	3
P52296	Importin subunit beta-1	2
Q7TQ16	Cytochrome b-c1 complex subunit 8	3
Q6AY04	Uncharacterized protein C2orf47 homolog, mitochondrial	3
Q5EBA1	ATP-dependent RNA helicase SUPV3L1, mitochondrial	2
P26772	10 kDa heat shock protein, mitochondrial	2
Q4V8I7	Volume-regulated anion channel subunit LRRC8A	2
Q9WV97	Mitochondrial import inner membrane translocase subunit Tim9	3
O88989	Malate dehydrogenase, cytoplasmic	3
Q9R1T5	Aspartoacylase	2

Q4KM49	Tyrosine--tRNA ligase, cytoplasmic	2
Q9EQV6	Tripeptidyl-peptidase 1	3
Q5XII9	Mitochondrial fission regulator 1-like	2
P08413	Calcium/calmodulin-dependent protein kinase type II subunit beta	2
P12346	Serotransferrin	3
Q63081	Protein disulfide-isomerase A6	2
Q6AYR2	Protein NDRG3	3
Q68FR6	Elongation factor 1-gamma	3
P86252	Transcriptional activator protein Pur-alpha (Fragments)	2
O35264	Platelet-activating factor acetylhydrolase IB subunit beta	2
B5DF89	Cullin-3	2
P97846	Contactin-associated protein 1	1
A2VCX1	TIP41-like protein	2
P54287	Voltage-dependent L-type calcium channel subunit beta-3	2
Q925N3	PEX5-related protein	3
Q6PCU8	NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial	2
Q62888	Neuroligin-2	2
P23977	Sodium-dependent dopamine transporter	3
Q8K3V3	G-protein coupled receptor 56	2
Q99MZ8	LIM and SH3 domain protein 1	2
P11661	NADH-ubiquinone oxidoreductase chain 5	2
Q6XVN8	Microtubule-associated proteins 1A/1B light chain 3A	2
Q641Y8	ATP-dependent RNA helicase DDX1	3
P62142	Serine/threonine-protein phosphatase PP1-beta catalytic subunit	2
P52504	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial	3
Q9JLA3	UDP-glucose:glycoprotein glucosyltransferase 1	2
P97544	Lipid phosphate phosphohydrolase 3	3
Q8K4K5	Lethal(2) giant larvae protein homolog 1	2
P08644	GTPase KRas	2
A6JFQ6	Clavesin-1	2
F1LQX4	Rho GTPase-activating protein 44	2
Q62868	Rho-associated protein kinase 2	3
Q07205	Eukaryotic translation initiation factor 5	2
P08592	Amyloid beta A4 protein	2
Q8VHW5	Voltage-dependent calcium channel gamma-8 subunit	3
P08426	Cationic trypsin-3	3
D4A6D7	Tetratricopeptide repeat protein 19, mitochondrial	2
Q9Z2L9	Protein NDRG4	2
P62494	Ras-related protein Rab-11A	2
Q9JJW3	Up-regulated during skeletal muscle growth protein 5	3
Q6Q629	Inactive dipeptidyl peptidase 10	2
Q4QQT4	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform	2
P10688	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-1	2
P29418	ATP synthase subunit epsilon, mitochondrial	2
Q6IG02	Keratin, type II cytoskeletal 2 epidermal	3
Q4KM45	UPF0687 protein C20orf27 homolog	2
P05503	Cytochrome c oxidase subunit 1	2
P00762	Anionic trypsin-1	3

P50408	V-type proton ATPase subunit F	2
P27867	Sorbitol dehydrogenase	2
Q71S46	ATP synthase F(0) complex subunit C3, mitochondrial	2
P62083	40S ribosomal protein S7	2
P04631	Protein S100-B	2
P11517	Hemoglobin subunit beta-2	2
Q6EIX2	Mitochondrial import inner membrane translocase subunit TIM16	2
P35171	Cytochrome c oxidase subunit 7A2, mitochondrial	3
Q9JKE3	Secretory carrier-associated membrane protein 5	2
Q99P82	Claudin-11	2
Q9WTR7	Signal peptidase complex catalytic subunit SEC11C	2
Q1AAU6	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 1	1
Q91XR8	Phospholipid hydroperoxide glutathione peroxidase, nuclear	2
O88658	Kinesin-like protein KIF1B	2
B5DFN3	Ubiquinol-cytochrome-c reductase complex assembly factor 2	2
P62898	Cytochrome c, somatic	2
Q9QUH6	Ras/Rap GTPase-activating protein SynGAP	2
Q3KRD5	Mitochondrial import receptor subunit TOM34	2
Q99JD4	CLIP-associating protein 2	3
Q3T1K5	F-actin-capping protein subunit alpha-2	2
Q4KM98	Mitochondrial fission factor	2
P37285	Kinesin light chain 1	2
D3ZML2	Serine/threonine-protein kinase BRSK2	2
Q9WV63	Kinesin-like protein KIF2A	2
Q6P7R8	Estradiol 17-beta-dehydrogenase 12	2
Q3KRD0	Aspartate--tRNA ligase, mitochondrial	2
P49187	Mitogen-activated protein kinase 10	2
Q7TP47	Heterogeneous nuclear ribonucleoprotein Q	2
O70441	Synapsin-3	2
Q6AYN4	Phytanoyl-CoA hydroxylase-interacting protein-like	2
Q62936	Disks large homolog 3	2
P11915	Non-specific lipid-transfer protein	2
P0C5W1	Microtubule-associated protein 1S	2
P23565	Alpha-internexin	2
P04905	Glutathione S-transferase Mu 1	2
Q5FVI4	Cell cycle exit and neuronal differentiation protein 1	2
P80432	Cytochrome c oxidase subunit 7C, mitochondrial	2
P61203	COP9 signalosome complex subunit 2	2
Q5XIC2	Evolutionarily conserved signaling intermediate in Toll pathway, mitochondrial	2
Q63468	Phosphoribosyl pyrophosphate synthase-associated protein 1	2
D4A6L0	Probable G-protein coupled receptor 158	2
Q6RJR6	Reticulon-3	1
Q9Z1Z3	Epsin-2	2
P10818	Cytochrome c oxidase subunit 6A1, mitochondrial	2
Q02589	[Protein ADP-ribosylarginine] hydrolase	2
Q4V897	Coiled-coil domain-containing protein 90B, mitochondrial	2
P97519	Hydroxymethylglutaryl-CoA lyase, mitochondrial	2
Q9QZM5	Abl interactor 1	2

D3ZPX4	Plexin-A3	2
P25113	Phosphoglycerate mutase 1	2
Q68FU3	Electron transfer flavoprotein subunit beta	2
P63035	Cytohesin-2	2
B2RZ78	Vacuolar protein sorting-associated protein 29	2
P17425	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	2
O88370	Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma	2
D3ZN43	NADH dehydrogenase (ubiquinone) complex I, assembly factor 6	2
P17105	Inositol-trisphosphate 3-kinase A	2
O54701	Sodium/potassium/calcium exchanger 2	2
P12368	cAMP-dependent protein kinase type II-alpha regulatory subunit	1
Q641Z2	Tyrosine-protein phosphatase non-receptor type 9	2
B0BN94	Protein FAM136A	2
Q5U2U0	ATP-dependent Clp protease ATP-binding subunit clpX-like, mitochondrial	2
Q32PX9	Lactation elevated protein 1	2
B1WBW4	Armadillo repeat-containing protein 10	2
Q792I0	Protein lin-7 homolog C	2
Q5XIA8	Growth hormone-inducible transmembrane protein	2
Q64232	Very-long-chain enoyl-CoA reductase	2
P07153	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	2
Q4KLH5	Arf-GAP domain and FG repeat-containing protein 1	2
O89046	Coronin-1B	2
Q9QWN8	Spectrin beta chain, non-erythrocytic 2	1
Q3B7U9	Peptidyl-prolyl cis-trans isomerase FKBP8	2
P19139	Casein kinase II subunit alpha	2
F1LP90	Misshapen-like kinase 1	1
O88637	Ethanolamine-phosphate cytidylyltransferase	1
Q62609	Noelin	1
Q6PCT3	Tumor protein D54	1
P62024	Phosphatase and actin regulator 1	2
P19491	Glutamate receptor 2	2
O54975	Xaa-Pro aminopeptidase 1	1
Q06486	Casein kinase I isoform delta	1
P35280	Ras-related protein Rab-8A	1
Q9ERQ6	Chondroitin sulfate proteoglycan 5	2
P10111	Peptidyl-prolyl cis-trans isomerase A	1
Q5XIJ4	Protein FAM210A	1
Q6P7S1	Acid ceramidase	1
Q5PQL7	Integral membrane protein 2C	1
P10824	Guanine nucleotide-binding protein G(i) subunit alpha-1	1
P55068	Brevican core protein	2
Q562B5	Serine/threonine-protein phosphatase PGAM5, mitochondrial	1
Q5HZY2	GTP-binding protein SAR1b	1
Q66H15	Regulator of microtubule dynamics protein 3	1
Q5RJR2	Twinfilin-1	1
P63012	Ras-related protein Rab-3A	1
Q9WVJ4	Synaptojanin-2-binding protein	1
A1L1I3	Numb-like protein	1

Q91ZW1	Transcription factor A, mitochondrial	3
P19468	Glutamate--cysteine ligase catalytic subunit	1
Q4V8B7	Inactive hydroxysteroid dehydrogenase-like protein 1	1
Q812E9	Neuronal membrane glycoprotein M6-a	1
P14173	Aromatic-L-amino-acid decarboxylase	1
P11167	Solute carrier family 2, facilitated glucose transporter member 1	1
P55067	Neurocan core protein	2
P48303	Sphingosine 1-phosphate receptor 1	1
Q9JM53	Apoptosis-inducing factor 1, mitochondrial	1
P85108	Tubulin beta-2A chain	2
O35952	Hydroxyacylglutathione hydrolase, mitochondrial	2
Q8R500	Mitofusin-2	1
Q9Z142	Transmembrane protein 33	1
Q91ZW6	Trimethyllysine dioxygenase, mitochondrial	1
Q5M819	Phosphoserine phosphatase	1
Q4FZU2	Keratin, type II cytoskeletal 6A	1
P62870	Transcription elongation factor B polypeptide 2	1
P03889	NADH-ubiquinone oxidoreductase chain 1	1
P62074	Mitochondrial import inner membrane translocase subunit Tim10	1
Q5BKC9	Ephexin-1	1
Q4KLF8	Actin-related protein 2/3 complex subunit 5	1
Q4VSI4	Ubiquitin carboxyl-terminal hydrolase 7	1
Q6AYQ3	Phenylalanine--tRNA ligase, mitochondrial	1
P84060	Dystrobrevin beta	1
Q9Z269	Vesicle-associated membrane protein-associated protein B	1
Q62829	Serine/threonine-protein kinase PAK 3	1
Q5FWT5	Glutamyl-tRNA(Gln) amidotransferase subunit A, mitochondrial	1
P20650	Protein phosphatase 1A	1
Q66X93	Staphylococcal nuclease domain-containing protein 1	1
P18298	S-adenosylmethionine synthase isoform type-2	1
P05508	NADH-ubiquinone oxidoreductase chain 4	1
Q8VBU2	Protein NDRG2	1
Q99MC0	Protein phosphatase 1 regulatory subunit 14A	1
Q63159	Hexaprenyldihydroxybenzoate methyltransferase, mitochondrial	1
Q62753	Syntaxin-binding protein 2	1
Q66HD0	Endoplasmin	1
P47861	Synaptotagmin-5	1
Q8K4M9	Oxysterol-binding protein-related protein 1	1
Q7TSA0	Mitochondrial Rho GTPase 2	1
P45479	Palmitoyl-protein thioesterase 1	1
P45953	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	1
P68035	Actin, alpha cardiac muscle 1	1
Q6Q7Y5	Guanine nucleotide-binding protein subunit alpha-13	1
B2DD29	Serine/threonine-protein kinase BRSK1	1
P18508	Gamma-aminobutyric acid receptor subunit gamma-2	1
P63079	Gamma-aminobutyric acid receptor subunit beta-3	1
P00564	Creatine kinase M-type	1
P63319	Protein kinase C gamma type	1
Q5FWU3	Autophagy-related protein 9A	1

P97710	Tyrosine-protein phosphatase non-receptor type substrate 1	1
Q6QLM7	Kinesin heavy chain isoform 5A	1
P29457	Serpin H1	1
Q5XIL3	DNA-directed RNA polymerase III subunit RPC3	1
P21263	Nestin	1
Q925Q9	SH3 domain-containing kinase-binding protein 1	1
Q62656	Receptor-type tyrosine-protein phosphatase zeta	2
Q6AYH5	Dynactin subunit 2	1
O89049	Thioredoxin reductase 1, cytoplasmic	1
Q9JK15	Arf-GAP with dual PH domain-containing protein 2	1
Q5U1X1	Oligoribonuclease, mitochondrial	1
P00159	Cytochrome b	1
Q91XU8	Phosphatidate cytidylyltransferase 2	1
Q5PPJ9	Endophilin-B2	1
Q5BK62	Protein Mpv17	1
P27952	40S ribosomal protein S2	1
P80431	Cytochrome c oxidase subunit 7B, mitochondrial	1
P97538	Ras-related protein M-Ras	1
P61023	Calcineurin B homologous protein 1	1
P29266	3-hydroxyisobutyrate dehydrogenase, mitochondrial	1
Q924S1	1-acyl-sn-glycerol-3-phosphate acyltransferase delta	1
P62138	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	1
Q8VHV7	Heterogeneous nuclear ribonucleoprotein H	1
Q5U2U2	Crk-like protein	1
P30713	Glutathione S-transferase theta-2	1
A2RUW1	Toll-interacting protein	1
Q6AXQ5	2',5'-phosphodiesterase 12	1
P42667	Signal peptidase complex catalytic subunit SEC11A	1
P26453	Basigin	1
P04906	Glutathione S-transferase P	1
Q5U1Z7	Centromere protein R	1
P28572	Sodium- and chloride-dependent glycine transporter 1	1
D3ZKF5	Serine protease HTR4	1
O35353	Guanine nucleotide-binding protein subunit beta-4	1
P84039	Ectonucleotide pyrophosphatase/phosphodiesterase family member 5	1
Q9R1B1	Mitochondrial import inner membrane translocase subunit Tim10 B	1
Q62876	Synaptogyrin-1	1
P56603	Secretory carrier-associated membrane protein 1	1
D3Z9R8	6.8 kDa mitochondrial proteolipid	1
Q6AYQ4	Transmembrane protein 109	1
P62250	40S ribosomal protein S16	1
P11232	Thioredoxin	1
Q9JJK1	Neuronal membrane glycoprotein M6-b	1
Q6MGD0	Protein CutA	1
Q5FVL2	ER membrane protein complex subunit 8	1
P62919	60S ribosomal protein L8	1
Q6AY23	Pyrroline-5-carboxylate reductase 2	1
P07872	Peroxisomal acyl-coenzyme A oxidase 1	1

Q6PEC4	S-phase kinase-associated protein 1	1
Q5M7T2	SPRY domain-containing protein 7	1
Q9ERH3	WD repeat-containing protein 7	1
P97603	Neogenin (Fragment)	1
Q9Z272	ARF GTPase-activating protein GIT1	1
P08010	Glutathione S-transferase Mu 2	1
P08050	Gap junction alpha-1 protein	1
P05506	NADH-ubiquinone oxidoreductase chain 3	1
P60868	40S ribosomal protein S20	1
P31422	Metabotropic glutamate receptor 3	1
P15205	Microtubule-associated protein 1B	1
Q8R4A1	ERO1-like protein alpha	1
Q6P7B0	Tryptophan--tRNA ligase, cytoplasmic	1
O35824	DnaJ homolog subfamily A member 2	1
Q68FT1	Ubiquinone biosynthesis protein COQ9, mitochondrial	1
Q7TMB7	Lipid phosphate phosphatase-related protein type 4	2
Q5RJY4	Dehydrogenase/reductase SDR family member 7B	1
P56558	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	1
Q63060	Glycerol kinase	1
Q5XII0	Mammalian ependymin-related protein 1	1
Q6MG82	Proline-rich transmembrane protein 1	1

Table 6.3 – Fold-change of proteins showing significant differences between SAPAP3 WT and SAPAP3 K910R

Protein	SAPAP3 WT (mean intensities)	SAPAP3 K910R (mean intensities)	Fold-change log2
KPCE_RAT	0.0018	0.0006	-1.46
GFAP_RAT	0.0249	0.0119	-1.07
SYT5_RAT	0.0267	0.0141	-0.92
TBB2B_RAT	0.0028	0.0015	-0.92
CYH2_RAT	0.0129	0.0075	-0.78
NDUA5_RAT	0.0164	0.0101	-0.70
TBB4B_RAT	0.0480	0.0302	-0.67
IDHP_RAT	0.2544	0.1646	-0.63
TBA1A_RAT	0.4579	0.3025	-0.60
PI42A_RAT	0.0059	0.0039	-0.60
TBA4A_RAT	0.1710	0.1132	-0.60
ADDA_RAT	0.0237	0.0157	-0.60
ARPC2_RAT	0.0272	0.0186	-0.55
PDIA3_RAT	0.0855	0.0588	-0.54
ATAD1_RAT	0.0144	0.0099	-0.54
QCR2_RAT	0.4381	0.3077	-0.51
TBB5_RAT	0.0327	0.0233	-0.49
GPDA_RAT	0.0743	0.0533	-0.48
MP2K1_RAT	0.1703	0.1230	-0.47
NDRG4_RAT	0.0108	0.0078	-0.46
SFXN3_RAT	0.0770	0.0570	-0.43

GNAO_RAT	0.2815	0.2089	-0.43
HXK1_RAT	1.1346	0.8445	-0.43
TCPG_RAT	0.0660	0.0495	-0.42
QCR1_RAT	0.3633	0.2736	-0.41
ALDOA_RAT	1.3082	0.9894	-0.40
OXR1_RAT	0.0278	0.0211	-0.40
PACN1_RAT	0.1268	0.0967	-0.39
VATB2_RAT	0.1119	0.0855	-0.39
DYN1_RAT	0.3334	0.2556	-0.38
SFXN1_RAT	0.0319	0.0246	-0.38
NSF_RAT	0.2803	0.2180	-0.36
NDUA9_RAT	0.1687	0.1316	-0.36
ACTB_RAT	0.4546	0.3550	-0.36
PRIO_RAT	0.0026	0.0020	-0.35
AT5F1_RAT	0.9583	0.7584	-0.34
PP2BB_RAT	0.0316	0.0250	-0.34
NDUS2_RAT	0.2141	0.1729	-0.31
SSDH_RAT	0.1726	0.1420	-0.28
SEPT7_RAT	0.2192	01805	-0.28
NDUS4_RAT	0.0715	0.0589	-0.28
NPTN_RAT	0.0719	0.0619	-0.22

Table 6.4 – Fold-change of proteins showing significant differences between SAPAP3 WT and SAPAP3 R770L

Protein	SAPAP3 WT (mean intensities)	SAPAP3 R770L (mean intensities)	Fold-change log2
CISY	0.0723	0.0252	-1.52
NOE1	0.0037	0.0017	-1.11
MDHM	0.1416	0.0691	-1.03
ENPL	0.0055	0.0029	-0.94
NDUA5	0.0164	0.0089	-0.89
AATM	0.1443	0.0807	-0.84
HS90A	0.0170	0.0097	-0.81
S2551	0.0535	0.0314	-0.77
NDRG4	0.0108	0.0064	-0.76
KPCE	0.0018	0.0011	-0.75
USMG5	0.0195	0.0117	-0.74
1433Z	0.0139	0.0084	-0.73
CYH2	0.0129	0.0078	-0.72
CSKP	0.0042	0.0026	-0.71
TTYH1	0.0097	0.0060	-0.70
GRP78	0.0222	0.0137	-0.70
ANK3	0.0353	0.0219	-0.69
ODP2	0.0513	0.0319	-0.69
TOM70	0.0292	0.0183	-0.67

CX7A2	0.0149	0.0093	-0.67
1433G	0.0127	0.0080	-0.67
ARP5L	0.0357	0.0225	-0.66
NDUA9	0.1687	0.1069	-0.66
ATPB	2.6073	1.6545	-0.66
CADM3	0.0333	0.0213	-0.65
GUAD	0.0316	0.0203	-0.64
ADDA	0.0237	0.0152	-0.64
SEPT9	0.0107	0.0069	-0.63
ODBA	0.0067	0.0043	-0.63
ATPA	2.7122	1.7555	-0.63
NDUS1	0.3949	0.2556	-0.63
CH60	0.2389	0.1550	-0.62
NDUS4	0.0715	0.0465	-0.62
CLVS1	0.0044	0.0029	-0.62
NDUAA	0.1113	0.0729	-0.61
ARP3	0.0264	0.0173	-0.61
PIMT	0.0071	0.0046	-0.61
AL1B1	0.0605	0.0397	-0.61
KCRB	0.6000	0.3946	-0.60
COX41	0.8112	0.5342	-0.60
NDUS2	0.2141	0.1412	-0.60
OAT	0.0669	0.0443	-0.59
PP2BA	0.0560	0.0372	-0.59
SCPDL	0.2631	0.1767	-0.57
ATPO	0.7786	0.5238	-0.57
SYSC	0.0102	0.0069	-0.57
COQ6	0.0254	0.0171	-0.57
ORN	0.0043	0.0029	-0.57
HOME1	0.0082	0.0056	-0.56
HSP7C	0.5750	0.3901	-0.56
MMSA	0.1118	0.0758	-0.56
QCR1	0.3633	0.2471	-0.56
ODPB	0.1912	0.1302	-0.55
AT5F1	0.9583	0.6529	-0.55
AT1B3	0.0178	0.0122	-0.55
ALDOC	0.2369	0.1625	-0.54
LETM1	0.2223	0.1526	-0.54
PYGM	0.0550	0.0378	-0.54
IDHP	0.2544	0.1751	-0.54
PA1B2	0.0538	0.0371	-0.54
CAND1	0.0774	0.0536	-0.53
PDIA3	0.0855	0.0593	-0.53
ACTB	0.4546	0.3157	-0.53
ACON	2.3755	1.6512	-0.52
G6PI	0.2310	0.1606	-0.52

DPYL2	0.6065	0.4237	-0.52
THIL	0.1620	0.1132	-0.52
NDUV2	0.1161	0.0812	-0.52
SEP11	0.0969	0.0678	-0.51
NSF	0.2803	0.1965	-0.51
SUV3	0.0135	0.0095	-0.51
RAB5A	0.0192	0.0135	-0.51
GNAZ	0.0033	0.0023	-0.51
CF136	0.0283	0.0200	-0.50
ODO1	0.2463	0.1743	-0.50
TMM11	0.0296	0.0210	-0.49
DYN1	0.3334	0.2370	-0.49
MAP6	0.1003	0.0714	-0.49
HCD2	0.2220	0.1589	-0.48
DHE3	1.5022	1.0756	-0.48
ALDOA	1.3082	0.9370	-0.48
PRDX2	0.0956	0.0686	-0.48
IDH3A	0.2421	0.1736	-0.48
THTR	0.0293	0.0210	-0.48
SCOT1	0.0348	0.0250	-0.48
G3P	1.2236	0.8789	-0.48
GLSK	0.2100	0.1509	-0.48
OPA1	0.1836	0.1320	-0.48
AP2A2	0.4486	0.3229	-0.47
SYN2	0.5780	0.4172	-0.47
ATP5H	0.2456	0.1774	-0.47
AP2B1	0.2764	0.2000	-0.47
GPDM	0.0875	0.0635	-0.46
SEPT8	0.0410	0.0298	-0.46
QCR2	0.4381	0.3181	-0.46
AT1B1	0.3433	0.2495	-0.46
ACBG1	0.0611	0.0446	-0.46
LPPRC	0.1205	0.0879	-0.45
D2HDH	0.0227	0.0166	-0.45
CISD1	0.0125	0.0091	-0.45
PYGB	0.1546	0.1134	-0.45
LDHA	0.1899	0.1394	-0.45
DNM1L	0.1534	0.1129	-0.44
ETFD	0.0211	0.0157	-0.43
GNAO	0.2815	0.2091	-0.43
AMPH	0.1545	0.1155	-0.42
SSDH	0.1726	0.1295	-0.42
GSK3B	0.0121	0.0091	-0.41
GSHR	0.0272	0.0205	-0.41
ECHA	0.0726	0.0552	-0.39
ARF3	0.0196	0.0150	-0.39

MP2K1	0.1703	0.1303	-0.39
OXR1	0.0278	0.0213	-0.38
GBB5	0.0145	0.0112	-0.38
EF2	0.0231	0.0179	-0.37
GHC2	0.1950	0.1518	-0.36
PP2BB	0.0316	0.0247	-0.35
RUFY3	0.0040	0.0031	-0.35
TCPG	0.0660	0.0520	-0.34
ECHM	0.0740	0.0585	-0.34
PACN1	0.1268	0.1004	-0.34
DPYL1	0.0758	0.0600	-0.34
EFTU	0.2084	0.1743	-0.26
PTN9	0.0004	0.0023	2.43

Table 6.5 – List of the proteins identified for the WT SAPAP3 domain

Uniprot Accession Number	Protein Name	No identified peptides w/ 95 % conf
P61765	Syntaxin-binding protein 1	56
Q9ER34	Aconitate hydratase, mitochondrial	51
P21575	Dynamin-1	51
P05708	Hexokinase-1	42
P06687	Sodium/potassium-transporting ATPase subunit alpha-3	41
P09951	Synapsin-1	48
P11980	Pyruvate kinase PKM	40
P62944	AP-2 complex subunit beta	38
P15999	ATP synthase subunit alpha, mitochondrial	38
Q62910	Synaptojanin-1	36
P10860	Glutamate dehydrogenase 1, mitochondrial	34
P47942	Dihydropyrimidinase-related protein 2	37
P13233	2',3'-cyclic-nucleotide 3'-phosphodiesterase	33
P05065	Fructose-bisphosphate aldolase A	34
P11505	Plasma membrane calcium-transporting ATPase 1	30
P0AEY0	Maltose-binding periplasmic protein	33
P10719	ATP synthase subunit beta, mitochondrial	32
P18484	AP-2 complex subunit alpha-2	32
Q2TA68	Dynamin-like 120 kDa protein, mitochondrial	29
Q5XIN6	LETM1 and EF-hand domain-containing protein 1, mitochondrial	30
P04797	Glyceraldehyde-3-phosphate dehydrogenase	32
Q66HF1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	28
P50554	4-aminobutyrate aminotransferase, mitochondrial	27
P13264	Glutaminase kidney isoform, mitochondrial	29
Q9QUL6	Vesicle-fusing ATPase	23
P53534	Glycogen phosphorylase, brain form (Fragment)	26
Q6P6V0	Glucose-6-phosphate isomerase	26
P63039	60 kDa heat shock protein, mitochondrial	25
Q62950	Dihydropyrimidinase-related protein 1	24
P63018	Heat shock cognate 71 kDa protein	24

P06685	Sodium/potassium-transporting ATPase subunit alpha-1	23
Q63537	Synapsin-2	24
Q6P6R2	Dihydrolipoyl dehydrogenase, mitochondrial	24
Q5XI78	2-oxoglutarate dehydrogenase, mitochondrial	22
Q9WVC0	Septin-7	22
Q09073	ADP/ATP translocase 2	23
Q5SGE0	Leucine-rich PPR motif-containing protein, mitochondrial	22
P85108	Tubulin beta-2A chain	22
P25809	Creatine kinase U-type, mitochondrial	24
P56574	Isocitrate dehydrogenase [NADP], mitochondrial	21
P14408	Fumarate hydratase, mitochondrial	20
P63259	Actin, cytoplasmic 2	19
Q6P9V9	Tubulin alpha-1B chain	20
Q68FY0	Cytochrome b-c1 complex subunit 1, mitochondrial	20
P32551	Cytochrome b-c1 complex subunit 2, mitochondrial	19
Q641Y2	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial	19
P07335	Creatine kinase B-type	19
P85834	Elongation factor Tu, mitochondrial	19
000001		29
Q5XHZ0	Heat shock protein 75 kDa, mitochondrial	19
P04636	Malate dehydrogenase, mitochondrial	18
P19511	ATP synthase F(0) complex subunit B1, mitochondrial	18
P42212	Green fluorescent protein	26
P09117	Fructose-bisphosphate aldolase C	19
P02688	Myelin basic protein	27
O35303	Dynamin-1-like protein	18
Q9JHU0	Dihydropyrimidinase-related protein 5	17
P21707	Synaptotagmin-1	17
P21396	Amine oxidase [flavin-containing] A	17
P48721	Stress-70 protein, mitochondrial	17
P62815	V-type proton ATPase subunit B, brain isoform	16
P35571	Glycerol-3-phosphate dehydrogenase, mitochondrial	16
Q68FX0	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	16
O08838	Amphiphysin	15
P51650	Succinate-semialdehyde dehydrogenase, mitochondrial	16
P61265	Syntaxin-1B	17
O70351	3-hydroxyacyl-CoA dehydrogenase type-2	16
P50137	Transketolase	16
Q99NA5	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	16
P31596	Excitatory amino acid transporter 2	17
Q5BK63	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial	14
P97685	Neurofascin	16
P05197	Elongation factor 2	14
Q06647	ATP synthase subunit O, mitochondrial	18
P19332	Microtubule-associated protein tau	15
Q66HF8	Aldehyde dehydrogenase X, mitochondrial	14
P11506	Plasma membrane calcium-transporting ATPase 2	14
P97700	Mitochondrial 2-oxoglutarate/malate carrier protein	14

P13596	Neural cell adhesion molecule 1	15
Q02253	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	14
Q63198	Contactin-1	14
P13086	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	14
P27139	Carbonic anhydrase 2	16
Q3KR86	MIC	13
B3GNI6	Septin-11	13
P41565	Isocitrate dehydrogenase [NAD] subunit gamma 1, mitochondrial	14
Q9Z0W5	Protein kinase C and casein kinase substrate in neurons protein 1	12
P26284	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	15
Q9JJM9	Septin-5	13
P52873	Pyruvate carboxylase, mitochondrial	11
Q01986	Dual specificity mitogen-activated protein kinase kinase 1	13
P35435	ATP synthase subunit gamma, mitochondrial	14
P00507	Aspartate aminotransferase, mitochondrial	14
O08839	Myc box-dependent-interacting protein 1	12
Q63560	Microtubule-associated protein 6	12
P20788	Cytochrome b-c1 complex subunit Rieske, mitochondrial	14
Q561S0	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial	13
P20651	Serine/threonine-protein phosphatase 2B catalytic subunit beta isoform	13
Q5RKI0	WD repeat-containing protein 1	13
Q62951	Dihydropyrimidinase-related protein 4 (Fragment)	13
P84092	AP-2 complex subunit mu	12
P67779	Prohibitin	12
O35077	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	12
Q64536	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 2, mitochondrial	11
P31399	ATP synthase subunit d, mitochondrial	11
P10888	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	13
O35179	Endophilin-A1	12
P63086	Mitogen-activated protein kinase 1	13
P08461	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	12
Q63633	Solute carrier family 12 member 5	12
P17764	Acetyl-CoA acetyltransferase, mitochondrial	10
P62632	Elongation factor 1-alpha 2	12
Q9JHY2	Sideroflexin-3	12
Q5XIM9	T-complex protein 1 subunit beta	10
P61107	Ras-related protein Rab-14	11
P56571	ES1 protein homolog, mitochondrial	11
P59215	Guanine nucleotide-binding protein G(o) subunit alpha	11
P04642	L-lactate dehydrogenase A chain	11
P47709	Rabphilin-3A	11
P11598	Protein disulfide-isomerase A3	12
P15146	Microtubule-associated protein 2	10
Q02563	Synaptic vesicle glycoprotein 2A	12
P09606	Glutamine synthetase	11
Q9Z2L0	Voltage-dependent anion-selective channel protein 1	10

Q704S8	Carnitine O-acetyltransferase	10
Q05546	Tenascin-R	10
P16036	Phosphate carrier protein, mitochondrial	11
P85845	Fascin	13
P11884	Aldehyde dehydrogenase, mitochondrial	10
P06761	78 kDa glucose-regulated protein	10
P97536	Cullin-associated NEDD8-dissociated protein 1	11
P06686	Sodium/potassium-transporting ATPase subunit alpha-2	12
Q63716	Peroxiredoxin-1	9
P04177	Tyrosine 3-monooxygenase	11
Q05140	Clathrin coat assembly protein AP180	11
P56522	NADPH:adrenodoxin oxidoreductase, mitochondrial	11
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	12
P12007	Isovaleryl-CoA dehydrogenase, mitochondrial	10
P0C2X9	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	9
P33124	Long-chain-fatty-acid--CoA ligase 6	10
Q63065	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 1, mitochondrial	10
P47858	ATP-dependent 6-phosphofructokinase, muscle type	10
Q64428	Trifunctional enzyme subunit alpha, mitochondrial	10
Q60587	Trifunctional enzyme subunit beta, mitochondrial	9
P14604	Enoyl-CoA hydratase, mitochondrial	10
P16086	Spectrin alpha chain, non-erythrocytic 1	9
Q7TPB1	T-complex protein 1 subunit delta	9
P60203	Myelin proteolipid protein	13
Q924S5	Lon protease homolog, mitochondrial	9
P85970	Actin-related protein 2/3 complex subunit 2	10
Q8K4Y5	Leucine-rich glioma-inactivated protein 1	10
Q6AY30	Saccharopine dehydrogenase-like oxidoreductase	9
Q5M7U6	Actin-related protein 2	10
P29147	D-beta-hydroxybutyrate dehydrogenase, mitochondrial	8
Q5XIT1	Microtubule-associated protein RP/EB family member 3	9
P15650	Long-chain specific acyl-CoA dehydrogenase, mitochondrial	9
P04182	Ornithine aminotransferase, mitochondrial	8
Q4QRB4	Tubulin beta-3 chain	9
Q6P502	T-complex protein 1 subunit gamma	9
P14669	Annexin A3	9
P42123	L-lactate dehydrogenase B chain	8
Q3KRE0	ATPase family AAA domain-containing protein 3	11
P07895	Superoxide dismutase [Mn], mitochondrial	9
Q4V7C7	Actin-related protein 3	7
Q9Z0V6	Thioredoxin-dependent peroxide reductase, mitochondrial	9
Q5FVI6	V-type proton ATPase subunit C 1	9
Q794F9	4F2 cell-surface antigen heavy chain	8
P32851	Syntaxin-1A	8
P34058	Heat shock protein HSP 90-beta	9
P19804	Nucleoside diphosphate kinase B	8
P49432	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	8
P29411	GTP:AMP phosphotransferase AK3, mitochondrial	8
P60881	Synaptosomal-associated protein 25	8

P63102	14-3-3 protein zeta/delta	8
P07722	Myelin-associated glycoprotein	8
P05712	Ras-related protein Rab-2A	8
Q8CFD0	Sideroflexin-5	10
Q62717	Calcium-dependent secretion activator 1	9
Q920L2	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	9
Q9Z2F5	C-terminal-binding protein 1	8
P35704	Peroxiredoxin-2	8
P97546	Neuroplastin	8
B2GV06	Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	7
Q499N5	Acyl-CoA synthetase family member 2, mitochondrial	8
P07340	Sodium/potassium-transporting ATPase subunit beta-1	8
Q91Y81	Septin-2	7
O35796	Complement component 1 Q subcomponent-binding protein, mitochondrial	9
P09812	Glycogen phosphorylase, muscle form	6
Q8VHF5	Citrate synthase, mitochondrial	8
Q62952	Dihydropyrimidinase-related protein 3	7
P81155	Voltage-dependent anion-selective channel protein 2	7
Q05962	ADP/ATP translocase 1	7
P46462	Transitional endoplasmic reticulum ATPase	7
Q9QYJ6	cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A	7
P48500	Triosephosphate isomerase	7
P11442	Clathrin heavy chain 1	7
P11507	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	7
P61206	ADP-ribosylation factor 3	7
Q8R491	EH domain-containing protein 3	7
Q64057	Alpha-aminoadipic semialdehyde dehydrogenase	6
P54313	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	7
Q9Z270	Vesicle-associated membrane protein-associated protein A	7
Q5EB81	NADH-cytochrome b5 reductase 1	7
Q01205	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	7
P47860	ATP-dependent 6-phosphofructokinase, platelet type	6
P46101	Dipeptidyl aminopeptidase-like protein 6	7
Q63028	Alpha-adducin	7
P85973	Purine nucleoside phosphorylase	6
P02091	Hemoglobin subunit beta-1	6
P97532	3-mercaptopyruvate sulfurtransferase	6
Q6UPE1	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	6
Q6IFW6	Keratin, type I cytoskeletal 10	7
P19643	Amine oxidase [flavin-containing] B	7
Q5XIH7	Prohibitin-2	7
Q9WTT6	Guanine deaminase	6
Q9WU34	Neuronal-specific septin-3	7
Q6MGB5	Estradiol 17-beta-dehydrogenase 8	6
Q05683	Glutamate decarboxylase 2	6
P00406	Cytochrome c oxidase subunit 2	8
Q75Q39	Mitochondrial import receptor subunit TOM70	6

D4AAT7	ATP-dependent (S)-NAD(P)H-hydrate dehydratase	6
P23965	Enoyl-CoA delta isomerase 1, mitochondrial	6
Q3MIE4	Synaptic vesicle membrane protein VAT-1 homolog	6
P11275	Calcium/calmodulin-dependent protein kinase type II subunit alpha	7
P28480	T-complex protein 1 subunit alpha	6
Q5XIF3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	6
Q9QX69	LanC-like protein 1	6
P16638	ATP-citrate synthase	6
Q4V8B0	Oxidation resistance protein 1	6
P52303	AP-1 complex subunit beta-1	6
P24942	Excitatory amino acid transporter 1	6
P85969	Beta-soluble NSF attachment protein	6
Q5GFD9	Protein IMPACT	5
P27605	Hypoxanthine-guanine phosphoribosyltransferase	6
Q8R431	Monoglyceride lipase	6
Q7TS56	Carbonyl reductase family member 4	5
P19234	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial	5
Q6PDU7	ATP synthase subunit g, mitochondrial	7
P21913	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	6
P69897	Tubulin beta-5 chain	5
Q9R063	Peroxiredoxin-5, mitochondrial	5
P24368	Peptidyl-prolyl cis-trans isomerase B	5
Q6PCU2	V-type proton ATPase subunit E 1	5
P51647	Retinal dehydrogenase 1	5
Q08163	Adenylyl cyclase-associated protein 1	5
B0BNF1	Septin-8	6
P00388	NADPH--cytochrome P450 reductase	7
Q05695	Neural cell adhesion molecule L1	5
Q99P39	Cysteine desulfurase, mitochondrial	6
Q5FVH2	Phospholipase D3	5
Q924N5	Long-chain-fatty-acid--CoA ligase ACSBG1	6
Q9WVK7	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	5
Q63965	Sideroflexin-1	5
P0DJJ3	SH3-containing GRB2-like protein 3-interacting protein 1	5
P31662	Sodium-dependent neutral amino acid transporter SLC6A17	6
Q52KK3	Solute carrier family 25 member 51	5
P11951	Cytochrome c oxidase subunit 6C-2	5
Q6AYS7	Aminoacylase-1A	5
Q08877	Dynamin-3	5
P24473	Glutathione S-transferase kappa 1	5
P32089	Tricarboxylate transport protein, mitochondrial	5
Q4V8F9	Hydroxysteroid dehydrogenase-like protein 2	5
O88767	Protein DJ-1	6
P01946	Hemoglobin subunit alpha-1/2	5
Q704E8	ATP-binding cassette sub-family B member 7, mitochondrial	5
Q3T1J1	Eukaryotic translation initiation factor 5A-1	5
P0C1X8	AP2-associated protein kinase 1	5
Q01062	cGMP-dependent 3',5'-cyclic phosphodiesterase	5

Q9JI66	Electrogenic sodium bicarbonate cotransporter 1	4
O70511	Ankyrin-3	5
Q6IMF3	Keratin, type II cytoskeletal 1	5
P49803	Regulator of G-protein signaling 7	5
Q9R0I8	Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha	4
Q63345	Myelin-oligodendrocyte glycoprotein	4
Q9QZR6	Septin-9	5
Q5RJQ4	NAD-dependent protein deacetylase sirtuin-2	4
Q5M7A7	CB1 cannabinoid receptor-interacting protein 1	4
P61983	14-3-3 protein gamma	4
Q9R1Z0	Voltage-dependent anion-selective channel protein 3	4
Q9QXQ0	Alpha-actinin-4	4
Q6PST4	Atlastin-1	4
Q9WU82	Catenin beta-1	4
Q68FX9	NAD-dependent protein deacylase sirtuin-5, mitochondrial	4
P05370	Glucose-6-phosphate 1-dehydrogenase	4
P48768	Sodium/calcium exchanger 2	4
O88483	[Pyruvate dehydrogenase [acetyl-transferring]]-phosphatase 1, mitochondrial	6
O88994	Mitochondrial amidoxime reducing component 2	4
P12075	Cytochrome c oxidase subunit 5B, mitochondrial	8
Q6QIX3	Zinc transporter 3	4
P21708	Mitogen-activated protein kinase 3	4
Q4QQV3	Protein FAM162A	4
Q5I0G4	Glycine--tRNA ligase (Fragment)	4
Q5XIE6	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	4
Q63362	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	4
Q4G064	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial	4
Q1WIM3	Cell adhesion molecule 3	5
Q91ZN1	Coronin-1A	4
Q62847	Gamma-adducin	4
P30835	ATP-dependent 6-phosphofructokinase, liver type	4
P47875	Cysteine and glycine-rich protein 1	4
P41499	Tyrosine-protein phosphatase non-receptor type 11	4
P31016	Disks large homolog 4	4
P62630	Elongation factor 1-alpha 1	4
O35331	Pyridoxal kinase	3
O55012	Phosphatidylinositol-binding clathrin assembly protein	4
Q1HCL7	NAD kinase 2, mitochondrial	4
Q5M821	Protein phosphatase 1H	4
P52481	Adenylyl cyclase-associated protein 2	4
B2RYW9	Fumarylacetoacetate hydrolase domain-containing protein 2	3
Q5FVJ0	Protein RUFY3	4
P84850	D-2-hydroxyglutarate dehydrogenase, mitochondrial	5
Q568Z9	Phytanoyl-CoA hydroxylase-interacting protein	3
P47819	Glial fibrillary acidic protein	3
P45592	Cofilin-1	4
Q9Z0J5	Thioredoxin reductase 2, mitochondrial	4
Q3B8Q0	Microtubule-associated protein RP/EB family member 2	3
P24329	Thiosulfate sulfurtransferase	3

P62260	14-3-3 protein epsilon	3
P35434	ATP synthase subunit delta, mitochondrial	4
Q5I6B8	Phosphatidylinositol 4-phosphate 5-kinase type-1 gamma	4
P18265	Glycogen synthase kinase-3 alpha	3
P34926	Microtubule-associated protein 1A	3
P63100	Calcineurin subunit B type 1	3
P22062	Protein-L-isoaspartate(D-aspartate) O-methyltransferase	3
Q64591	2,4-dienoyl-CoA reductase, mitochondrial	3
Q6AYT7	Monoacylglycerol lipase ABHD12	3
Q4FZU2	Keratin, type II cytoskeletal 6A	3
Q6P9T8	Tubulin beta-4B chain	3
Q5XIF6	Tubulin alpha-4A chain	3
M0RC99	Ras-related protein Rab-5A	3
P11608	ATP synthase protein 8	3
P25286	V-type proton ATPase 116 kDa subunit a isoform 1	4
P63004	Platelet-activating factor acetylhydrolase IB subunit alpha	3
O08776	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 3	3
P69682	Adaptin ear-binding coat-associated protein 1	4
Q6Q0N1	Cytosolic non-specific dipeptidase	4
Q68FW7	Threonine--tRNA ligase, mitochondrial	4
P62882	Guanine nucleotide-binding protein subunit beta-5	3
Q6AXT5	Ras-related protein Rab-21	3
O08618	Phosphoribosyl pyrophosphate synthase-associated protein 2	3
Q5RKI8	ATP-binding cassette sub-family B member 8, mitochondrial	3
P63045	Vesicle-associated membrane protein 2	3
Q63564	Synaptic vesicle glycoprotein 2B	3
Q04400	Adenylate cyclase type 5	3
Q62888	Neuroligin-2	3
Q9Z214	Homer protein homolog 1	3
P0C089	Phosphatidylglycerophosphatase and protein-tyrosine phosphatase 1	3
D3ZQG6	Tripartite motif-containing protein 2	3
P18088	Glutamate decarboxylase 1	4
Q66HA8	Heat shock protein 105 kDa	3
Q64568	Plasma membrane calcium-transporting ATPase 3	4
P60522	Gamma-aminobutyric acid receptor-associated protein-like 2	3
P10760	Adenosylhomocysteinase	3
P63329	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	3
Q920F5	Malonyl-CoA decarboxylase, mitochondrial	3
Q05982	Nucleoside diphosphate kinase A	3
P85515	Alpha-centractin	3
Q5XI22	Acetyl-CoA acetyltransferase, cytosolic	2
Q505J6	Mitochondrial glutamate carrier 2	3
P62483	Voltage-gated potassium channel subunit beta-2	3
Q5XI79	NADH dehydrogenase [ubiquinone] complex I, assembly factor 7	2
P97576	GrpE protein homolog 1, mitochondrial	3
P62332	ADP-ribosylation factor 6	3
P13638	Sodium/potassium-transporting ATPase subunit beta-2	4
P31977	Ezrin	3

Q62915	Peripheral plasma membrane protein CASK	4
O35180	Endophilin-A3	2
A1L108	Actin-related protein 2/3 complex subunit 5-like protein	2
P08503	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	3
P38718	Mitochondrial pyruvate carrier 2	3
Q5FVQ8	NLR family member X1	2
Q64542	Plasma membrane calcium-transporting ATPase 4	4
B0BN86	Transmembrane protein 11, mitochondrial	3
Q1WIM1	Cell adhesion molecule 4	2
Q6P799	Serine--tRNA ligase, cytoplasmic	2
P62813	Gamma-aminobutyric acid receptor subunit alpha-1	3
Q6JE36	Protein NDRG1	3
P23928	Alpha-crystallin B chain	2
O35264	Platelet-activating factor acetylhydrolase IB subunit beta	2
P60892	Ribose-phosphate pyrophosphokinase 1	2
Q5M7W5	Microtubule-associated protein 4	2
Q62634	Vesicular glutamate transporter 1	3
P62986	Ubiquitin-60S ribosomal protein L40	3
B0K020	CDGSH iron-sulfur domain-containing protein 1	3
Q5PPI4	Lysosome-associated membrane glycoprotein 5	2
P52504	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial	2
Q9Z250	Protein lin-7 homolog A	3
D3ZAF6	ATP synthase subunit f, mitochondrial	2
P08413	Calcium/calmodulin-dependent protein kinase type II subunit beta	2
P29419	ATP synthase subunit e, mitochondrial	2
B2RYT9	Translational activator of cytochrome c oxidase 1	2
Q6MG12	Uncharacterized protein C6orf136 homolog	2
P35213	14-3-3 protein beta/alpha	2
Q9NQR8	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 4	2
Q64559	Cytosolic acyl coenzyme A thioester hydrolase	2
P68370	Tubulin alpha-1A chain	2
Q2PQA9	Kinesin-1 heavy chain	2
P00762	Anionic trypsin-1	3
Q6AYQ8	Acylpyruvase FAHD1, mitochondrial	2
Q63092	CaM kinase-like vesicle-associated protein	2
Q5M9I5	Cytochrome b-c1 complex subunit 6, mitochondrial	3
Q4KM45	UPF0687 protein C20orf27 homolog	2
Q9JKE3	Secretory carrier-associated membrane protein 5	2
Q505J9	ATPase family AAA domain-containing protein 1	2
P63331	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform	2
P50408	V-type proton ATPase subunit F	2
Q9JJW3	Up-regulated during skeletal muscle growth protein 5	2
P86252	Transcriptional activator protein Pur-alpha (Fragments)	2
P32232	Cystathionine beta-synthase	2
Q5EB62	Solute carrier family 25 member 46	2
Q9Z1N4	3'(2'),5'-bisphosphate nucleotidase 1	3
P70619	Glutathione reductase (Fragment)	2
O88588	Phosphofurin acidic cluster sorting protein 1	2

Q6XVN8	Microtubule-associated proteins 1A/1B light chain 3A	2
P07825	Synaptophysin	3
Q6IG02	Keratin, type II cytoskeletal 2 epidermal	3
Q99PD4	Actin-related protein 2/3 complex subunit 1A	3
Q6EIX2	Mitochondrial import inner membrane translocase subunit TIM16	2
P04631	Protein S100-B	2
Q5I0K5	Mycophenolic acid acyl-glucuronide esterase, mitochondrial	2
Q03346	Mitochondrial-processing peptidase subunit beta	2
P07936	Neuromodulin	2
Q68FQ0	T-complex protein 1 subunit epsilon	1
P04764	Alpha-enolase	2
Q9WVK3	Peroxisomal trans-2-enoyl-CoA reductase	3
Q62696	Disks large homolog 1	4
P35171	Cytochrome c oxidase subunit 7A2, mitochondrial	2
P29418	ATP synthase subunit epsilon, mitochondrial	2
P31647	Sodium- and chloride-dependent GABA transporter 3	2
P38650	Cytoplasmic dynein 1 heavy chain 1	2
P0C5X8	Protein tweety homolog 1	2
P08644	GTPase KRas	2
Q9WUS0	Adenylate kinase 4, mitochondrial	2
O08662	Phosphatidylinositol 4-kinase alpha	2
P11960	2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial (Fragment)	2
P0C546	Mitochondrial coenzyme A transporter SLC25A42	2
P68403	Protein kinase C beta type	3
Q62784	Type I inositol 3,4-bisphosphate 4-phosphatase	2
O35094	Mitochondrial import inner membrane translocase subunit TIM44	2
P54311	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	2
Q9WU49	Calcium-regulated heat stable protein 1	2
P85968	6-phosphogluconate dehydrogenase, decarboxylating	2
Q9JJ19	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	2
P04218	OX-2 membrane glycoprotein	2
Q5RJL0	Ermin	2
P80432	Cytochrome c oxidase subunit 7C, mitochondrial	2
Q5U2X7	Mitochondrial import inner membrane translocase subunit Tim21	2
Q9WTP0	Band 4.1-like protein 1	4
Q5PPN7	Coiled-coil domain-containing protein 51	2
Q5M7W7	Probable proline--tRNA ligase, mitochondrial	3
P07633	Propionyl-CoA carboxylase beta chain, mitochondrial	2
Q5BJU7	Wiskott-Aldrich syndrome protein family member 1	2
P11661	NADH-ubiquinone oxidoreductase chain 5	3
Q66HA6	ADP-ribosylation factor-like protein 8B	2
Q7TQ16	Cytochrome b-c1 complex subunit 8	2
Q68FT3	Pyridine nucleotide-disulfide oxidoreductase domain-containing protein 2	2
P26772	10 kDa heat shock protein, mitochondrial	2
Q9QZM5	Abl interactor 1	2
P62744	AP-2 complex subunit sigma	2
P19627	Guanine nucleotide-binding protein G(z) subunit alpha	2
Q5EBA1	ATP-dependent RNA helicase SUPV3L1, mitochondrial	1

Q5BJQ0	Chaperone activity of bc1 complex-like, mitochondrial	1
Q6AY19	AarF domain-containing protein kinase 4	2
P13437	3-ketoacyl-CoA thiolase, mitochondrial	1
Q63372	Neurexin-1	2
Q3KRD5	Mitochondrial import receptor subunit TOM34	2
Q63377	Sodium/potassium-transporting ATPase subunit beta-3	1
O54701	Sodium/potassium/calcium exchanger 2	2
P35738	2-oxoisovalerate dehydrogenase subunit beta, mitochondrial	1
P17425	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	1
Q64548	Reticulon-1	2
Q5I0L3	Tyrosine--tRNA ligase, mitochondrial	1
Q07803	Elongation factor G, mitochondrial	2
P97544	Lipid phosphate phosphohydrolase 3	1
Q5XIM4	ATP synthase subunit s, mitochondrial	2
Q62936	Disks large homolog 3	1
Q6P6Q9	Calcium uptake protein 1, mitochondrial	1
O35964	Endophilin-A2	1
P35281	Ras-related protein Rab-10	2
A1L1I3	Numb-like protein	1
O88917	Latrophilin-1	1
Q63617	Hypoxia up-regulated protein 1	1
Q68FR6	Elongation factor 1-gamma	2
P18266	Glycogen synthase kinase-3 beta	1
Q6AXW1	Glutaredoxin-2, mitochondrial	1
Q32PX9	Lactation elevated protein 1	1
O35095	Neurochondrin	1
Q9R1T5	Aspartoacylase	1
Q66HR2	Microtubule-associated protein RP/EB family member 1	1
Q8VHK2	Caskin-1	1
O88377	Phosphatidylinositol 5-phosphate 4-kinase type-2 beta	1
P45953	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	1
P00564	Creatine kinase M-type	1
Q6GQP4	Ras-related protein Rab-31	1
P63079	Gamma-aminobutyric acid receptor subunit beta-3	1
Q99MC0	Protein phosphatase 1 regulatory subunit 14A	1
Q3ZB98	Breast carcinoma-amplified sequence 1 homolog (Fragment)	2
Q9JM53	Apoptosis-inducing factor 1, mitochondrial	1
Q63622	Disks large homolog 2	1
O35952	Hydroxyacylglutathione hydrolase, mitochondrial	1
Q5PQM2	Kinesin light chain 4	1
O88954	MAGUK p55 subfamily member 3	1
P68035	Actin, alpha cardiac muscle 1	1
O35764	Neuronal pentraxin receptor	1
D3ZPX4	Plexin-A3	1
Q6Q7Y5	Guanine nucleotide-binding protein subunit alpha-13	1
P26817	Beta-adrenergic receptor kinase 1	1
Q4V882	Epsin-3	1
A6JUQ6	Clavesin-2	1
A2VCX1	TIP41-like protein	1

Q99MZ8	LIM and SH3 domain protein 1	1
Q6PEC0	Bis(5'-nucleosyl)-tetraphosphatase [asymmetrical]	1
Q62753	Syntaxin-binding protein 2	1
P62494	Ras-related protein Rab-11A	1
P05503	Cytochrome c oxidase subunit 1	1
Q9JK15	Arf-GAP with dual PH domain-containing protein 2	1
D3ZN43	NADH dehydrogenase (ubiquinone) complex I, assembly factor 6	1
Q5XIJ4	Protein FAM210A	1
P97694	Cytohesin-1	1
P84060	Dystrobrevin beta	1
D4A6D7	Tetratricopeptide repeat protein 19, mitochondrial	1
Q5FVI3	Leucine-rich repeat-containing protein 57	1
Q4V8I7	Volume-regulated anion channel subunit LRRC8A	1
P80431	Cytochrome c oxidase subunit 7B, mitochondrial	1
Q6AXQ5	2',5'-phosphodiesterase 12	1
P11517	Hemoglobin subunit beta-2	1
P47861	Synaptotagmin-5	1
Q6Q629	Inactive dipeptidyl peptidase 10	1
Q8VHV7	Heterogeneous nuclear ribonucleoprotein H	1
B5DFN3	Ubiquinol-cytochrome-c reductase complex assembly factor 2	1
Q5HZY2	GTP-binding protein SAR1b	1
P16617	Phosphoglycerate kinase 1	1
O88370	Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma	1
A2RUW1	Toll-interacting protein	1
Q9WTV5	26S proteasome non-ATPase regulatory subunit 9	1
Q925Q9	SH3 domain-containing kinase-binding protein 1	1
Q5BJP6	Ribosome-releasing factor 2, mitochondrial	1
Q3T1K5	F-actin-capping protein subunit alpha-2	1
P04905	Glutathione S-transferase Mu 1	1
B2DD29	Serine/threonine-protein kinase BRSK1	1
Q9Z142	Transmembrane protein 33	1
Q91XR8	Phospholipid hydroperoxide glutathione peroxidase, nuclear	1
Q6AYQ4	Transmembrane protein 109	1
Q63081	Protein disulfide-isomerase A6	1
Q5XIL3	DNA-directed RNA polymerase III subunit RPC3	1
Q5FVI4	Cell cycle exit and neuronal differentiation protein 1	1
Q4FZX5	Methionine-R-sulfoxide reductase B2, mitochondrial	1
P97710	Tyrosine-protein phosphatase non-receptor type substrate 1	1
P62083	40S ribosomal protein S7	1
P62074	Mitochondrial import inner membrane translocase subunit Tim10	1
P61203	COP9 signalosome complex subunit 2	1
P61023	Calcineurin B homologous protein 1	1
P56603	Secretory carrier-associated membrane protein 1	1
P13852	Major prion protein	1
B0BN94	Protein FAM136A	1
Q924S1	1-acyl-sn-glycerol-3-phosphate acyltransferase delta	1
Q05764	Beta-adducin	1
Q5RJR2	Twinfilin-1	1
P82995	Heat shock protein HSP 90-alpha	2

P41498	Low molecular weight phosphotyrosine protein phosphatase	2
P08010	Glutathione S-transferase Mu 2	1
Q91ZW6	Trimethyllysine dioxygenase, mitochondrial	1
P20069	Mitochondrial-processing peptidase subunit alpha	1
B2RYW8	MIC	1
Q6AY04	Uncharacterized protein C2orf47 homolog, mitochondrial	1
Q5M7T2	SPRY domain-containing protein 7	1
Q8K3P6	Calcium-binding mitochondrial carrier protein SCaMC-2	1
Q91ZW1	Transcription factor A, mitochondrial	1
Q91ZN7	Serine/threonine-protein kinase Chk1	1
Q5XIT9	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	1
Q68FU7	Ubiquinone biosynthesis monooxygenase COQ6	1
P36876	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	1
P10818	Cytochrome c oxidase subunit 6A1, mitochondrial	1
Q8VGC3	Voltage-dependent L-type calcium channel subunit beta-2	1
Q5XIA8	Growth hormone-inducible transmembrane protein	2
Q05511	Serine protease hepsin	1
Q62889	Neuroligin-3	1
Q6P6Q2	Keratin, type II cytoskeletal 5	1
P31422	Metabotropic glutamate receptor 3	1
P17105	Inositol-trisphosphate 3-kinase A	1
Q62656	Receptor-type tyrosine-protein phosphatase zeta	1
P03889	NADH-ubiquinone oxidoreductase chain 1	1
D3ZFB6	Proline-rich transmembrane protein 2	2
P55161	Nck-associated protein 1	2
P62898	Cytochrome c, somatic	1
Q68FS4	Cytosol aminopeptidase	2
P14173	Aromatic-L-amino-acid decarboxylase	2
P20650	Protein phosphatase 1A	1
Q62636	Ras-related protein Rap-1b	1
P27653	C-1-tetrahydrofolate synthase, cytoplasmic	1
Q9Z2Z8	7-dehydrocholesterol reductase	1
Q99P82	Claudin-11	1
Q6AYN4	Phytanoyl-CoA hydroxylase-interacting protein-like	1
Q07205	Eukaryotic translation initiation factor 5	1
P04775	Sodium channel protein type 2 subunit alpha	1
Q6RJR6	Reticulon-3	1
Q4KLF8	Actin-related protein 2/3 complex subunit 5	1
Q99JD4	CLIP-associating protein 2	1
Q64232	Very-long-chain enoyl-CoA reductase	1
P05508	NADH-ubiquinone oxidoreductase chain 4	1