Sarah STRYECK, BSc

Molecular Mechanisms of Neurodegenerative Diseases

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science

Masterstudium Biochemie und Molekulare Biomedizin

eingereicht an der

Technischen Universität Graz

Betreuer

Assoc. Prof. Dr. Tobias Madl

Institut für Molekularbiologie und Biochemie

Medizinische Universität Graz

Graz, Juni 2016

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Datum

Unterschrift

Content

Abbreviations

°C	degree celsius
^{1}H	hydrogen isotope with 1 proton
ADMA	ω-N^{G}, N^{G} dimethyl-L-arginine
ALL	acute lymphoblastic leukemia
ALS	amyotrophic lateral sclerosis
APS	ammonium persulfate
CBS	Cystathione beta synthase
CBS $^{-/-}$	homologous knockout of CBS
CIRP	cold inducible RNA binding protein
CIS	clinically isolated syndrome
CPMG	Carr- Purcell- Meiboom-Gill
CO_2	carbon dioxide
E. coli	Escherichia coli
EDTA	Ethylenediaminetetaacetic acid
EGTA	Ethylenglycol-bis(aminoethylether)-N,N,N′,N′-tetraessigsäure
EMT	epithelial mesenchymal transition
fALS	familial ALS
FTD	Frontotemporal dementia
FUS	Fused in Sarcoma
H_2O	hydrogen dioxide
HCl	hydrochloride
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
His-tag	polyhistidin-tag
hnRNP	heterogenous ribonucleoprotein
ILVM	isoleucine leucine valine methionine
K	Kelvin
K_d	dissociation constant
kDa	kilo dalton
LB	lysogeny broth
IPTG	isopropyl ß-D-1-thiogalactopyranoside
M	molar
MeOH	methanol
MES	2-(N-Morpholino)ethansulfonsäure
$MgSO_4$	magnesium sulfate
ml	milliliter
mM	millimolar
MMA	ω-N^{G} monomethyl-L-arginine
$MnCl_2$	mangan chloride
MOPS	3-(N-morpholino)propanesulfonic acid

mRNA	messenger ribonucleic acid
MS	Multiple Sclerosis
N_2	molecular nitrogen
Na_2CO_3	natrium carbonate
NaOH	natriumhydroxide
nm	nanometer
NMR	nuclear magnetic resonance
OD	optical density
O-PLS-DA	Orthogonal-Partial Least Square-Discriminant Analysis
P525L	amino acid exchange prolin-to-leucine at position 525
PAGE	polyacrylamide gelelectrophoresis
PCA	Principal Component Analysis
PEG MMA 2000	Polyethylenglycol Methylmethacrylate 2000
PRMT1	protein arginine methyltransferase
PY-NLS	Prolin-tyrosine nuclear localization signal
RbCl	rubidium chloride
RBM3	RNA binding motif protein 3
RGG	arginine-glycine rich region
RNA	ribonucleic acid
SAH	S-adenosyl-homocysteine
SAM	S-adenosyl-L-methionine
Sam68	Src-associated in mitosis 68 kDa protein
SAXS	Small angle X-ray scattering
SDMA	ω-N^G, $N^{G'}$ dimethyl-L-arginine
SDS	sodium dodecyl sulfate
SRSF1	serine arginine rich splicing factor 1
SY	serine-tyrosine rich region
TEV	Tobacco Etch Virus
TFB1	Transformation buffer 1
TFB2	Transformation buffer 2
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	Tetramethylethylendiamin
TNPO1	Transportin-1
TNPO1Δlinker	Transportin-1 lacking flexible loop region
Tris	Tris(hydroxymethyl)aminomethane
TSP	sodium trimethylsilyl [2,2,3,3-$^2H_4$] propionate
TWIST1	twist related protein 1
X-ray	x-radiation

Zusammenfassung

Die Arginin-Methylierung ist eine post-translationale Modifikation, bei welcher die Methylgruppe des Donormoleküls S-Adenosylmethionin (SAM) an eines oder mehrere Guanidinostickstoffatome in der Seitenkette der Aminosäure Arginin übertragen wird. Diese Reaktion wird hauptsächlich von der Protein Arginin Methyltransferase 1 (PRMT1) katalysiert. [1–3] Es wurde gezeigt, dass Proteine der hnRNP-Proteinfamilie, welche Arginin- und Glycin-reiche Sequenzregionen haben (RG/RGG-Motive), von Protein Arginin Methyltransferasen methyliert werden. [4,5] Bei dem Protein Fused–in-Sarcoma (FUS) reguliert die Methylierung des RGG-Motivs die Interaktion mit dem Karyopherin Transportin-1 und dadurch den Kernimport. Bei einer Misregulation dieser post-translationalen Modifikation kommt es zur Mislokalisation der Proteine und in weiterer Folge zur Bildung von zytoplasmatischen Proteinaggregaten, was ein zentrales Merkmal der neurodegenerativen Erkrankungen Amyotrophe Lateralsklerose (ALS) und Frontotemporale Dementia (FTD) ist. [4,6] Außerdem ist auch eine große Anzahl anderer RG/RGG Proteine in anderen neurodegenerativen Erkrankungen und proliferativen Krebserkrankungen involviert. [7–9]

Die Hypothese dieser Studie ist, dass eine verringerte Arginin-Methylierung zu neurodegenerativen Erkrankungen führt, wohingegen eine Hypermethylierung für die Entstehung von Krebserkrankungen verantwortlich ist. Das Ziel dieser Arbeit ist die Aufklärung von molekularen Schlüsselmechanismen, welche zu neurodegenerativen Erkrankungen führen. Dafür werden regulatorische Mechanismen des Kernimports von RG/RGG Proteinen der hnRNP Familie anhand des Beispiels Cold-inducible RNA binding protein (CIRP) untersucht. Um das Wissen über die regulatorische Arginin-Methylierung und damit potentiell assoziierte Biomarker in Patienten nachweisen zu können, wurde ein Protokoll für die NMR-basierte metabolische Phänotypisierung erarbeitet und an Proben einer Patientenkohorte getestet, um Metabolitenprofile zu untersuchen bzw. Metaboliten der Argininmethylierung nachzuweisen.

Eine Kombination aus biochemischen und strukturbiologischen Techniken wird in dieser Arbeit verwendet, um humane Proteinkonstrukte (CIRP und Transportin-1) heterolog in *Escherichia coli* BL21 DE3 Zellen zu exprimieren und mittels immobilisierter Metallionenaffinitätschromatographie und Größenausschlusschromatographie

aufgereinigt. Bindungsstudien von CIRP-, als auch FUS-Proteinkonstrukten an Transportin-1 werden mit isothermer Titrationskalorimetrie und Kernspinresonanzspektroskopie (NMR – nuclear magnetic resonance) durchgeführt. In weiterer Folge wurde NMR-Metabolomik getestet, um Metaboliten des Arginin-Methylierungsstoffwechsels, als auch Biomarker generell zu identifizieren und zu quantifizieren. Diese metabolische Phänotypisierung wurde erweitert mit Festphasenextraktion und optimierten NMR Pulssequenzen, um die Anzahl an Metaboliten, die identifiziert werden können, zu erhöhen.

Ergebnisse dieser Studien können auf das gesamte RG/RGG Proteom, welches mehr als 1000 Proteine umfasst, übertragen werden. Viele dieser Proteine sind involviert in einer Vielzahl an Krankheiten wie Neurodegeneration und Krebs. Das RGG Motiv und die SY Region (Serin- und Tyrosin-reiche Region) von CIRP wurden als Transportin-1 bindende Elemente identifiziert, wodurch sie im Kernimport von CIRP eine zentrale Rolle spielen. Da die CIRP SY-Region auch ein potentielles Ziel für Proteinkinasen ist, könnte eine Phosphorylierung eine weitere wichtige post-translationale Modifikation sein, die den Kernimport reguliert. Das RGG-Motiv von CIRP zeigt eine stärkere Bindung an Transportin-1 als die SY Region, weshalb diese Bindung auch bei der Analyse des Volllängenkonstrukts dominiert. Die Analyse von Patientenproben mittels NMR-basierter metabolischer Phänotypisierung zeigte, dass Metaboliten aus dem Argininmethylierungsstoffwechsel neben einer Vielzahl an anderen Metaboliten in Bioflüssigkeiten identifiziert werden können. In einem Demonstrationsversuch konnte anhand einer Patientengruppe gezeigt werden, dass die etablierten Protokolle gut geeignet sind, um Änderungen in Metaboliten und damit Biomarker zu identifizieren.
Ein Wechselspiel der Phosphorylierung von CIRP-SY und der Methylierung von CIRP-RGG könnte die Affinität von CIRP zu Transportin-1 regulieren. Weitere Analysen von den Bioflüssigkeiten könnten detektierbare Unterschiede in Metaboliten des Arginin-Methylierungsstoffwechsels identifizieren, was ein wichtiges Diagnostikkriterium werden könnte.

Abstract

Arginine methylation is a post-translational modification which adds a methyl group of the donor molecule S-adenosyl-methionine (SAM) to one or several guanidino nitrogen atoms in the sidechain of amino acid arginine and is mainly catalyzed by protein arginine methyltransferase 1 (PRMT1). [1–3] It has been reported that members of the large family of hnRNPs harbouring an arginine-glycine-glycine repeat region (RG/RGG) are methylated by arginine methyltransferases.[4,5] In the case of the protein Fused-in-Sarcoma (FUS), the methylation of the RGG motif has been discovered to regulate the interaction with transportin-1, a member of the family of karyopherins, and therefore the nuclear import of the protein. A misregulation of this posttranslational modification leads to mislocalization of proteins and in further consequence to development of cytoplasmic inclusions which is the hallmark of neurodegenerative diseases including Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD).[4,6] In addition, a plethora of other RG/RGG proteins is associated with other neurodegenerative diseases as well as cancer. [7–9]

The broad perspective hypothesis of this study is, that a reduced arginine methylation is involved in neurodegenerative diseases, whereas a hypermethylation is responsible for the development of cancer. The central aim of my thesis is to reveal key molecular mechanisms causing neurodegenerative diseases. Therefore the mechanisms of regulatory nuclear import of RG/RGG proteins using the hnRNP model system Cold-inducible RNA binding protein (CIRP) is analyzed. To analyze these findings about regulatory arginine methylation and consequently associated biomarkers, a protocol for NMR-based metabolic phenotyping is applied to samples of patient cohort studies to investigate metabolic profiles and to detect metabolites linked to the arginine methylation metabolism.

A combination of biochemical and structural biology techniques was applied to express human protein constructs (CIRP, transportin-1) heterologously in *Escherichia coli* BL21 DE3 and to purify them using immobilized metal ion affinity chromatography (IMAC) and Size Exclusion Chromatography (SEC). Binding studies of CIRP- and FUS-constructs to transportin-1 constructs were carried out using Isothermal Titration Calorimetry (ITC) and Nuclear Magnetic Resonance spectroscopy (NMR). Thus far, NMR-based metabolomics has been tested to detect and to quantify metabolites linked to arginine methylation as well as biomarkers in general in patient samples. This metabolic phenotyping pipeline was further extended with solid phase extraction (SPE) and different optimized NMR pulse se-

quence setups to extend the number of covered metabolites linked to arginine methylation metabolism.

These findings can be translated to the entire RG/RGG proteome of more than 1000 proteins whose misregulation is involved in a plethora of human diseases including neurodegeneration and cancer. RGG-motif as well as SY-motif (serine-tyrosine rich region) of CIRP were identified as transportin-1 binding regions and are therefore involved in nuclear import of the protein. Since CIRP-SY region is also a target for protein kinases, a phosphorylation may be another post-translational modification regulating the nuclear import. CIRP-RGG shows a higher affinity to transportin-1 which also explains the dominance of this binding event in the full-length CIRP construct. Analysis of patients' samples using NMR-based metabolic phenotyping shows, that identification of metabolites of arginine methylation pathway in addition to a variety of other metabolites is possible. The demonstration test indicated that protocols are suitable to identify variations in metabolites and hence also in biomarkers.
Interplay of phosphorylation or methylation of CIRP-SY and CIRP-RGG, respectively may regulate affinity of CIRP to transportin-1 depending on type of modification. Further analysis of patients' biofluids may reveal detectable differences in metabolites of arginine methylation pathway which may be a promising diagnostic criterion.

1. Introduction

1.1. The RGG/RG proteome consists of thousands of proteins

The RGG/RG proteome is a large protein family in a variety of species including humans, which comprises proteins harboring arginine-glycine rich regions in their amino acid sequence. In the human proteome, there are thousands of proteins containing tri-RGG, di-RGG, tri-RG or di-RG sequence motifs (for examples see Fig. 1, 2). [10] These sequence motifs are known to be involved in RNA binding. [11] In addition, it has been reported that a variety of protein-protein interactions of members of the RGG/RG proteome are feasible due to the binding to interaction partners via its arginine-glycine rich sequence region. [12–15]

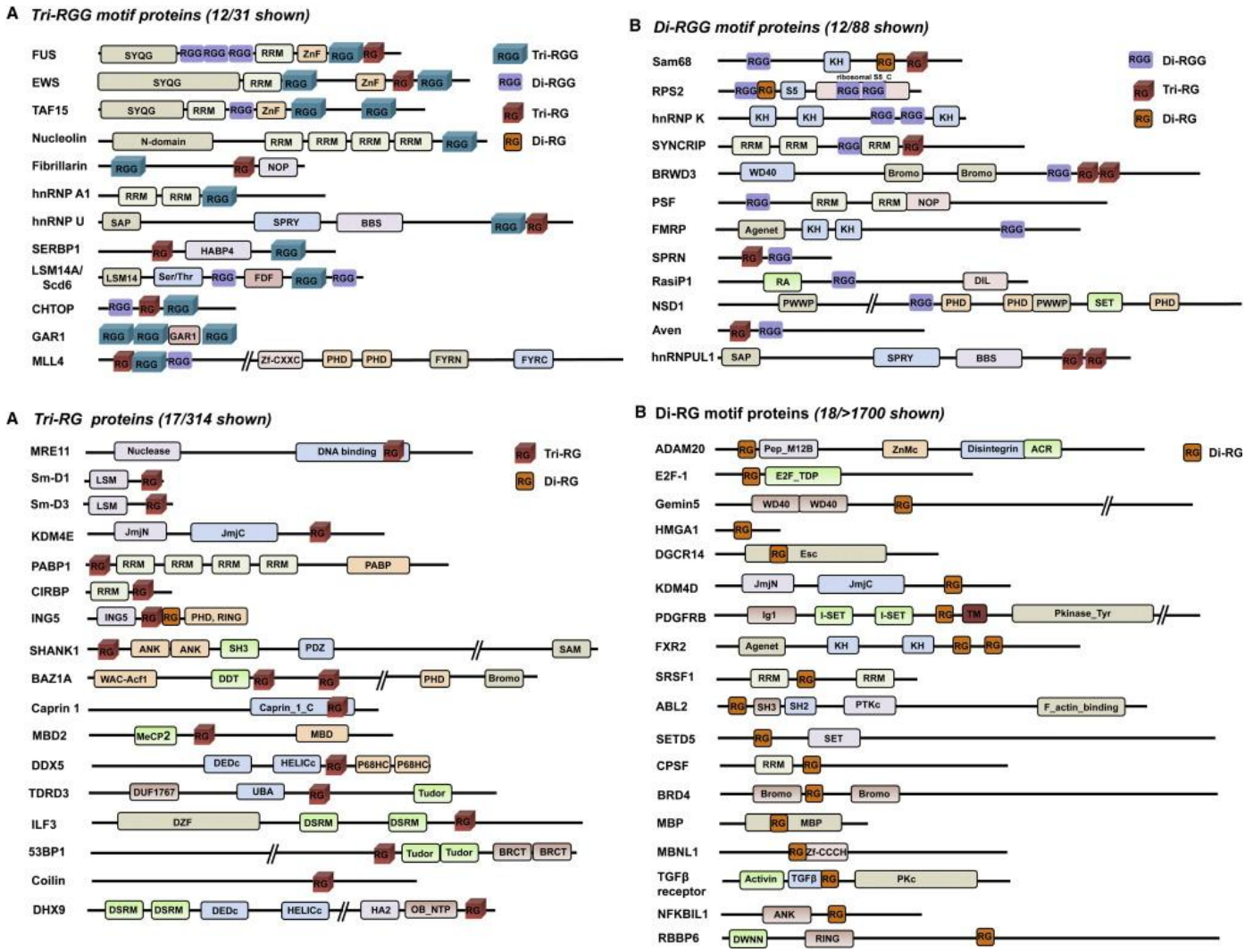


Fig. 1, Fig. 2: RGG and RG motif-containing proteins in the human genome. Adapted from Thandapani, P. et al 2013 [10]

1.2. Functions of RGG/RG proteins

Essential processes in cells are mediated by proteins harboring RGG/RG sequence motifs. Translational repression is a mechanism, which is regulated by proteins having arginine-glycine rich regions in their amino acid sequence. Nucleolin, a eukaryotic tri-RGG phosphoprotein, for example, represses the translation of the tumorsuppressor p53 mRNA upon DNA damage. The same protein is also involved in regulation of apoptosis. The mRNA of anti-apoptotic protein Bcl-2 is stabilized by binding of Nucleolin to the 3' untranslated region. [16–18] E2F-1, a transcription factor, is on the one hand involved regulation of apoptosis upon DNA damage and on the other hand in transcription control of genes expressed in synthesis phase of cell cycle. [19] hnRNPUL1 participates in DNA damage signaling via interaction with NBS1. [20]

Taken all together, the RGG/RG proteome is involved in a variety of different functions ensuring the quality and amount of proteins as well as of DNA, which is inevitable for a proper function of essential cellular mechanisms (Fig. 4).

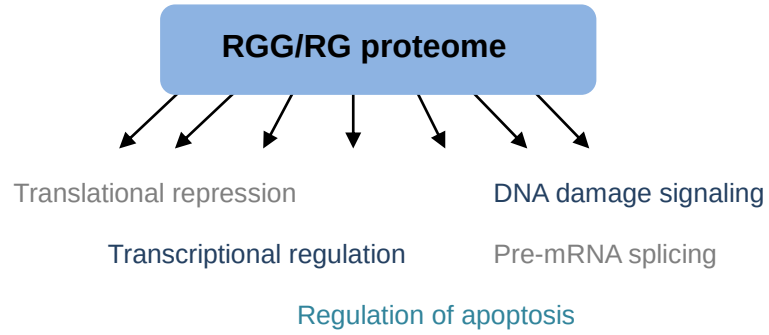


Fig. 4: Functions of RGG/RG proteins

1.3. RGG/RG Motifs are post-translationally modified by arginine methylation

Post-translational modifications extend the cell's repertoire of proteins having highly specific functions. The amino acid arginine is a residue which can be modified via enzymatic methylation. [1] Protein arginine methyltransferase 1 (PRMT1) was identified as the predominant enzyme capable of adding a methyl group to guanidino nitrogen atoms of arginine residues. [3,2] Since proteins of the RGG/RG proteome contain arginine-rich regions, they are a putative target of PRMT1. This post-translational modification regulates different functions of proteins. [14,21,22] There are three distinct forms of methylated arginine:

- ω-N^G-monomethyl-L-arginine (MMA)
- ω-N^G, N^G-dimethyl-L-arginine (ADMA) see Fig. 5
- ω-N^G, $N^{G'}$-dimethyl-L-arginine (SDMA) see Fig. 6

The most abundant methylation reaction on arginine residues mediated by PRMT1 is the production of ω-N^G, N^G-dimethyl-L-arginine (ADMA). [23]

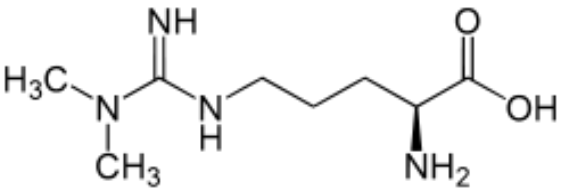


Fig. 5: Structure of ADMA

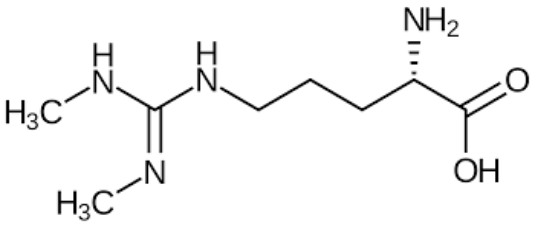


Fig. 6: Structure of SDMA

The methylgroup donor for arginine methylation is S-adenosyl-L-methionine (SAM), which is consequently converted to S-adenosyl-homocysteine (SAH). In addition, the regeneration to SAM comprises hydrolysis of SAH to homocysteine, methylation to methionine and finally adenylation to SAM. [24,25] (Fig. 7)

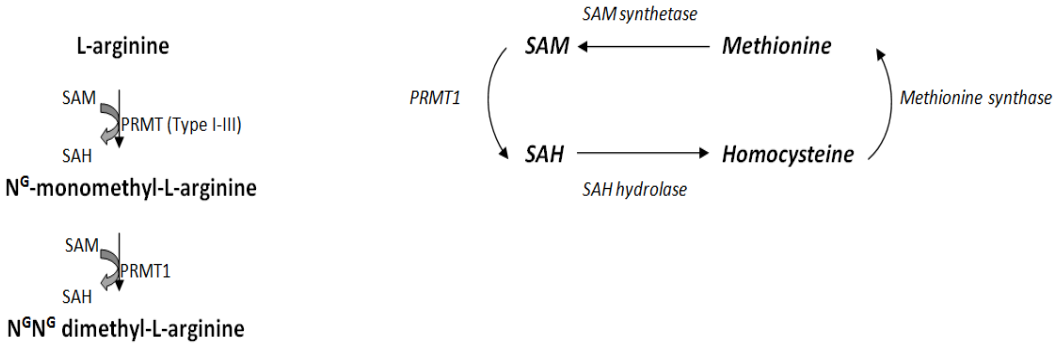


Fig. 7: Reaction scheme of arginine methylation and regeneration of SAM
L-arginine is monomethylated via different PRMTs and asymmetric dimethylated via PRMT1 using S-adenosyl-L-methionine (SAM) as methylgroup donor. SAM regeneration proceeds via S-adenosyl-homocysteine (SAH), homocysteine and methionine.

1.4. Arginine methylation regulates function of RGG/RG proteins

Protein-protein interactions can be positively or negatively influenced by methylation of arginine residues in RGG/RG proteins. For instance, SRC associated in mitosis of 68 kDa (Sam68) binding to the Src-like SH3 domain is drastically reduced upon arginine methylation. On the other hand, the Tudor domains of survival motor neuron protein (SMN) show increased affinity to methylated arginines. [15,26]

As discussed, hnRNPUL1 is involved in DNA damage signaling by interaction with NBS1. This interaction is dramatically reduced in case of hypomethylation. [20]

Despite the massive amount of biological data for RG/RGG proteins, the molecular mechanisms of these interactions are largely unknown.

1.4.1. Aberrant arginine methylation pattern in Homocysteinuria

Post-translational modifications such as arginine methylation specify proteins functions. Deregulation of these modification pathways may lead to a deficiency of different physiological functions which can result in the development of diseases

An example for a deregulation of the arginine methylation pathway is the neurological disorder Homocysteinuria. It is a metabolic disease characterized by mentally retardation, skeletal abnormalities, vascular disease and lens dislocation, due to intracellular accumulation of L- homocysteine. [27,28] This molecule is an intermediate of the S-adenosyl-L-methionine regeneration pathway of arginine methylation. An accumulation of downstream products in a metabolic pathway can result in a negative feedback due to enzyme inhibition. [29]

The cause of the disease is a deficiency of the enzyme Cystathione ß-synthase (CBS), which catalyzes the conversion of homocysteine to cystathione, the precursor of cysteine.[30] This leads to increasing homocysteine and S-adenosylhomocysteine levels in tissues. In mouse models ($CBS^{-/-}$) the CBS deficiency affects the global protein arginine methylation status (decrease of 10-35%). The hypomethylation of arginine residues may promote different pathophysiological changes causing neurodegeneration. [31]

1.4.2. Misregulation of PRMT1-mediated arginine methylation is involved in the development of ALS as well as FTD

The protein Fused-in-sarcoma/Translocated-in-Liposarcoma (FUS/TLS) is a member of the family of heterogeneous ribonucleoproteins (hnRNPs) and it has recently been reported that a deregulation of its arginine methylation pattern is a trigger for neurodegenerative diseases. [4,6] FUS is involved in different biological processes, such as transcription, splicing and microRNA processing; therefore the nuclear localization of this protein is inevitable. Nuclear import of hnRNPs is mainly mediated by the karyopherinß2 transportin-1. [32–37]

Mutant variants of the protein FUS are linked to familial amyotrophic lateral sclerosis (fALS), which is a motor neuron disease characterized by degeneration of neurons in the brainstem and spinal cord.[38] It is occurring in either a sporadic or a mutation dependent form; four percent of inherited fALS are caused by mutations in the C-terminal region of the FUS gene, affecting the proline-tyrosine nuclear localization signal (PY-NLS) [34,39].

Inclusions of the FUS protein in the cytoplasm of neurons and a reduced FUS occurrence in the nucleus are a hallmark of this disease.[38,40]

Frontotemporal dementia (FTD) is a prevalent neurodegenerative disease characterized by changes in behavior and personality. Causes for these symptoms are protein depositions of a variety of different proteins in the central nervous system, including FUS protein. In contrast to fALS, this medical condition does not comprise cytoplasmic inclusions of mutant FUS, but the wildtype protein. [41,42]

Although these two disorders share a common denominator, the deposition of FUS due to inefficient nuclear import, they have distinct disease mechanisms. Neuronal cytoplasmic inclusions in post mortem tissue of ALS patients contain a mutant form of FUS with asymmetrically dimethylated RGG motifs whereas the RGG motif of wildtype FUS in inclusions of FTLD patients is mono- or unmethylated.

The nuclear localization of the protein FUS is not only regulated by binding of the PY nuclear localization signal to transportin-1, but also by arginine methylation. Upon inhibition of PRMT1-dependent methylation of the arginine-glycine rich (RGG) region adjacent to a mutant PY (e.g. P525L) motif, the nuclear localization of the protein is restored. This RGG domain of FUS interacts tightly with transportin-1 and can compensate for a reduced binding affinity of the mutant PY-NLS in its non-methylated state. In FUS-FTLD patients a deficient methylation pattern in the wildtype protein is observed. In addition, cytoplasmic inclusions of patients contain transportin-1, which implicates an enhanced affinity of the non-modified arginine residues of FUS to the karyopherin. [4,6]

1.4.3. Members of RGG-motif harboring proteins involved in proliferative diseases

The family of RGG motif-harboring proteins comprises 1000 of different macromolecules. A variety of proteins in Homocysteinuria and FUS in ALS and FTD are linked to neurodegeneration, but there are also members which are associated to proliferative diseases.

Serine/arginine-rich splicing factor 1 (SRSF1) belongs to RNA-binding proteins, controls the splice site recognition and facilitates the export of spliced mRNAs. The nuclear translocation of this protein is, besides phosphorylation by cytosolic protein kinase SRPK,

regulated by arginine methylation of PRMT1 at residues R93, R97 and R109. SRSF1 is an oncoprotein and is up-regulated in different types of cancers. For instance, the expression level of SRSF1 as well as PRMT1 is increased in newly diagnosed and relapse bone marrow samples of patients with acute lymphoblastic leukemia (ALL). This may implicate an influence of inaccurate arginine methylation patterns in development of ALL.[7,43]

Twist-related protein 1 (TWIST1) is a transcription factor and together with PRMT1, it is involved in epithelial-mesenchymal transition (EMT). EMT is an important factor for invasion, progression and metastasis of tumors. The methylation of R34 in TWIST1 by PRMT1 may play an important role for the nuclear import of the protein. [8]

The nuclear localization of other RNA binding proteins, like SAM68 (SRC associated in mitosis of 68 kDa), is also regulated by PRMT1. SAM68 regulates mRNA metabolism and is involved in different signal transduction pathways. SAM68 is up-regulated in different types of cancer, like prostate and cervical cancer. [44]

1.5. Model system for arginine methylation

As discussed above, arginine methylation is an important post-translational modification in a large group of different proteins and a deregulation of this modification may lead to a plethora of diseases. Given the importance of these proteins in health and disease, the aim of this thesis was to study common molecular mechanisms among this large group of proteins in order to understand their biological function as well as mis-regulation in disease.

Therefore I chose a model system to investigate the arginine methylation or, more precisely, the transportin-1 and arginine methylation dependent nuclear import of the Cold Inducible RNA binding Protein.

1.5.1. Cold inducible RNA binding protein (CIRP) – a member of the RGG proteome

Arginine-glycine-rich sequence motifs (RGG-motifs) are favored regions for PRMT1-dependent arginine methylation of proteins. Cold inducible RNA binding protein (CIRP) is a member of the human hnRNPs harboring a C-terminal RGG motif. CIRP is down-regulated at elevated temperatures and up-regulated at low temperatures. The protein was first identified in male germ cells and facilitates the mRNA translation upon cold stress. [45–47] In addition, CIRP migrates to cytoplasmic stress granules upon cellular oxidative, osmotic or heat stress as well as in response to endoplasmic reticulum (ER) stress. Since CIRP is localized in the nucleus, it has to migrate to the cytoplasm to operate in stress granules. Furthermore arginine residues in the C-terminal arginine-glycine-rich region are important for the downregulation of mRNA translation. [5,48]

1.5.2. Knowledge about nuclear import of CIRP

There are many studies describing the role of CIRP in cold stress response, by migration to stress granules and by facilitating mRNA translation. [45–47] All of these processes implicate an inevitable nuclear transport of the protein. But the detailed mechanism still remains elusive.

CIRP is an RNA binding protein containing arginine-glycine rich regions as well as a RNA recognition motif (RRM). Screening of the sequence revealed an additional serine-tyrosine (SY) rich region on the C-terminus of CIRP. [49]

It has been reported that the RGG region of the protein CIRP is a target for PRMT1. A methylation of these arginine residues resulted in a cytoplasmic accumulation of CIRP: This arginine-glycine rich region of CIRP may be a putative transportin-1 substrate. [48]

In addition, the SY- region of CIRP may share common functions with FUS-PY and since it contains serine and tyrosine residues, which are putative phosphorylation sites, these post-translational modifications may be an additional regulatory factor.

1.6. Studying arginine methylation metabolism *in vivo*

1.6.1. Metabolic profiling

Arginine methylation is a metabolic pathway which regulates a variety of proteins having different functions. As discussed, several diseases could be linked to a mis-regulation in arginine methylation. [4,6–8,27,44] These deregulations in the metabolism could be analyzed via metabolic profiling focusing on small molecule metabolites associated with this pathway (Fig. 8).

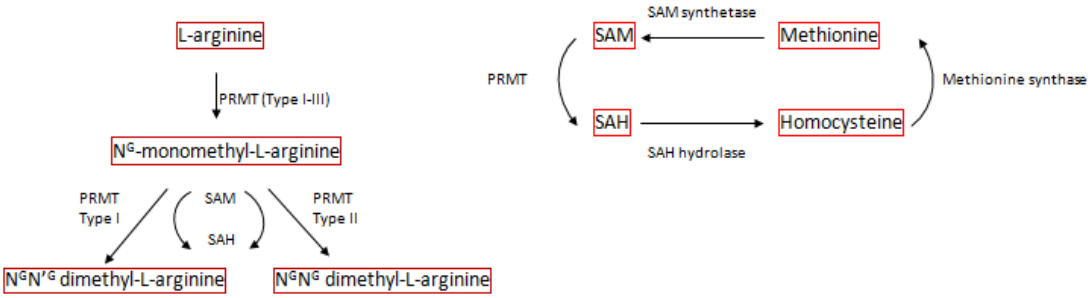


Fig. 8: Reaction scheme of arginine methylation and regeneration of SAM
L-arginine is monomethylated via different PRMTs and asymmetric dimethylated via PRMT1 using S-adenosyl-L-methionine (SAM) as methylgroup donor. SAM regeneration proceeds via S-adenosyl-homocysteine (SAH), homocysteine and methionine. Metabolites of interest are labeled.

Metabolic profiling is the measurement of the response to genetic modifications as well as different stimuli.[50] There are thousands of different metabolites in the human metabolome, which can be identified and quantified using techniques like mass spectrometry or nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy is a quantitative method; the integrals of peaks are directly proportional to the metabolite concentration. Furthermore NMR-based metabolomics is a non-destructive method with a straightforward sample preparation, including protein removal by methanol precipitation or ultrafiltration, and fast one dimensional measurements. In addition, two-dimensional measurements provide further information about metabolites and in case of J-resolved experiments, the complexity of spectra is reduced. A disadvantage of NMR-based metabolomics is a low sensitivity; metabolites should be in µM range.

Biofluids, including blood, urine and cerebrospinal fluid, as well as cell lysates and tissue lysates can be analyzed. [51]

1.6.2. Proof-of-principle metabolic profiling study

Analysis of the metabolome includes the measurement of small molecules in different biofluids, cell- or tissue lysates. This metabolic profile is important in different fields of biological sciences, including the development of promising diagnostic tools. A standard approach of metabolomics comprises a sample preparation, followed by the measurement, processing and analysis of data. These steps lead to identification of metabolites and a biological interpretation. This workflow is an untargeted approach and gives a profile of all metabolites present in the samples. [52]

Since arginine methylation is associated with several metabolic pathways and given that this post-translational modification is stable (due to the lack of an enzyme that removes the modification) we hypothesize that several related metabolites can be detected in biofluids and that these metabolites reflect the degree or arginine methylation. In fact, it has already been reported that metabolites associated with arginine methylation can be detected. Dhar et al. showed that a loss of PRMT1, which is the primary PRMT producing ADMA, results in global levels of methylated arginine. Upon decreased PRMT1 activity, the ADMA level is reduced, whereas SDMA and MMA levels accumulate. [53]

An increased amount of asymmetrically dimethylated arginine in biofluids may implicate an increased level of arginine methylation in members of the RG/RGG proteome and therefore a higher risk for cancer, whereas a reduced arginine methylation level may be a sign for neurodegeneration.

To this end, an experiment for metabolic profiling of biofluid samples was setup and tested for metabolites associated with arginine methylation metabolism. Furthermore, the setup was tested in a proof-of-principle experiment on a cohort containing control samples and samples of patients with multiple sclerosis. Serum and cerebrospinal fluid (CSF) samples of patients suffering from multiple sclerosis (MS) are analyzed via metabolic profiling and due to the untargeted approach, a general pattern of the metabolites is obtained.

MS is a chronic inflammatory disease of the central nervous system (CNS). Demyelination in the CNS leads to a wide range of neurological symptoms including muscle weakness and visual loss. The disease course is highly variable in patients and therefore four different clinical courses are distinguished: relapsing-remitting (RR), secondary progressive (SP), primary progressive (PP) and progressive relapsing (PR) multiple sclerosis. [54–57]

Since there is a plethora of neurological diseases, diagnosis of MS still remains difficult.

Demyelinating diseases including Neuromyelitis optica and Acute Disseminated Encephalomyelitis share common symptoms with multiple sclerosis.[58] To avoid misdiagnosis, McDonald Criteria for diagnosis of multiple sclerosis are followed. These criteria involve the evidence of demyelinating lesions disseminated in time and in space (DIT and DIS), using Magnetic Resonance Imaging (MRI) and measuring CSF parameters. However, although the McDonald Criteria specify diagnostic parameters, a slight uncertainty still remains and diagnosis is made because there is no better explanation. [55,58]

Due to the absence of a highly reliable and fast diagnostic tool for multiple sclerosis patients a lot of effort is made in the field of biomarker research. Different biofluids of patients as well as of animal models (e.g. serum, cerebrospinal fluid (CSF) and urine) are screened for accumulation or lack of different metabolites. These aberrant metabolite patterns may shed light on metabolic pathways which are affected in multiple sclerosis.

Different studies of urine metabolite analysis already indicated an alteration of a variety of different metabolites such as acetate, creatinine, hippurate, 3-hydroxybutyrate, malonate and oxaloacetate. These metabolites are involved in metabolic pathways of energy metabolism, fatty acid synthesis and gut microflora The energy consumption of the brain is quite high (25% of available glucose and ketone bodies as alternative energy source) and since the brain chemistry is changed in MS it may explain the effect on the metabolites in energy metabolism. In experimental autoimmune encephalomyelitis (EAE) mice the level of a metabolite from the gut microflora, indoxyl sulfate, is influenced. Given that gut microorganisms are involved in induction and progression of EAE, it could also be associated with MS.[59,60] Serum and CSF analysis also identified an aberrant energy, fatty acid and tryptophan metabolism as well as nitric oxide synthesis and levels of endogenous antioxidants in multiple sclerosis patients. [61–63]

A relapsing remitting-EAE mouse model identified primary bile acid biosynthesis, tryptophan, D-arginine and D-ornithine metabolism as altered pathways in relapsing remitting MS. Dickens et al (2014) identified a type 2 biomarker that distinguishes RR from SP multiple sclerosis, which may implicate that metabolic phenotyping could enable not only a diagnosis of multiple sclerosis but also a differentiation of the different clinical courses of the disease. This would shorten the time from first symptoms to diagnosis and highly improve the therapeutic success [64,65]

1.7. Objectives

The central aim of my thesis is to provide insights into key molecular mechanisms causing neurodegenerative diseases. Based on recent findings [4,6], arginine methylation is an important factor for the development of neurodegenerative diseases (i.e. FTD, ALS). A hypothesis is that these mechanisms can be translated to the entire family of RG/RGG proteins. The RGG proteome comprises a variety of more than 1000 different members which contain arginine-glycine-rich motifs that are preferred targets for PRMT1-dependent methylation. The nuclear import of all these members may be regulated by the same mechanism and understanding of this mechanism will provide new insights into the development of neurodegenerative diseases.

PRMT1 is known to be the main enzyme capable of producing asymmetric dimethylated arginine residues in proteins. [2,6,7,20,66,67] This enzyme specifically binds to substrates rich in Phe, Arg and Gly (Fig. 9) and transfers methylgroups to arginine residues in the sequence. PRMT1 has a highly conserved methyltransferase domain, whereas the N-terminal region is very variable and therefore an essential factor for substrate specificity. Nevertheless, all the different isoforms of PRMT1 bind to and post-translationally modify a similar set of proteins. [68,69]

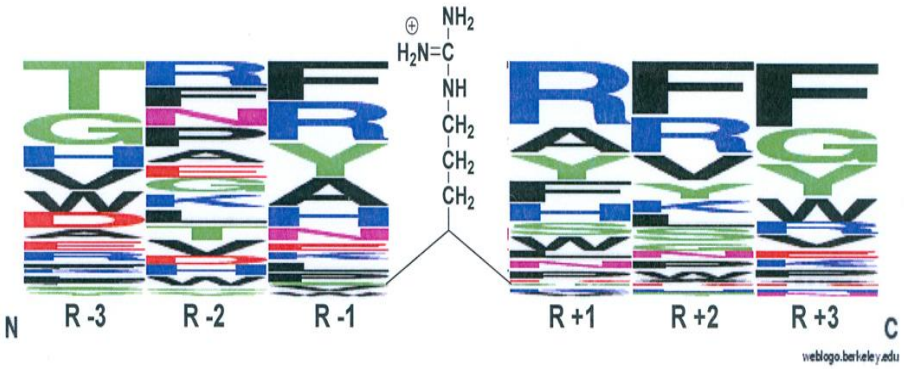


Fig. 9: Substrate specificity of PRMT1[68]

The broad perspective hypothesis is that a deregulation of arginine methylation is involved in the development of neurodegenerative as well as proliferative diseases. An up-regulation of different RGG motif containing proteins proteins (i.e. TWIST1, SRSF1) as well as PRMT1-dependent methylation are known to be involved in different types of cancer, whereas other RGG-proteins (i.e. FUS) are hypomethylated in neurodegenerative

diseases like FTD. [6–9,44] The mis-regulation of this post-translational modification may entail profound consequences, in a neurodegenerative way for hypomethylation and in a proliferative way for hypermethylation (Fig. 10).

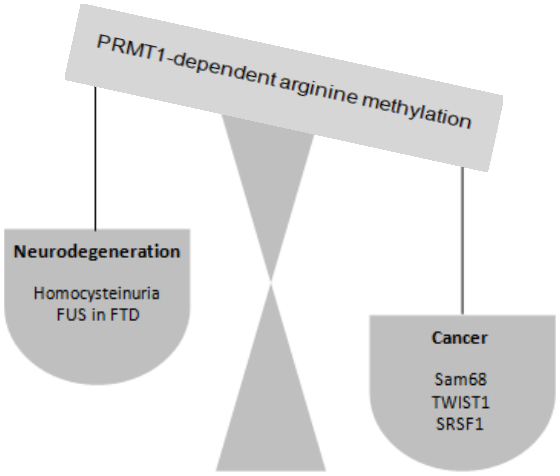


Fig. 10: Hypothesis of misregulated arginine methylation in diseases
Hypomethylation of proteins may lead to neurodegenerative diseases like homocysteinuria or FTD, hypermethylation of proteins (i.e. Sam68, TWIST1, and SRSF1) to cancer.

The aim of this thesis was to characterize the mechanisms of regulatory nuclear import of RGG proteins using the model system CIRP (Cold-inducible RNA binding protein) in order to study the hypothesis that the transportin-1 dependent nuclear import is a key regulatory mechanism for the entire RG/RGG proteome. Since CIRP harbors another sequence motif, the SY-region the role of this region in terms of nuclear import was analyzed. This motif contains several putative serine and tyrosine phosphorylation sites, therefore phosphorylation may regulate the interaction with transportin-1. Thus, first I carried out phosphorylation assays in cell lysate and localized the phosphorylation sites. Since phoshporylation was detected, mutants were designed for binding studies of phosphorylated CIRP to study the molecular mechanisms of regulatory phosphorylation on the CIRP-transportin-1 complex. Furthermore, I aimed to study the atomic details of the CIRP-transportin-1 interaction by solving the structure of the CIRP-transportin-1 complex in order to further translate my findings to the entire RG/RGG proteome.

These findings were applied to a model system and samples of patient cohort studies to detect metabolites linked to the arginine methylation metabolism. A setup for metabolic profiling of biofluid samples was tested on a cohort. This proof-of-principle experiment for NMR-based metabolomics enables a detection of metabolites of the arginine methylation metabolism.

Since the approach of NMR-based metabolomics is untargeted, it allows a profiling of the entire metabolome; many more metabolites, including possible biomarkers for diagnosis and monitoring of different diseases, can, in principle, be identified and quantified.

2. Materials and Methods

2.1. Buffers and Solutions

Gel buffer for SDS PAGE (3 x)

1 M Bis (2-hydroxyethyl)aminotris(hydroxymethyl)methane (Sigma 98%)

pH 6.5-6.8 adjusted with HCl (12.2 M) and NaOH (5 M)

MES SDS running buffer (20 x)

1 M MES (Sigma Aldrich >99%)

1 M Trizma® base (Sigma Aldrich >99.8%)

2 % (w/v) SDS (Applichem Panreac 10 %)

20 mM EDTA (Alfa Aesar 99+%)

No need to adjust pH

Coomassie Staining solution

40% desalted H_2O

40% MeOH (VWR 99.9%)

10 % glacial acetic acid (Roth >99%)

Tip of a spatula of Coomassie R-250 dye powder (Thermo Fischer)

Coomassie Destaining solution

40% desalted H_2O

40% MeOH (VWR 99.9%)

10 % glacial acetic acid (Roth >99%)

Kanamycin (1000 x; 50 mg/ml)

5 g Kanamycin (Applichem Panreac) in 100 ml of MilliQ H_2O

Filter sterilize

(Selective) LB medium

2 % (w/v) lysogeny broth (Roth)

(+ 0.1 % (v/v) Kanamycin (1000 x))

(Selective) LB Agar plates

2 % (w/v) lysogeny broth (Roth)

2 % (w/v) agar (Amresco pure, bacterial grade)

(+ 0.1 % (v/v) Kanamycin (1000 x))

10 x Salt

1 M KH_2PO_4 (VWR 99.8%)

0.5 M K_2HPO_4 (VWR 99.8%)

0.6 M Na_2HPO_4 (Applichem Panreac anhydrous >99%)

0.14 M K_2SO_4 (VWR 99-101%)

pH 7.2 adjusted with HCl (12.2 M) and NaOH (5 M)

Trace element solution

41 mM $CaCl_2$ x 2 H_2O (Applichem Panreac 97%)

22 mM $FeSO_4$ x 7 H_2O (VWR 84%)

6 mM $MnCl_2$ x 4 H_2O (Applichem Panreac)

3 mM $CoCl_2$ x 6 H_2O (Applichem Panreac)

1 mM $ZnSO_4$ x 7 H_2O (VWR 99.9%)

0.1 mM $CuCl_2$ x 2 H_2O (VWR)

0.2 mM $(NH_4)_6Mo_7O_{24}$ x 4 H_2O (VWR 98.5%)

17 mM EDTA (Alfa Aesar 99+%)

Minimal medium

It was used for production of labeled or non-labeled proteins. In case of labeling, ^{15}N and/or ^{13}C isotopes were used.

For expression of 13C-methyl-ILVM transportin-1, 100 mg of 2-Keto-3-methylbutyric acid 5-^{13}C (Sigma), 50 mg of 2-Ketobutyric acid 4-^{13}C (Sigma) and 50 mg of L-methionine-methyl-^{13}C (Cambridge Isotope Laboratories) was added before induction.

900 ml MilliQ H_2O
100 ml 10 x Salt
5 ml Trace element solution
5 ml 1 M $MgCl_2$ (Alfa Aesar 99%)
1g $^{15}NH_4Cl$ (Sigma Aldrich 98%)/2 g $^{14}NH_4Cl$ (Alfa Aesar 99.5%)
2g $^{13}C_6H_{12}O_6$ (Cambridge Isotope Laboratories 99%)/
6 g $^{12}C_6H_{12}O_6$ (Roth 99%)
0.1% (w/w) Kanamycin (1000x)

TFB1

30 mM KCH_3COOH (VWR 99-101%)
10 mM $CaCl_2$ x $2H_2O$ (Applichem Panreac 97%)
50 mM $MnCl_2$ x $4H_2O$ (Applichem Panreac)
100 mM RbCl (Amresco biotechnology grade)
15 % glycerol (VWR 99%)
pH 5.8 adjusted with 1 M acetic acid
filter sterilize

TFB 2

10 mM MOPS (Roth)
75 mM $CaCl_2$ x $2H_2O$ (Applichem Panreac 97%)
10 mM RbCl (Amresco biotechnology grade)
15% glycerol (VWR 99%)

Chook buffer (lysis buffer for structured proteins)

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
20 % glycerol (VWR 99%)
20 mM imidazole (Millipore 99.0%)

Urea buffer (lysis buffer for disordered proteins)

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
6 M Urea (Roth >99.5%)
20 M imidazole (Millipore 99.0%)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

Buffer A

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
20 mM imidazole (Millipore 99.0%)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

Buffer A'

1 M NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
20 mM imidazole (Millipore 99.0%)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

Buffer B

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
200 mM imidazole (Millipore 99.0%)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

Buffer B’

1 M NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
500 mM imidazole (Millipore 99.0%)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

ITC buffer

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
2 mM TCEP (Roth 98%)
pH 7.5 adjusted with HCl (12.2 M) and NaOH (5 M)

NMR buffer

150 mM NaCl (VWR 99.6%)
50 mM Tris (VWR ultra pure)
2 mM TCEP (Roth 98%)
pH 6.7 adjusted with HCl (12.2 M) and NaOH (5 M)

Methylation buffer

50 mM Na_2HPO_4/NaH_2PO_4 (Applichem Panreac 98%)
150 mM NaCl (VWR 99.6%)
2 mM DTT
pH adjusted to 8.0 with HCl (12.2 M) and NaOH (5 M)

Lysis buffer for HEK293 cells

50 M Tris (VWR ultra pure)
150 mM NaCl (VWR 99.6%)
5 mM $MgCl_2$
2 mM TCEP (Roth 98%)
1% TritonX®-100 (Roth)
1 (v/v) % Proteaseinhibitor
(Protease-Inhibitor Mix HP - Serva)
1 tablet Phosphataseinhibitor (PhosSTOP™ - Roche)/10 ml

Crystallization buffer

20 mM HEPES (Pufferan® Roth 99.5%)

110 mM KCH_3COOH (VWR 99-101%)

2 mM $Mg(CH_3COOH)_2$

1 mM EGTA (Roth 99%)

2 mM DTT (Roth >99%)

pH adjusted to 7.3 with HCl (12.2 M) and NaOH (5 M)

Reservoir solution

1M succinic acid (Roth >99%)

0.1 M HEPES (Pufferan® Roth 99.5%)

1 % (w/v) PEG MME 2000 (Sigma Aldrich)

pH adjusted to 7.0 with HCl (12.2 M) and NaOH (5 M)

NMR metabolomics buffer

70 mM Na_2HPO_4 (Applichem Panreac 98%)

0.025% v/v NaN_3 (Merck 99%)

0.8% w/v TSP(Alfa Aesar 98%)

pH 7.4 adjusted with HCl (8 M) and NaOH (5 M)

+ ^2H_2O

2.2. Model system for arginine methylation

2.2.1. SDS PAGE

For resolving and stacking gels a 3x gel buffer (1 M BisTris; pH 6.5-6.8 adjusted with HCl (12.2 M) and NaOH (5 M)) was prepared.

Resolving gel: 1/3 volume 3x gel buffer
12-15 % acrylamide/bisacrylamide 37.5:1 (40%; Serva)
Add desalted H_2O to final volume
25 µl APS (10%) (Emsure® 98%)
10 µl TEMED (Roth >99%)

Gel was poured to the appropriate level, covered with water and polymerized.

Stacking gel 1/3 volume 3x gel buffer
5 % acrylamide/bisacrylamide 37.5:1 (40%; Serva)
Add desalted H_2O to final volume
20 µl APS (10%) (Emsure® 98%)
10 µl TEMED (Roth >99%)

Comb was inserted and polymerized.

20 x MES SDS running buffer (1 M MES, 1 M Tris base, 2 % (w/v) SDS, 20 mM EDTA; no need to adjust pH) was diluted to 1 x with desalted H_2O and filled in gel chamber.
Samples were mixed with Laemmli Loading dye in a ratio of 1:6 and heated for 5-10 min to 95 °C. A constant voltage of 100 V was applied for approximately 1 hour.

Coomassie staining was used to visualize proteins on the SDS gel. The gel was incubated with Coomassie staining solution (40% desalted H_2O, 40% MeOH, 10 % glacial acetic acid, tip of a spatula of Coomassie R-250 dye powder) for one hour and destained in destaining solution (40 % desalted H_2O, 40 % MeOH, 10 % glacial acetic acid) over night shaking.

2.2.2. Production of competent *E.coli* cells

2.5 ml of LB medium were inoculated with 1 µl of competent cells (*E.coli* BL21 DE3, *E.coli* pLysS, *E.coli* Top10 – New England BioLabs) and grown overnight at 37 °C with 200 rpm shaking. The overnight culture was used to inoculate 250 ml of LB medium containing 20 mM $MgSO_4$ in an one liter flask. This culture was grown to an optical density 600 of 0.4 – 0.6 and harvested by centrifugation at 4500 rpm for 5 minutes at 4 °C. The supernatant was removed and the pellet was resuspended in 100 ml of ice cold TFB1 (30 mM potassium acetate, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15 % glycerol; pH adjusted to 5.8 with 1 M acetic acid and filter sterilized). The cells were incubated on ice for 5 minutes and the remaining steps were made on ice as well. Cells were pelleted at 4500 rpm for 5 minutes at 4 °C, supernatant was removed and pellet of *E.coli* cells was resuspended in 10 ml of ice cold TFB2 (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% glycerol). Cellsolution was incubated on ice for 15 – 60 minutes and aliquoted to prechilled tubes. Tubes were snap frozen in liquid N_2 and stored at – 80°C.

2.2.3. Transformation

DNA constructs were obtained from Genescript®. Genes of interest were cloned in a modified pETM11 vector containing a T7 promotor, a N-terminal Polyhistidine Tag, a protein A tag as well as a Tobacco Etch Virus (TEV) cleavage site for purification (Z-Tag).
50 ng of circular DNA was added to competent *E.coli* DE3 or *E.coli* pLysS, respectively. After an incubation step of 10 min on ice, a heat shock of 42°C for 45 sec was applied in an Eppendorf thermoblock. For regeneration, cells were incubated for one hour in 1 ml of LB media without an antibiotic selection marker at 37°C shaking in an Eppendorf thermoblock. Competent *E.coli* were spun and the pellet plated on selective LB agar plates. Growth followed overnight at 37°C in an incubator.

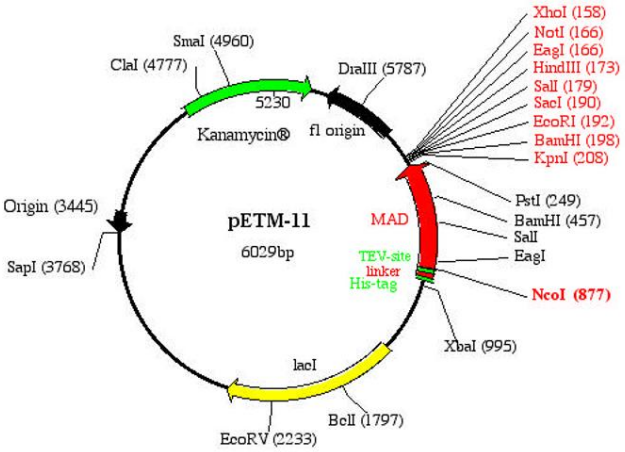


Fig. 11: pETM-11 vector for expression of proteins containing Z-Tag in *E.coli*[80]

2.2.4. Preparation of glycerol stocks

Glycerol stocks of transformed *E.coli* cells were prepared by mixing 1 ml of overnight culture (LB media + Kanamycin for selection; inoculated with one colony of transformation plate) with 1 ml of 50% glycerol. Cells are frozen at -80°C.

2.2.5. Protein expression in *E.coli* strains

An overnight culture (10 ml of LB media + Kanamycin for selection; inoculated with glycerol stock) of *E.coli* BL21 DE3 with the genetic information of the protein of interest on a plasmid containing a selection marker was used to inoculate a main culture (1 L LB media or minimal medium + Kanamycin for selection). When grown to an OD_{600} (optical density at 600 nm) of approximately 0.8 at 37°C, the protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranosid (IPTG). For transportion-1 constructs, the main culture was grown to a high OD_{600} at room temperature for two days and diluted to OD_{600} of 0.8. The induced main culture was incubated at 16°C – 20°C shaking for 16-20 hours. For smaller proteins (e.g. CIRP-SY, FUS-PY) the expression was performed for 4-5 hours at room temperature. After expression of proteins, cells were harvested by centrifugation (15 min, 6000 rpm, 4°C). Pelleted cells can be stored at -20°C or -80°C.

2.2.6. Lysis of *E.coli* cells

E.coli cells were lysed to isolate the recombinant expressed proteins. The pelleted cells were thawed on ice and resuspended in 20 ml of Lysis buffer (Tab. 1). RNA contamination of CIRP full length was avoided by adding 20 µl of RNase A (1 µg/ml; Thermo Fisher) and 10 mM $MgCl_2$. Cells were sonicated for 15 minutes (1 sec on/1 sec off; 50 % intensity) on ice and centrifuged for 30-45 minutes (12 000 rpm). The supernatant should contain all the soluble cellular protein.

2.2.7. Protein purification

In the first step proteins were purified via HisTrap FF on ÄktaFPLC system (5 ml, GE Healthcare; for PRMT1) or Ni-NTA (Ni-NTA Agarose as stationary phase, Thermo Fisher). The purification is based on binding of His-tagged proteins to Ni^{2+} residues in the column. The column was equilibrated with Buffer A (approximately 25 ml) and the supernatant of lysed *E.coli* cells was applied on the column. The flow through was collected and applied a second time to attain a higher yield. Non-specific bound proteins were removed by washing the column with approximately 50 ml of different washing buffers. The elution of non-specific proteins was monitored using NanoDrop™ (Peqlab), Bradford Proteinassay or absorption at 280 nm on Äkta systems (GE Healthcare). The protein of interest was eluted with 10-15 ml of Elution buffer. Washing and elution protocols for the different constructs are shown in Tab. 1. PRMT1 was eluted from the HisTrap FF using a gradient of elution buffer (100% in 15 min).

Tab. 1: Washing and elution protocols for different constructs

Protein	Washing step				Elution
CIRP-SY	Urea buffer	Buffer A	Buffer A'	Buffer A	Buffer B
FUS-PY	Urea buffer	Buffer A	Buffer A'	Buffer A	Buffer B
CIRP-RGG	Urea buffer	Buffer A	Buffer A'		Buffer B'
TNPO1Δlinker	Buffer A				Buffer B
TNPO1	Buffer A				Buffer B
CIRP-FL	Buffer A				Buffer B
PRMT1	Buffer A				Buffer B'

2.2.8. TEV cleavage

To avoid a non-native conformation of the protein, the tag was removed. The concentration of the eluted protein was determined using NanoDrop^{TM} (Peqlab; see 2.2.10.) and 2 (w/w) % of TEV protease was added to the protein solution. The sample was incubated for at least 4 hours at 4 °C for cleavage.

2.2.9. Buffer exchange of purified proteins

The buffer exchange of purified proteins was either performed with PD10 columns (GE Healthcare) or using Äkta pure FPLC system (GE Healthcare) with Desalting columns.
PD10 columns were equilibrated using 25 ml of the new buffer. 2.5 ml of protein sample were applied and proteins were eluted with 3.5 ml of new buffer.
Desalting columns (HiTrap Desalting 1 ml or HiPrep 26/10 Desalting 50 ml, GE Healthcare) were equilibrated with 5 ml or 60 ml of new buffer, respectively at a flow rate of 5 ml/min. Volumes from 1 ml to at maximum 15 ml were applied and eluted fractions were collected depending on the UV absorbtion at 280 nm measured by Äkta system detector.

2.2.10. Determination of concentration

Concentration of purified proteins was determined using Nanodrop (Peqlab) for measurement of absorption at 280 nm. 2 µl of buffer served as a blank and 2 µl of protein were applied for measurement of absorption. According to the law of Lambeert Beer the absorption at 280 nm was divided by the molar extinction coefficient ε of the protein to calculate the molar concentration.

Tab. 2: Protein constructs, molecular weights and extinction coefficients

Construct	UniProt ID	Residues	Molecular weight [Da]	Extinction coefficient (ε; 280 nm)
CIRP-SY	Q14011	138-172	3949	7680
CIRP-RGG	Q14011	68-137	7247	3840
CIRP-FL	Q14011	1-172	18648	12800
TNPO1	Q92973	1-898	102355	102130
TNPO1Δlinker	Q92973	1-898 (depletion 328-381)	96213	102130
FUS-PY	P35637	502-526	2960	1280
PRMT1	Q63009	11-353	39576	51490

2.2.11. Concentration of proteins using Amicons

To obtain higher protein concentrations, Amicon Ultra Centrifugal Filters (Millipore) with a cut off of 10 kDa or 3 kDa were used. Centrifugation steps of 5 – 10 min at 4 °C at 3500 rpm were followed by mixing of the protein solution in the filter to avoid precipitation of the protein due to the concentration gradient. Since CIRP-SY tends to produce a jellylike solution at lower temperatures, it was concentrated at room temperature.

2.2.12. Size exclusion chromatography

Size Exclusion Chromatography (SEC) was performed to get rid of contaminants for ITC and NMR measurements. Size exclusion was performed on Äkta pure FPLC system using Superdex 200 16/600 (GE Healthcare) at a flow rate of 1 ml/min for large transportin-1 constructs, and Superdex peptide 10/300 GL (GE Healthcare) at a flow rate of 0.5 ml/min for CIRP-constructs. Sample volumes of 500 µl or 4 ml, respectively, were applied on Superdex 200 16/600 or Superdex peptide 10/300 GL.

2.2.13. PRMT1-dependent arginine methylation

For the methylation reaction the proteins were exchanged in methylation buffer (50 mM sodiumphosphate pH adjusted to 8.0 with NaOH and HCl, 150 mM NaCl and 2 mM DTT) (see 2.2.9.). 50 µM of CIRP-RGG was mixed with PRMT1 and SAM at a final concentration of 7 µM and 1 mM, respectively. Solutions of proteins at lower concentrations were concentrated using Amicons (see 2.2.10.). The methylation reaction was performed overnight at room temperature.

2.2.14. Isothermal titration calorimetry (ITC)

Binding affinities of CIRP-RGG or FUS-PY to TNPO1 or TNPO1Δlinker were determined using ITC on a VP-ITC Microcal calorimeter (Microcal, Northhampton, USA) at 25°C and for CIRP-SY to TNPO1 or TNPO1Δlinker at 10°C with 36 injections of 8 µl (first injection 4 µl). All proteins were dissolved in ITC buffer (50 mM Tris pH 7.5 adjusted with HCl and NaOH, 150 mM NaCl, 2 mM TCEP). Transportin-1 constructs were in the cell at a concentration of 30 µM and CIRP-SY at 300 µM. For CIRP-RGG and FUS-PY (100 µM), transportin-1 was at a concentration of 10 µM. For competition experiments, CIRP-RGG, CIRP-SY and FUS-PY were at a concentration of 50 µM in the cell, to have a saturated transportin-1 complex.

The ITC data were analyzed with the program MicroCal Origin software version 7.0 and single site binding model.

2.2.15. In Lysate - Phosphorylation assay of $^{15}N^{13}C$ labelled SY

HEK293 cells were lysed using HEK293 lysis buffer (50 mM Tris, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM TCEP, 1 % Triton, Proteaseinhibitor, Phosphataseinhibitor). After an incubation period of 30 minutes at room temperature, the lysate was spun for 30 minutes at 4°C. For the phosphorylation assay, HEK cell lysate with a final protein concentration of 15 mg/ml was mixed with 100 µM protein of interest (tagged CIRP-SY after elution of Ni-NTA) and 10 mM $MgCl_2$ and 10 mM ATP were added. The phosphorylation mixture was incubated overnight at room temperature.

Since the CIRP-SY is very temperature stable, a heatshock at 95 °C for 10 minutes served to precipitate all the HEK293-lysate protein. The lysate was centrifuged for some minutes and the supernatant was diluted in Buffer A and applied on a Ni-NTA equilibrated with 25 ml of Buffer A. After a washing step with buffer A and buffer A', the protein was eluted.

2.2.16. Quick Change PCR mutagenesis for phosphorylation mimic

Following serine residues were exchanged to glutamate residues for phosphorylation mimic:

^{136} MGSRDYYSSRSQSGGYSDRSSGGSYRDSYDSYATHNE ^{172}

Primers were designed using online tool http://nebasechanger.neb.com.

Tab. 2: Primer for CIRP-SY mutagenesis

Mutation	Primer	Annealing Temp. [°C]
S146E	Forward: 5'-CAGCTCTCGCGAACAGTCGGGCG-3' Reverse: 5'-TAATAATCACGACTGCCCATGG-3'	66
S148E	Forward: 5'-TCGCTCACAGGAAGGCGGTTATTA-3' Reverse: 5'-GAGCTGTAATAATCACGAC-3'	59
S155E	Forward: 5'-TTCCGACCGTGAATCCGGCGGTAG-3' Reverse: 5'-TAACCGCCCGACTGTGAG-3'	67
S159E	Forward: 5'-TTCCGGCGGTGAATACCGCGATAG-3' Reverse: 5'-CTACGGTCGGAATAACCG-3'	64

For the 50µl PCR the reagents were mixed as follows:

10 µl	Q5 High Fidelity DNA Polymerase Buffer (10x)
1 µl	dNTPs (10 mM)
2.5 µl	Forward Primer (10 µM)
2.5 µl	Reverse Primer (10 µM)
0.25 µl	Q5 High Fidelity DNA Polymerase (2 U/µl)
1 µl	DNA Template (20 ng/µl)
32.75 µl	Nuclease-free water

Cycling conditions:

Initial Denaturation (98°C)	2 minutes	
Denaturation (98°C)	30 seconds	}
Annealing (primer-dependent)	30 seconds	} 25 cycles
Extension (72°C)	4.5 minutes	
Final Extension (72°C)	2 minutes	
Hold (10°C)	∞	

1 µl of DpnI enzyme was added into the PCR tube and mixed by pipetting. After incubation at 37°C for one hour, 2 µl of the PCR reaction were mixed with 15 µl of Nuclease-free water and heated for 20 minutes at 80°C. 2 µl of 10X T4 Ligase buffer (NEB) and 1 µl of T4 polynucleotide Kinase (PNK-NEB) were added to the chilled tube and incubated for 30 minutes at 37°C. After a ligation step of 2 hours at room temperature using 1 µl of T4 Ligase (NEB), the plasmids were transformed into competent high copy *E.coli* cells (DH5α) and constructs were sent for sequencing to GATC biotech.

2.2.17. Protein crystallization for X-ray crystallography

TNPO1Δlinker (transportin-1 construct lacking the flexible loop region) as well as CIRP-SY were transferred into crystallization buffer individually and as a complex using size exclusion chromatography (2.2.12.). Crystallization was performed using Vapor diffusion in hanging drop format. This means a drop of protein solution (complex of TNPO1Δlinker

/CIRP-SY or 10 - 20 times excess of CIRP-SY to TNPO1Δlinker) in crystallization buffer was mixed with reservoir solution on a cover slip, which was placed inverted onto the reservoir containing 500 µl reservoir solution. Crystallization was checked after some days. Crystals were frozen in liquid nitrogen.

Tab. 3: Volumina in µL for crystallization of CIRP-SY/ TNPO1Δlinker

CIRP-SY (1.5 mM)	TNPO1Δlinker (100 µM)	CIRP-SY/ TNPO1Δlinker complex	Reservoir solution
0.2	1.2	/	1.5
0.4	1.2	/	1.5
/	/	1.2	1.2

2.2.18. Two dimensional NMR measurements

Samples for NMR titrations contained 50 µM of ILVM-labeled transportin-1, CIRP-FL, CIRP-SY, CIRP-RGG or FUS-PY in NMR buffer (150 mM NaCl, 50 mM Tris, 2 mM TCEP, pH 6.7 adjusted with HCl (12.2 M) and NaOH (5 M)). Stochiometric ratio (transportin-1:ligand) was 1:1. $^{15}N^{13}C$-CIRP-SY for HN(CO)CACB was at a concentration of 300 µM in NMR buffer (150 mM NaCl, 50 mM Tris, 2 mM TCEP, pH 6.7 adjusted with HCl (12.2 M) and NaOH (5 M)).

10 (v/v) % $^{2}H_2O$ was added for the lock signal. NMR measurements of proteins were recorded on 900 MHz NMR spectrometer (Bruker) at 298 K. Pulses of the nuclei were calibrated before each measurement.

Spectral changes were monitored by $^{13}C^{1}H$ HSQC for ILVM labeled transportin-1 and by $^{15}N^{1}H$ HSQC or HN(CO)CACB for $^{15}N^{13}C$ labeled CIRP-SY.

Overlays of spectra were made using software Sparky 3 (NMR Assignment and Integration Software).

2.3. Analysis of arginine methylation metabolites in human serum samples

2.3.1.1. Sample collection

Biofluid samples of patients and healthy controls (n = 30) were collected via vene- and lumbar puncture at University Hospital Graz, Department Neurology. Patients suffer from Multiple Sclerosis (MS; n = 8) or Clinically Isolated Syndrome (CIS; n = 22). The mean age at lumbar puncture for patients and healthy controls was 31.6 and 35.0 years, respectively. The median disease duration was 0.1 year and the age at disease onset was 30.5. 96.7 % (n = 29) of patients were oligoclonal band positive.

2.3.2. Sample preparation for NMR metabolomics

Biofluids, like serum or CSF, and cell lysates need to be processed in order to reduce undesired signals of proteins. For methanol precipitation of proteins, a volume of 200 µl MeOH (- 20°C) was added to a volume of 100 µl of biological sample. The samples were stored at -20°C for at least one hour. A centrifugation step of 30 min (4°C, 13 000 rpm) removed potential debris and the supernatant was mixed with 500 µl of NMR metabolomics buffer. After a lyophilization step, the sample was diluted in 500 µl D_2O and transferred into 5 mm NMR tubes.

2.3.3. Protein hydrolysis for NMR metabolomics

In order to analyze the arginine methylation (especially asymmetric dimethylation) in protein bound arginine residues of biofluids the samples were hydrolyzed with 9 M HCl for 18 hours (110°C). The solution was neutralized with 2 M Na_2CO_3 solution until CO_2 production stopped. The supernatant was lyophilized and resuspended in 500 µl of methanol to reduce high salt concentrations. These samples were centrifuged (30min, 13000 rpm) and the supernatant was lyophilized.

2.3.4. NMR measurement

One-dimensional ^{1}H NMR measurements of patients' samples were recorded on a Bruker 500 MHz spectrometer at 310 K to get a metabolic profile. Chemical shifts are reported in parts per million. TSP was used as internal standard for chemical shift referencing (set to 0 ppm). Carr- Purcell- Meiboom-Gill (CPMG) sequence was used for 1D ^{1}H measurements. For each sample, 1024 scans were made for an appropriate signal-to-noise ratio. In addition, two-dimensional, homonuclear ^{1}H J-resolved spectra were recorded to reduce spectral complexity. Samples in 5 mm tubes were measured in a 5 TXI probe (Tab. 4)

Tab. 4: Setup for NMR measurements

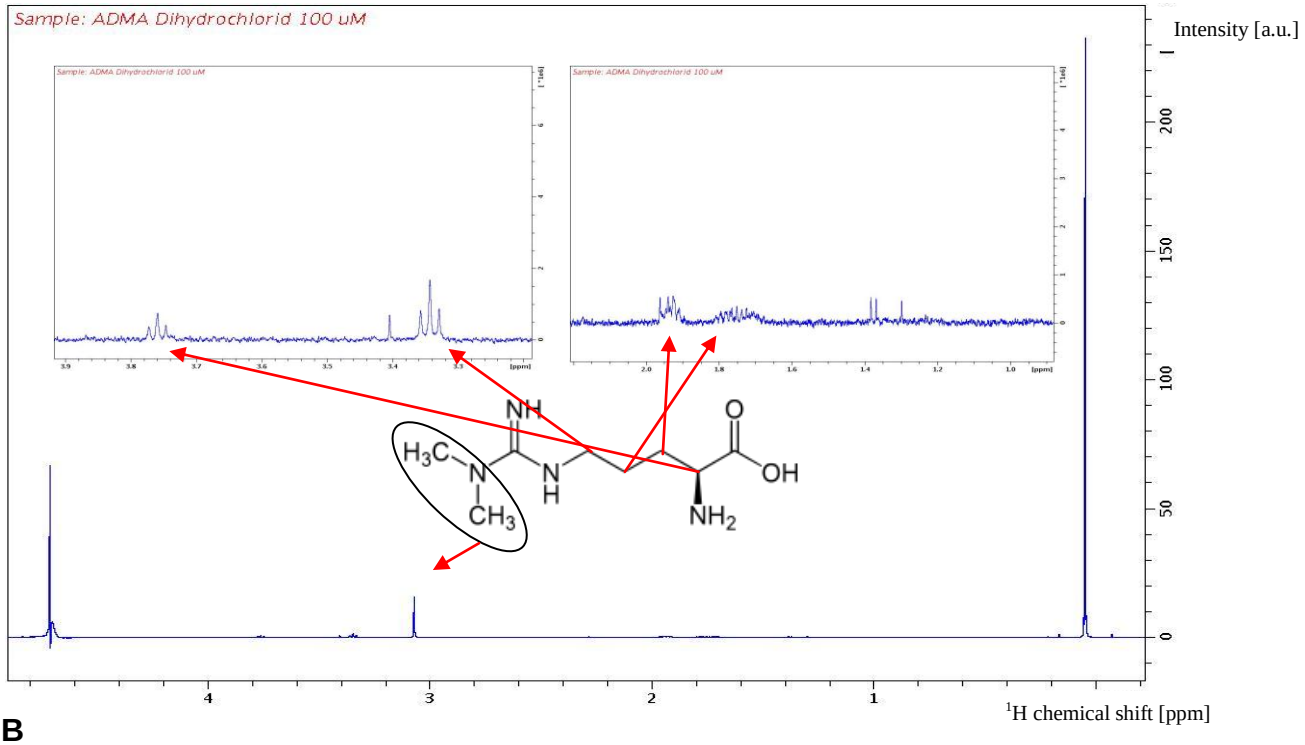
	CPMG	J-resolved
Pre scan delay [µsec]	6.5	6.5
Nr. of dummy scans	4	4
Nr. of scans	1024	128
Spectral width [Hz]	12 019	10 027
Spectral width [ppm]	24.03	20.05
P1 [µsec]	12.26	12.36
Acquisition time [sec]	3.07	0.8
Relaxation delay [sec]	2	1
Time domain points	73 728	16 384

2.3.5. Experimental setup of NMR-based metabolomics

Hundreds of different metabolites can be distinguished in these spectra. To analyze the level of PRMT1-mediated arginine methylation, ω-N^{G}, N^{G}-dimethyl-L-arginine (ADMA), ω-N^{G}, $N^{G'}$-dimethyl-L-arginine (SDMA), N^{G}-monomethyl-L-arginine (MMA), L-arginine, S-adenolylhomocysteine (SAH) and S-adenosyl-L-methionine (SAM) are metabolites of interest. Reference spectra of standard substances were recorded using CPMG pulse sequence. Examples of these standard spectra are shown in Fig. 12 and Fig. 13.
These spectra show the specific ppm values for the metabolites of interest and help to identify these metabolites in spectra of biofluids.

Different types of biofluids, namely cerebrospinal fluid (CSF), serum and urine, were measured. Since the signal intensity is directly linked to the concentrations of the different metabolites, identified metabolites could also be quantified.

A



B

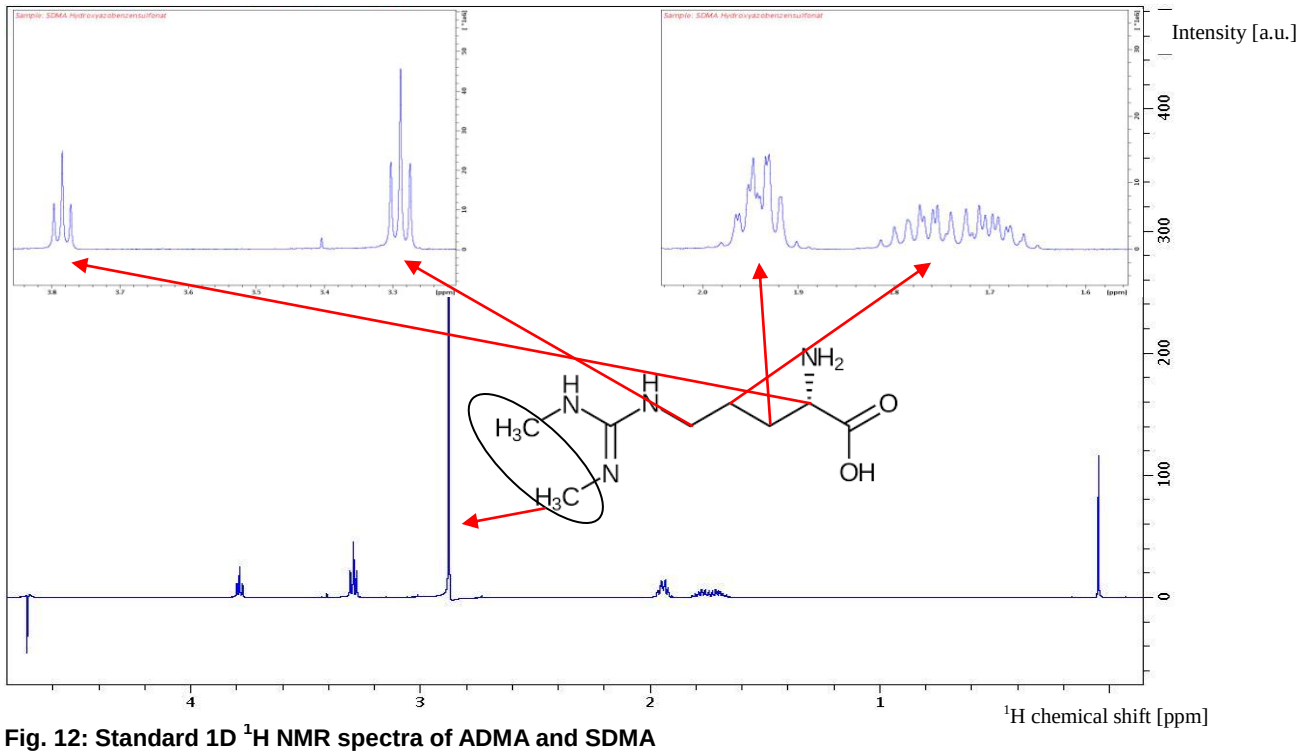
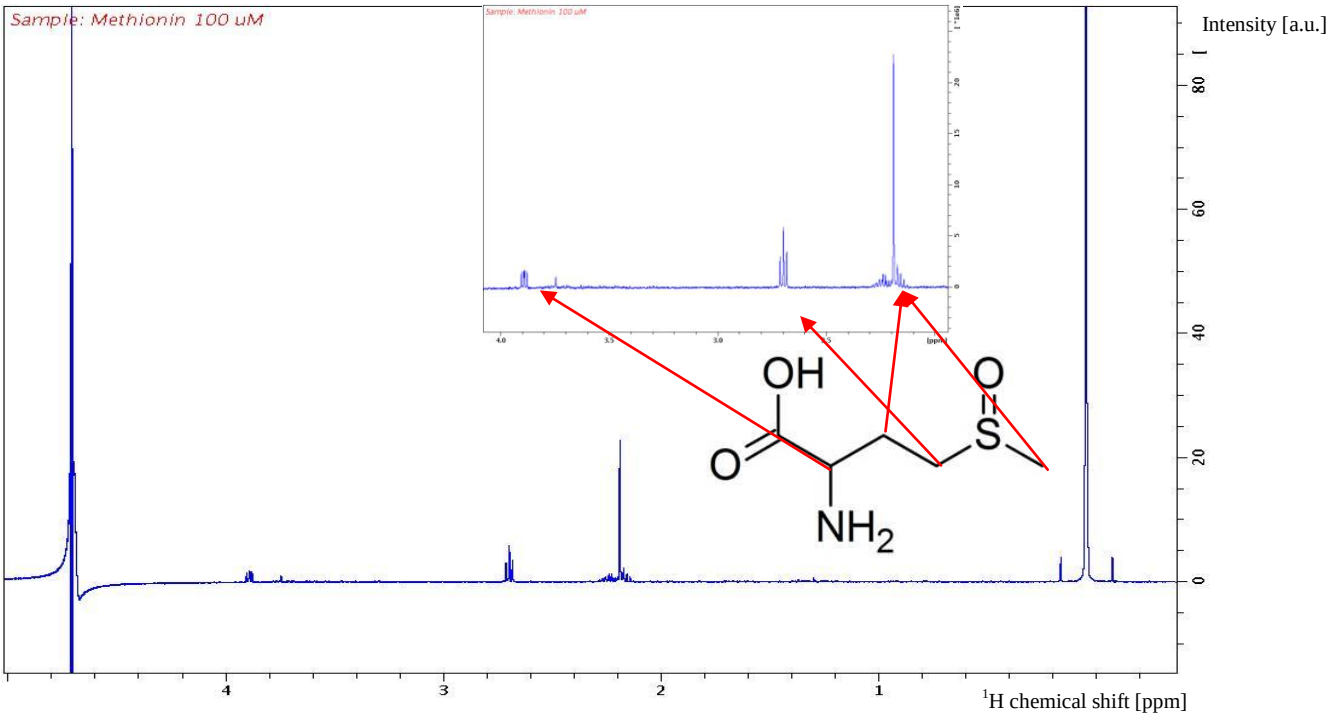


Fig. 12: Standard 1D ^{1}H NMR spectra of ADMA and SDMA

A: ω-N^{G}, N^{G}-dimethyl-L-arginine (ADMA) – Hydrochlorid [100 µM]
B: ω-N^{G}, $N^{G'}$-dimethyl-L-arginine (SDMA) – Hydroxyazobenzensulfonat [100 µM]
Signal intensity is plotted vs. ppm. Protons producing signals are indicated.

A



B

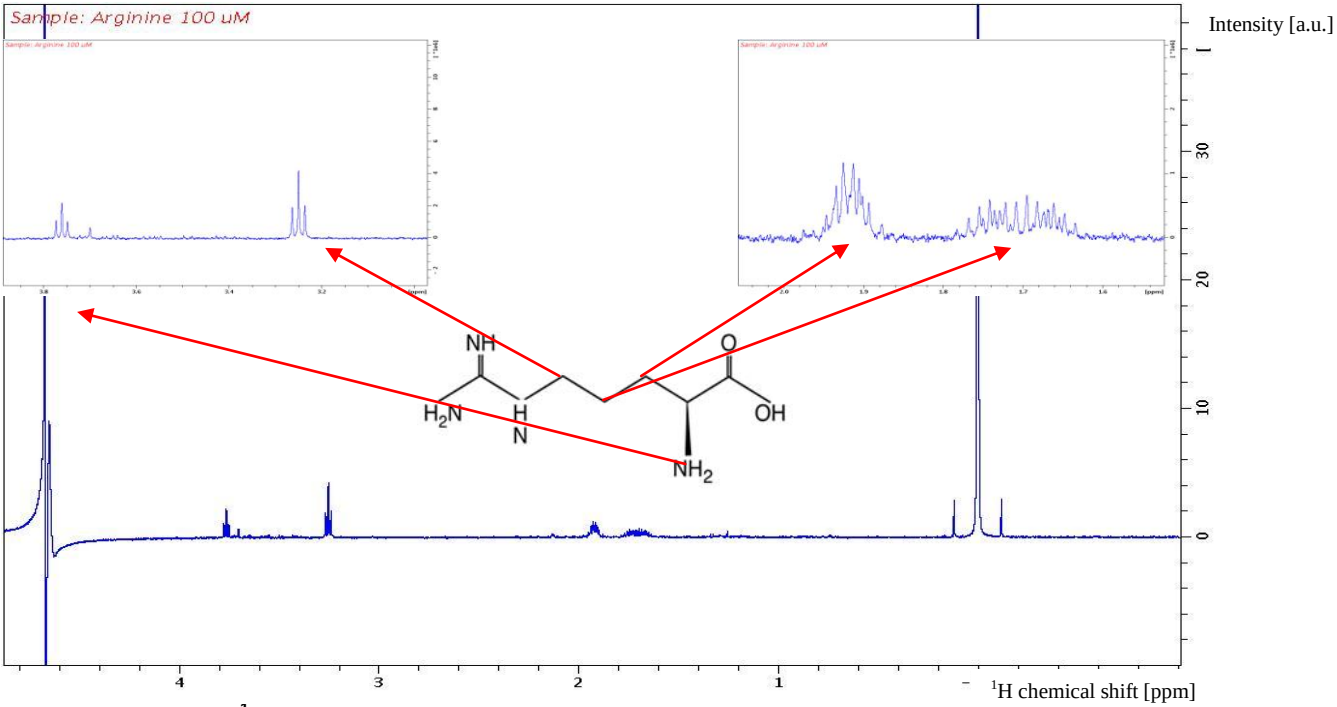
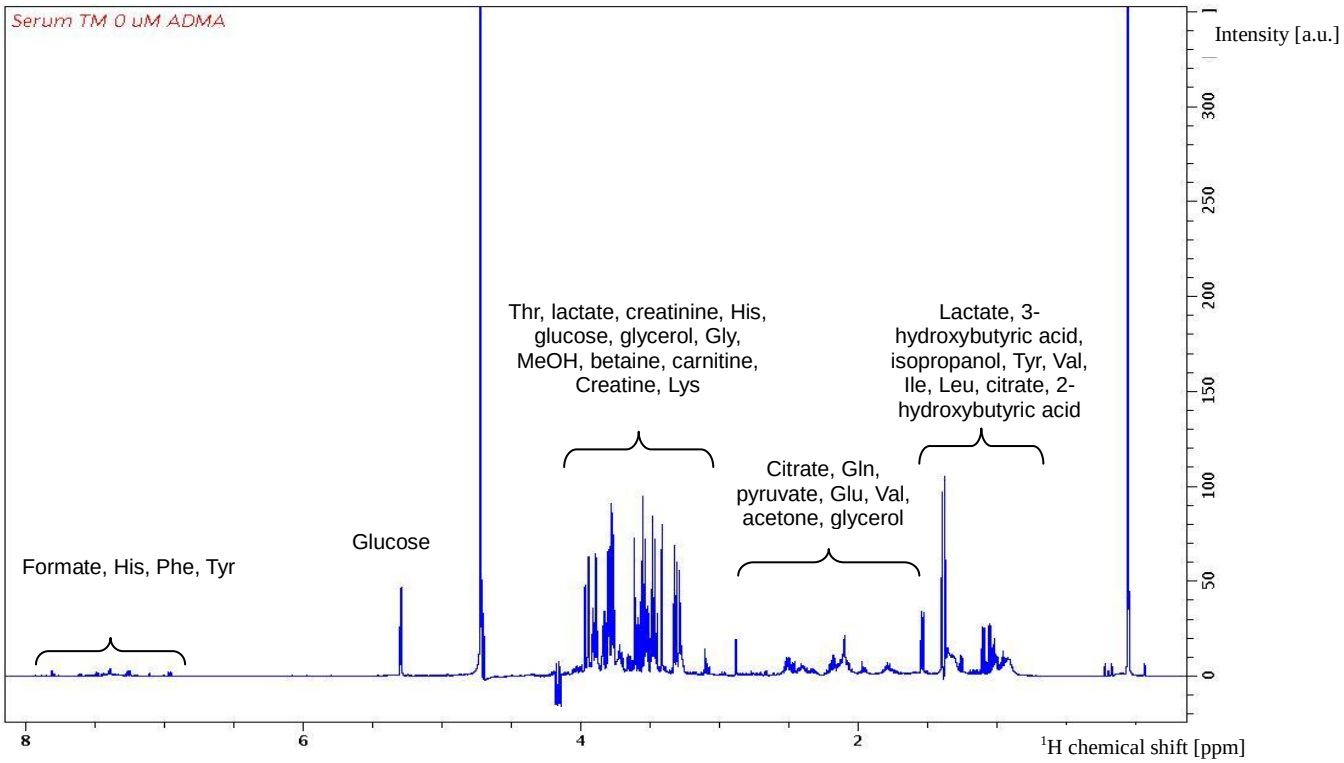


Fig. 13: Standard 1D ^{1}H NMR spectra of Methionine and Arginine

A: Methionine [100 µM]
B: Arginine [100 µM]
Signal intensity is plotted vs. ppm. Protons producing signals are indicated.

This NMR metabolomics protocol using CPMG pulse sequence is also applicable for different types of biofluids like serum and CSF. Each metabolite is represented by one or several peaks depending on the amount of ^{1}H atoms in chemically different environment (Fig. 14).

A



B

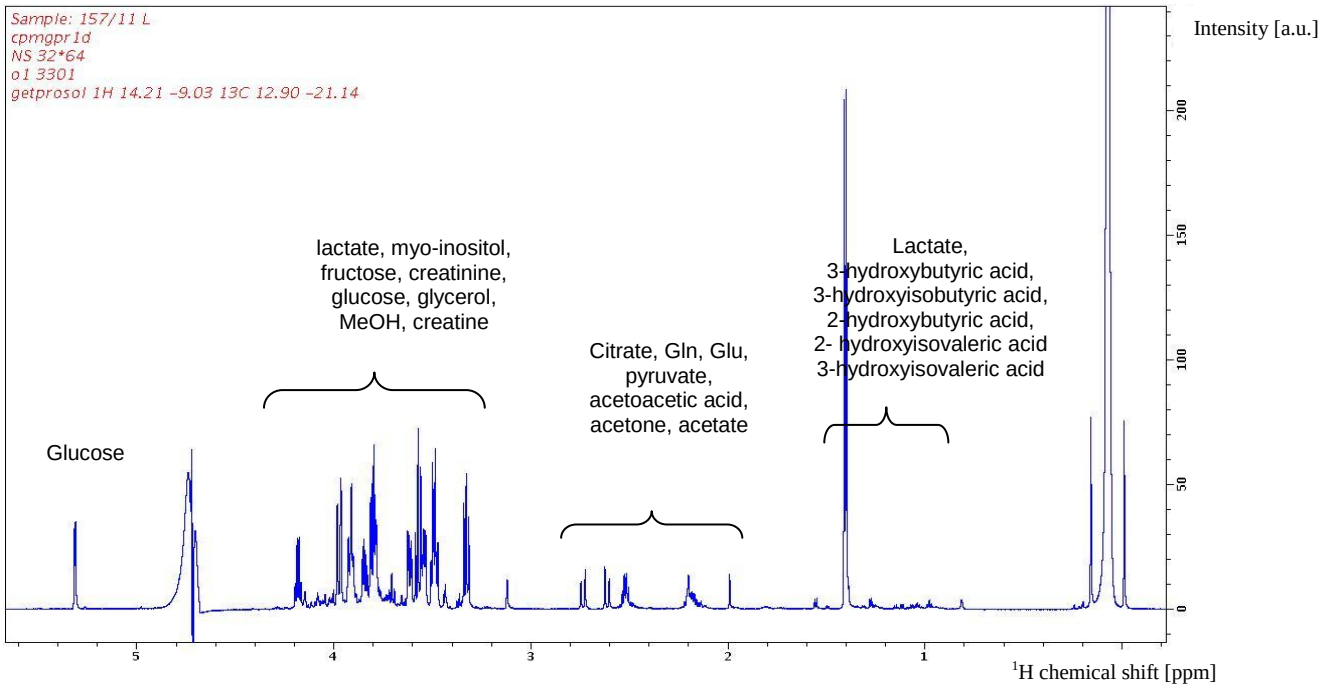
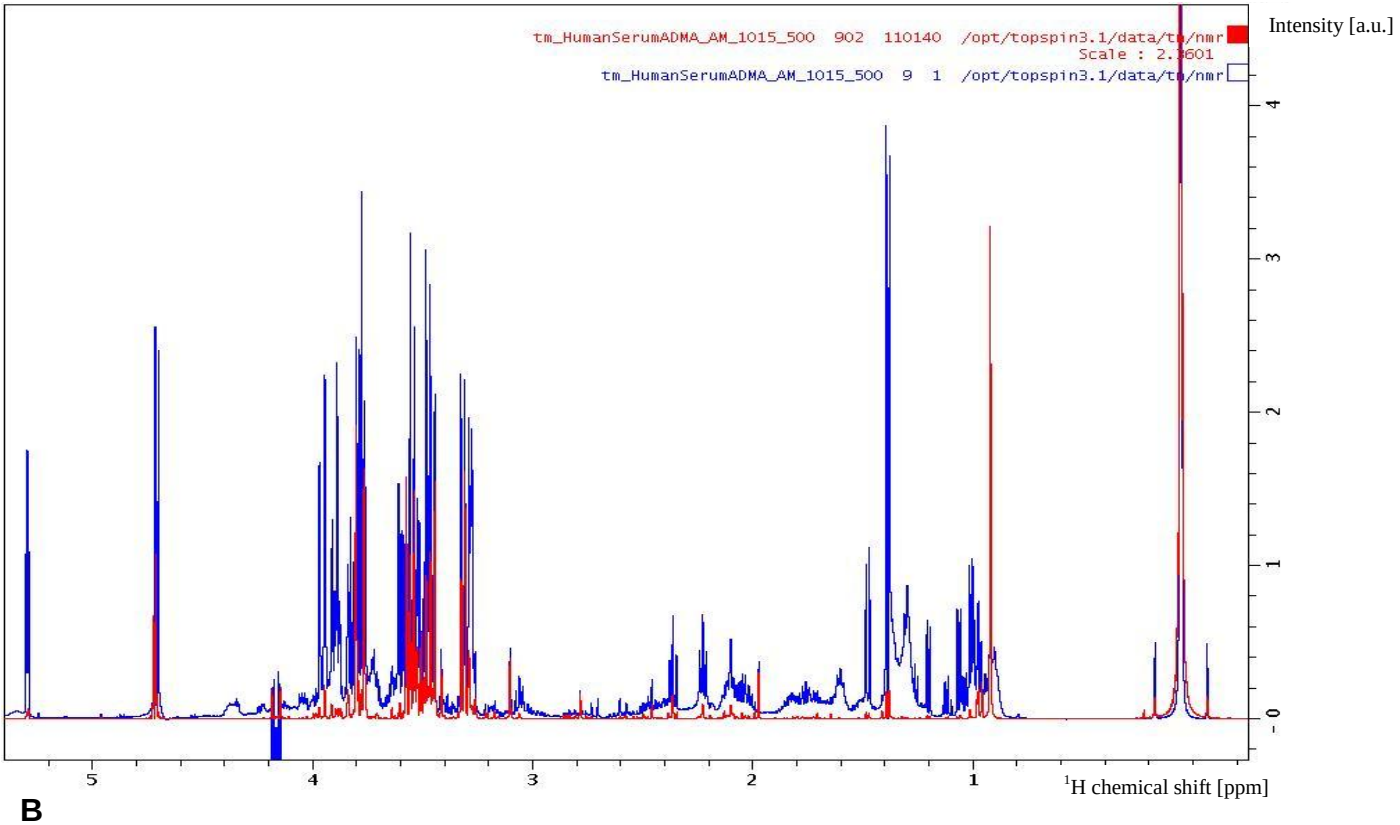


Fig. 14: 1D ^{1}H NMR spectra of serum and CSF

CPMG pulse sequence was used to record 1D ^{1}H NMR spectra of serum (A) and CSF (B). A variety of different metabolites could be identified in these spectra. Signal intensity is plotted vs. ppm.

In order to get a better resolution of ADMA signals, 2D J-resolved experiments were recorded (Fig. 15 A).

A



B

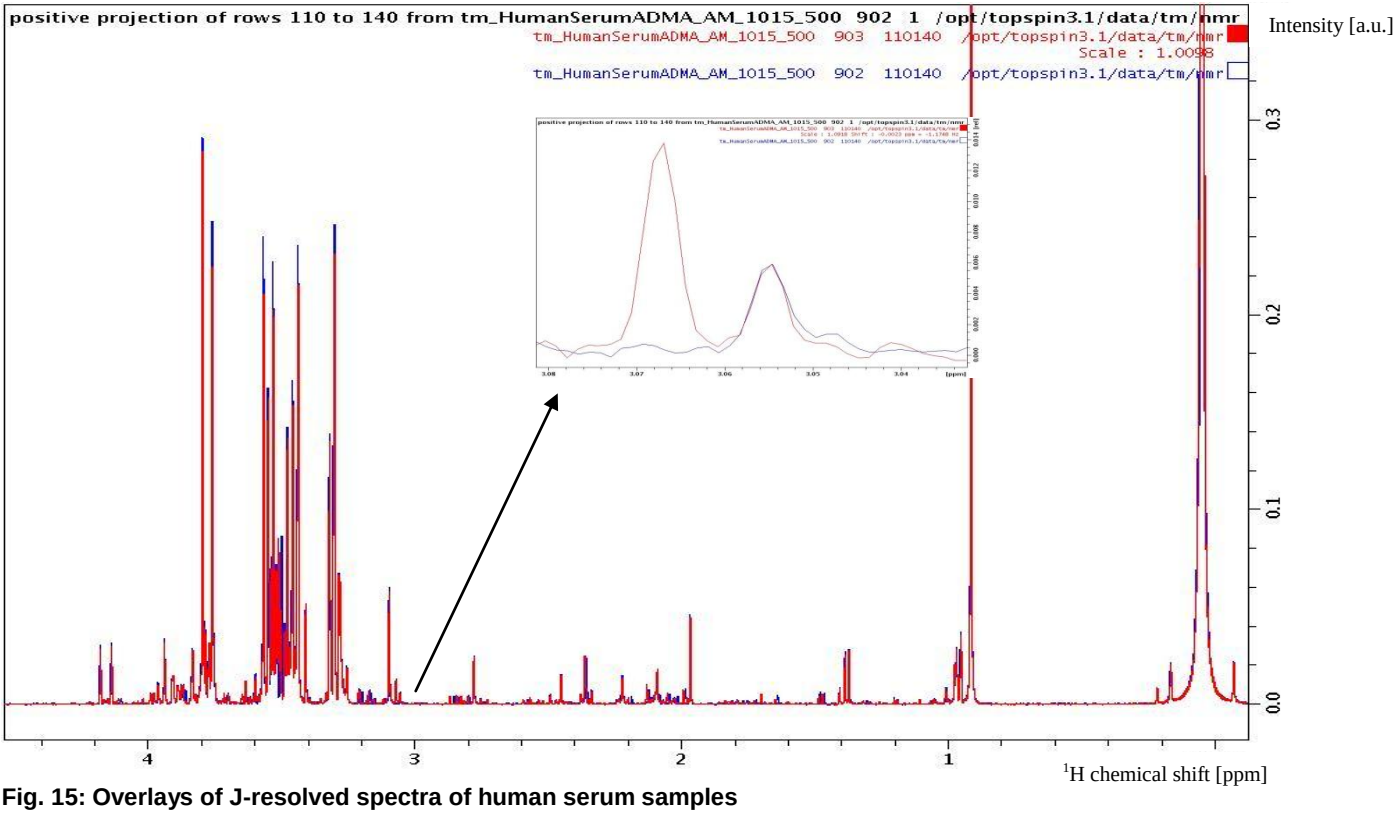


Fig. 15: Overlays of J-resolved spectra of human serum samples

JRES spectra were recorded in order to reduce the complexity of spectra. Overlays of CMPG and JRES spectra are shown in A. In addition, J-resolved spectra of ADMA spike-in experiments were recorded to ensure identification and retrieving of peaks of interest (B). Signal intensity is plotted vs. ppm.

These experiments reduce the complexity of NMR spectra in order to facilitate the identification of different metabolites.

In addition, J-resolved NMR spectra of spike-in experiments were recorded, to specifically identify additional peaks for the desired metabolites (Fig. 16). Positive projections of rows 110-140 showed information in 1D spectrum. (Fig. 15 B).

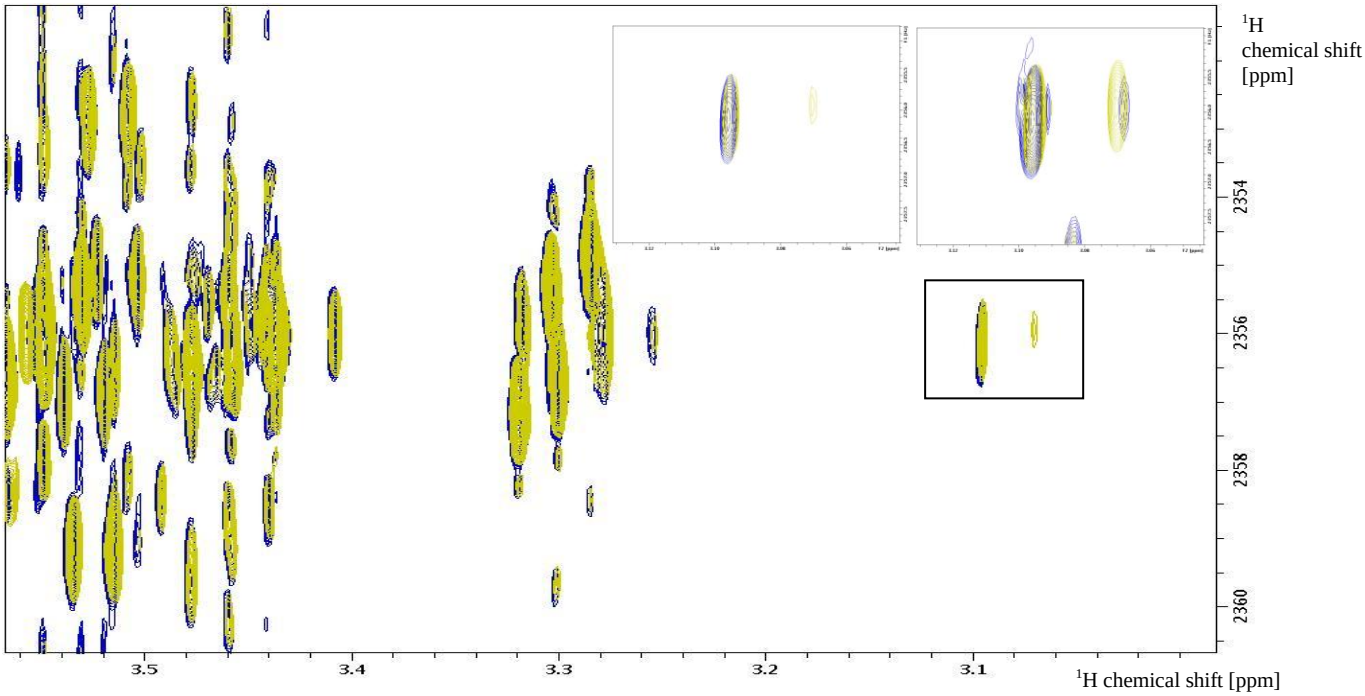


Fig. 16: Overlay of J-resolved NMR spectra

Overlay of J-resolved NMR spectra for ADMA spike-in experiments is shown.
Blue: serum without ADMA
Yellow: serum with 10 µM ADMA

2.3.6. ADMA extraction using SPE extraction cartridges

ADMA was extracted using SPE extraction cartridges (Fa. Waters: Oasis MCX 1cc). The columns were treated with a detachment solution (10% NH_3 saturated solution, 40 % MilliQ H_2O, 50 % MeOH), MeOH and PBS (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH adjusted with HCl and NaOH to 7.4). Subsequently the samples were applied and washed consecutively with MilliQ H_2O, 0.1 M HCl and MeOH. The concentrated ADMA was recovered using the detachment solution. Each step was performed with 1 ml of solution and via centrifugation at room temperature (2200 rpm for 1 minute). The solution was dried at 60°C with nitrogen and re-dissolved in deuterated NMR metabolomics buffer for NMR measurement.

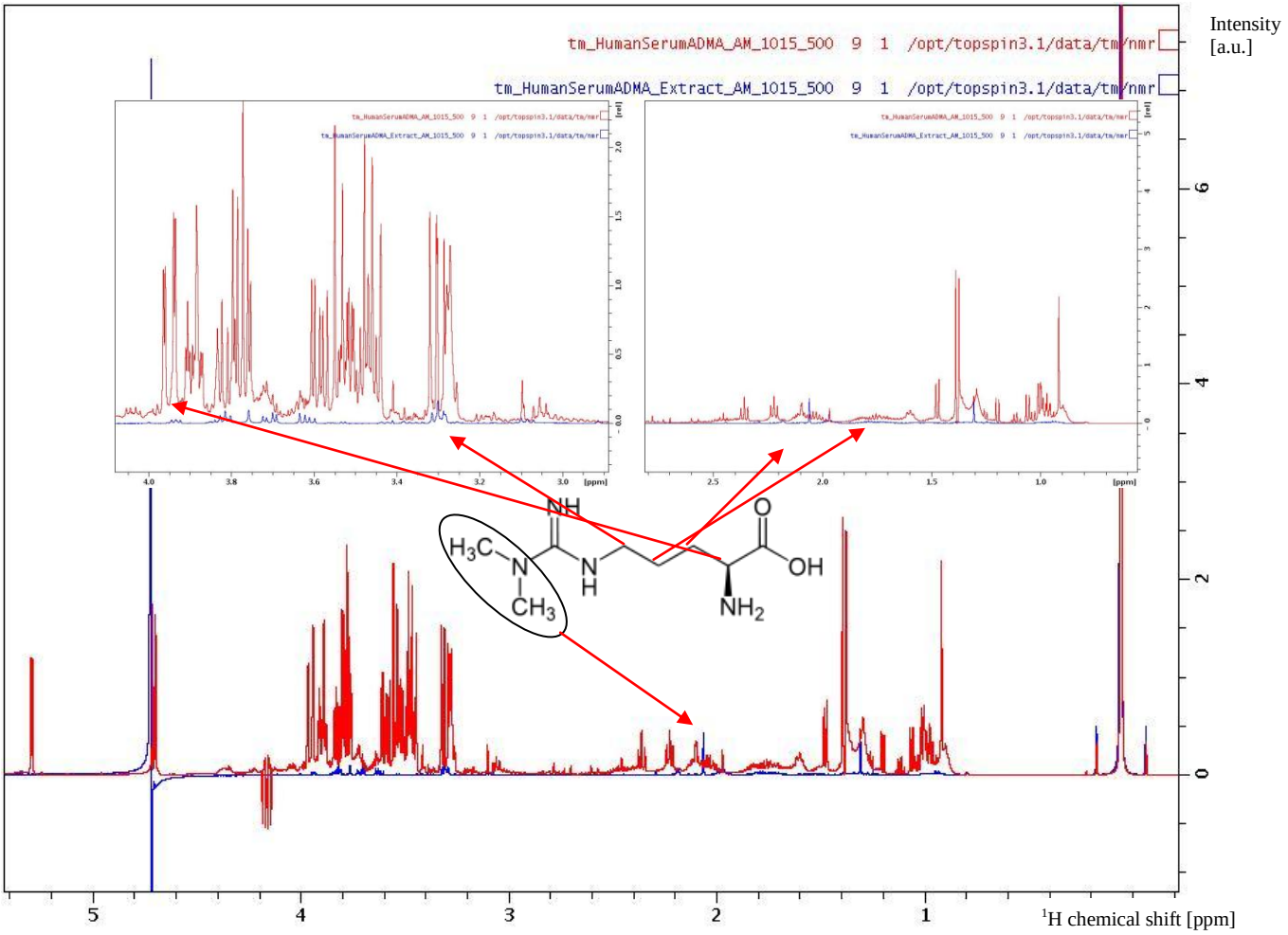


Fig. 17: Serum after solid phase extraction

Overlay of human serum (red) and human serum after solid phase extraction (blue). SPE resulted in a reduced complexity of the spectrum for a better identification and quantification of signals. Signal intensity is plotted vs. ppm.

Due to a similar chemical shift of ^{1}H atoms, peaks in some regions are overlapped, which leads to complex spectra. To improve the quantification of ADMA signals, SPE extraction cartridges were used. This protocol enables the concentration of ADMA and the removal of nonspecific signals in spectra. Spectra were recorded with the same setup using CPMG sequence. In fact, solid phase extraction for concentration of ADMA resulted in a reduced complexity of the spectrum.

2.3.7. Analysis of NMR metabolomics data

Raw data from Bruker Topspin were inserted into MestreNova for the analysis. It was used for phase correction, baseline correction and alignment of spectra.

Statistical analysis of metabolomics data was performed in a Matlab software using Principal component analysis (PCA) as well as Orthogonal-Partial Least Square-Discriminant Analysis (O-PLS-DA). O-PLS-DA plots were created with Matlab script according to the method of Cloarec et al. [70]

Bruker raw data were imported into Matlab, water (δ 4.45-5.15) and TSP (δ -0.2 – 0.2) signals were excluded and data points were aligned. Normalization reduced effects of dilution or different concentrations.

PCA was used to see similarities in spectra and to exclude outliers. Orthogonal projections on partial least square discriminant analysis were used to characterize differences in patients and controls. Statistical significance was cross-validated by leaving out a sample at a time.

3. Results

3.1. Purification of proteins

First purification step:

TEV protease and PRMT1 with polyhistidine tag as well as transportin-1 and CIRP constructs with Z-tag were expressed several times in *E.coli* BL 21 DE3 strain and purified using His Trap FF or Ni-NTA. The different steps of purification were checked using SDS polyacrylamide gelelectrophoresis (SDS PAGE; Fig. 18, Fig. 19, Fig. 20 A). Since the general procedure of purification, based on high affinity binding of polyhistidine residues to bivalent nickel ions, was the same for all proteins, only examples for protein purification are shown.

SDS PageRuler in Lane 1 is a pre-stained marker for determination of the molecular weight of proteins.

Lane 2 and 3 show the *E.coli* cell lysates before and after induction with isopropyl ß-D-1-thiogalactopyranoside (IPTG), respectively. After induction of overexpression of the different proteins of interest, an additional intense band of the specific molecular weight appeared in Lane 3. Consequently the heterologous expression was efficient and a high amount of protein was produced by *E.coli* BL 21 DE3. Although the overexpression bands were not very intense (Fig. 19, Fig. 20 A), a high amount of protein was purified. This may be due to saturation of the lanes with cell lysate and proteins of similar molecular weight.

Lane 4 corresponds to the supernatant that has passed the HisTrap/Ni-NTA and should contain only non-specific *E.coli* proteins lacking the polyhistidine tag (Flow through). Since this lane contained a high amount of proteins the lysis of the cells via sonication was efficient and proteins were soluble in the lysis buffer. Although the flow through was passed twice through the HisTrap to ensure a high binding rate there was still some of the protein of interest present which did not bind to the column. This could be due to the saturation of the Ni^{2+} ions or a too high flow rate when passing the lysate via the column.

The Wash fractions on Lane 5 (and 6 for Fig. 20 A) with buffer containing higher imidazole concentration removed non-specific bound protein of *E.coli* of the column. There was still some protein of interest which may be the remaining flow through in the columns at the beginning of the washing steps. In addition, a slight degradation in the region of the his-tag and therefore a reduced affinity to bivalent nickel ions, which does not affect the molecular

weight in a vigorous manner, may lead to an early elution of the protein.

The elution contained mainly the proteins of interest and some contaminations which may be *E.coli* proteins with His-rich regions having higher affinity to Ni^{2+} ions or, if smaller molecular weight, degradation products of the protein of interest.

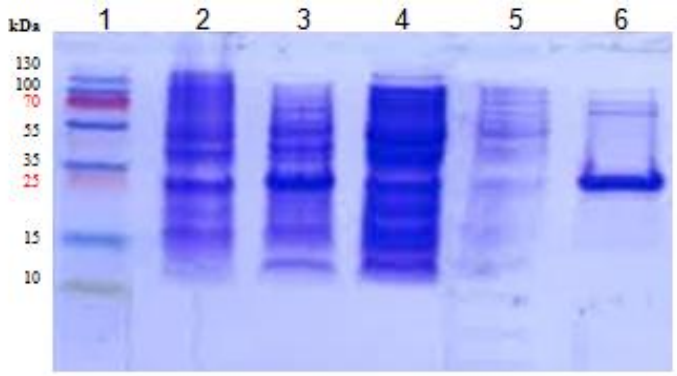


Fig. 18: Expression and purification of TEV protease Enzyme *Tobacco Etch Virus* protease for the TEV cleavage was expressed in *E.coli* BL21 cells and purified using His-Trap FF (GE Healthcare). Purification procedure was checked using SDS PAGE.

Lane 1: Prestained Proteinmarker, Lane 2: *E.coli* cell lysate before induction, Lane 3: *E.coli* cell lysate after induction with IPTG, Lane 4: Flow through, Lane 5: Wash fraction with higher imidazole concentration, Lane 6: elution of HisTrap

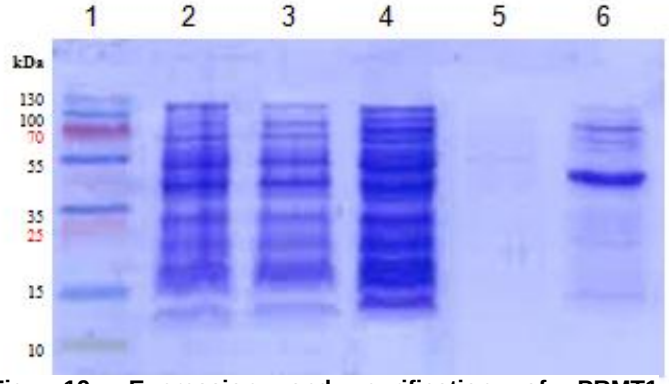


Fig. 19: Expression and purification of PRMT1 Enzyme PRMT1 for the methylation reaction was expressed in *E.coli* BL21 cells and purified using His-Trap FF (GE Healthcare). Purification procedure was checked using SDS PAGE.

Lane 1: Prestained Proteinmarker, Lane 2: *E.coli* cell lysate before induction, Lane 3: *E.coli* cell lysate after induction with IPTG, Lane 4: Flow through, Lane 5: Wash fraction with higher imidazole concenration, Lane 6: elution of HisTrap

Second purification step

To ensure a native conformation of the purified proteins, it was important to remove the Z-tag. The cleavage was performed overnight at 4 °C and checked via SDS PAGE (Fig. 20 B).

The molecular weight of the band in the elution (Lane 2) corresponds to the molecular weight of protein of interest + Z tag. After the cleavage reaction the molecular weight of the protein was reduced by the molecular weight of the Z-tag (10 kDa) (Lane 3). The protein solution was applied to the HisTrap to bind the TEV protease as well as the Tag and to obtain the purified protein without a tag as Flow through (Lane 4).

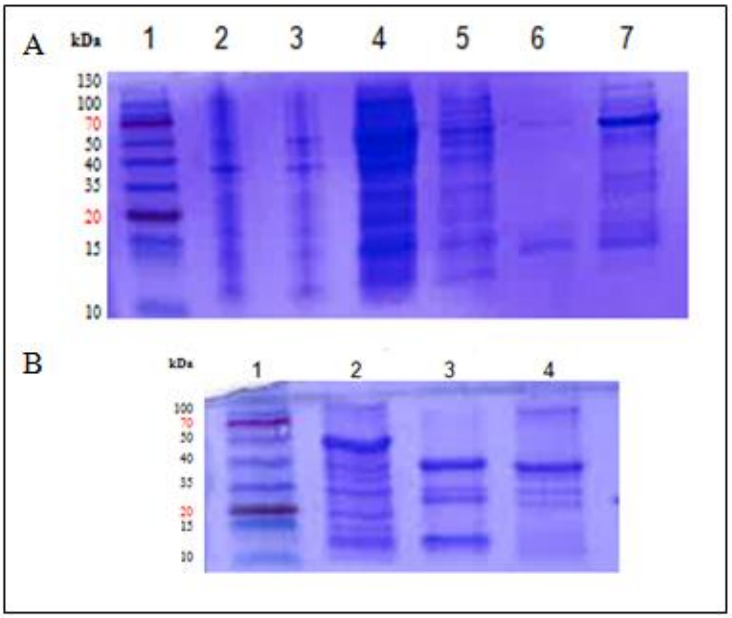


Fig. 20: Expression, Purification and TEV cleavage of purified proteins
Proteins was expressed in *E.coli* BL21 cells, purified using His-Trap FF (GE Healthcare) and the tag was removed using TEV protease. Purification procedure was checked using SDS PAGE.

A: Lane 1: Prestained Proteinmarker, Lane 2: *E.coli* cell lysate before induction, Lane 3: *E.coli* cell lysate after induction with IPTG, Lane 4: Flow through, Lane 5, 6: Wash fractions with higher imidazole concentrations, Lane 7: elution of HisTrap
B: Lane 1: Prestained Proteinmarker, Lane 2: protein with Z Tag, Lane 3: Protein + TEV protease after cleavage, Lane 4: cleaved and repurified protein

3.2. CIRP as a model system for RGG proteome

3.2.1. Distinct residues of transportin-1 are involved in CIRP-SY, CIRP-RGG and FUS-PY binding

NMR titration measurements of ILVM-labeled transportin-1 indicate a higher affinity in transportin-1 binding of FUS-PY than of CIRP-SY. A mutual titration of CIRP-SY and FUS-PY revealed that distinct residues are involved in the binding and a simultaneous binding is possible. Yet all the transportin-1 residues involved in the FUS-PY binding shifted back in presence of FUS-PY and CIRP-SY, indicating that the binding of FUS-PY has a higher affinity. Some transportin-1 residues are only affected by CIRP-SY and in presence of FUS-PY, they were still affected. (Fig. 21)

However NMR data show a simultaneous binding of CIRP-SY and FUS-PY to distinct residues of transportin-1, Isothermal titration calorimetry indicates a vastly higher affinity of FUS-PY than CIRP-SY to transportin-1 (48.1 ± 5.0 nM and 4.0 ± 0.8 µM, respectively). In presence of FUS-PY, there is no binding of CIRP-SY detectable. Vice versa, FUS-PY can displace the bound CIRP-SY and the affinity to transportin-1 is not significantly altered (33.1 ± 3.7 nM).

Transportin-1 is a structured protein containing a flexible loop region. For ITC experiments, CIRP-SY and FUS-PY were titrated to the full length transportin-1 (TNPO1) and to a construct lacking the flexible region (TNPO1Δlinker). For both constructs, the affinity did not change significantly (for FUS-PY 48.1 ± 5.0 nM and 61.7 ± 4.2 nM, respectively; for CIRP-SY 8.3 ± 0.5 µM and 4.0 ± 0.8 µM, respectively).

For CIRP-RGG, the affinity to the TNPO1Δlinker is reduced compared to the TNPO1 (363.6 ± 20.6 nM and 74.6 ± 11.7 nM, respectively). This indicates a binding of the RGG motif mainly to the flexible loop. Simultaneous binding of CIRP-SY and CIRP-RGG were tested and CIRP-RGG could bind to TNPO1 in presence of CIRP-SY with the same affinity (69.4 ± 2.3 nM), whereas no binding of CIRP-SY was detectable in presence of CIRP-RGG. (Tab. 5, Suppl. Tab. 1, Suppl. Fig. 1).

Tab. 5: ITC data of CIRP

	K_d [µM]
TNPO1 vs. CIRP-SY	4.0 ± 0.8
TNPO1 vs. FUS-PY	0.0481 ± 0.0050
TNPO1 vs. CIRP-RGG	0.0746 ± 0.0117
TNPO1Δlinker vs. CIRP-SY	8.3 ± 0.5
TNPO1Δlinker vs. FUS-PY	0.0617 ± 0.0042
TNPO1Δlinker vs. CIRP-RGG	0.3636 ± 0.0206
TNPO1/CIRP-SY vs. FUS-PY	0.0331 ± 0.0037
TNPO1/FUS-PY vs. CIRP-SY	No binding
TNPO1/CIRP-SY vs. CIRP-RGG	0.0694 ± 0.0023
TNPO1/CIRP-RGG vs. CIRP-SY	No binding

NMR titrations of CIRP-RGG, CIRP-SY and CIRP full length construct to ILVM-labeled transportin-1 reveal a binding of all the three constructs to the karyopherinβ2. However, the residues affected in the binding of CIRP-SY to transportin-1, shifted back or in another direction upon binding of CIRP full length, where CIRP-SY as well as CIRP-RGG motifs are present. This implicates that the CIRP-SY binding to TNPO1 is inhibited in the full length construct. This hypothesis is also supported by the fact that the affinity of CIRP-SY is weaker than the affinity of CIRP-RGG to transportin-1. (Tab. 5, Fig. 22)

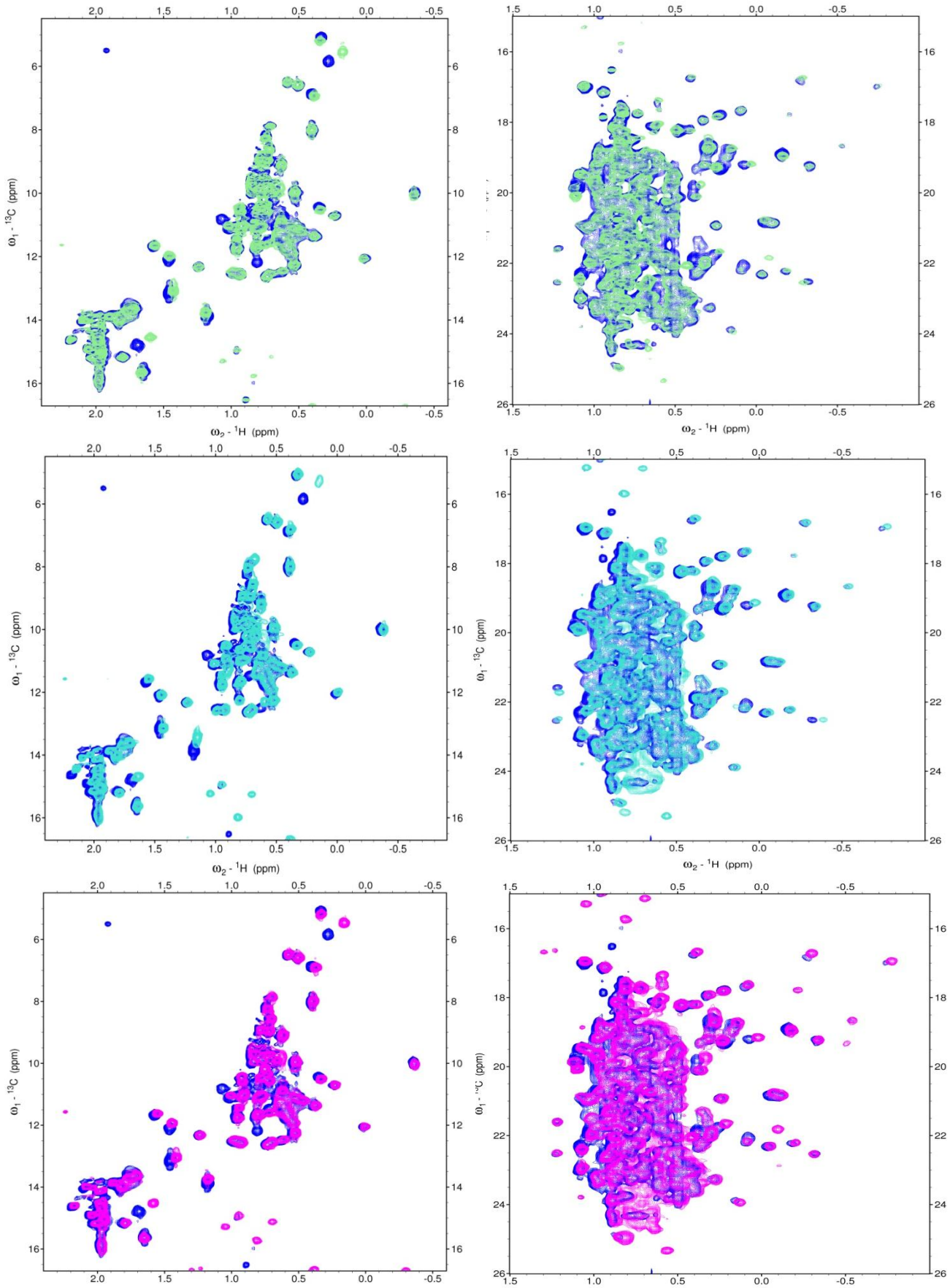


Fig. 21 Overlays of ^{13}C-^{1}H-NMR titrations of ILVM labeled transportin-1 with CIRP-SY, FUS-PY and CIRP-SY/FUS-PY, respectively

The spectrum for the ILVM labeled transportin-1 is shown in dark blue. Equimolar amounts of CIRP-SY and FUS-PY were added.

Light blue: transportin-1 + CIRP-SY, Turquoise: transportin-1 + FUS-PY, Magenta: transportin-1 + CIRP-SY/FUS-PY

Left: Ile, Leu, Meth; Right: Valin

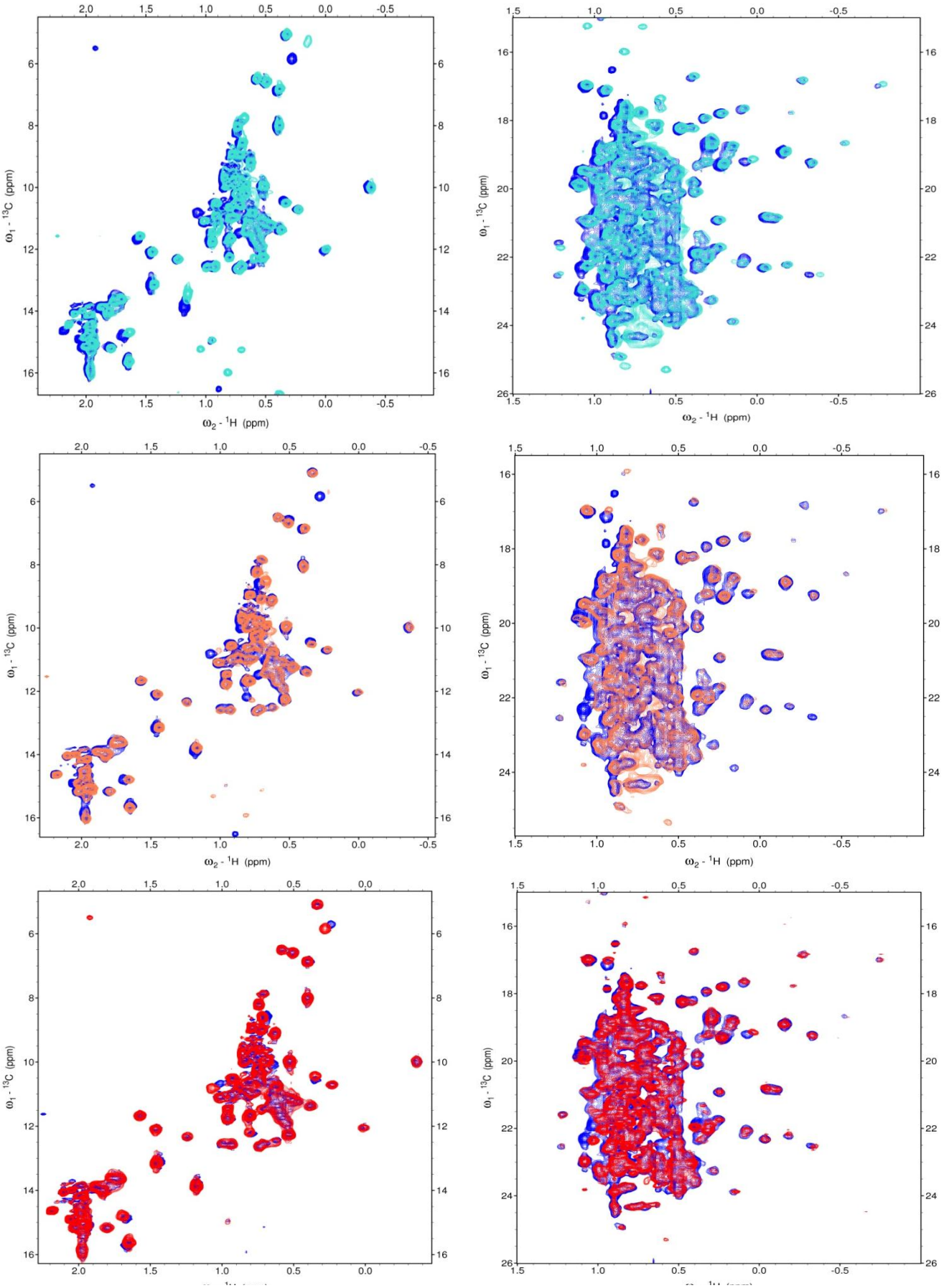


Fig. 22 Overlays of 13C-1H-NMR titrations of ILVM labeled transportin-1 with CIRP-SY, CIRP full length and CIRP-RGG, respectively

The spectrum for the ILVM labeled transportin-1 is shown in dark blue. Equimolar amounts of CIRP-SY, CIRP full length and CIRP-RGG were added.

light blue: transportin-1 + CIRP-SY; coral: transportin-1 + CIRP full length; red: transportin-1 + CIRP-RGG

Left: Ile, Leu, Meth; Right: Valin

3.2.2. CIRP-SY contains phosphorylation sites

$^{15}N^{13}C$-labeled CIRP-SY was phosphorylated using In-Lysate phosphorylation assay. Additional peaks were detected after the phosphorylation reaction, which implicates a post-translational phosphorylation of residues. However, in comparison with the non-phosphorylated CIRP-SY no peaks disappeared. Thus, the phosphorylation reaction was not fully quantitative (Fig. 23, Fig. 24).

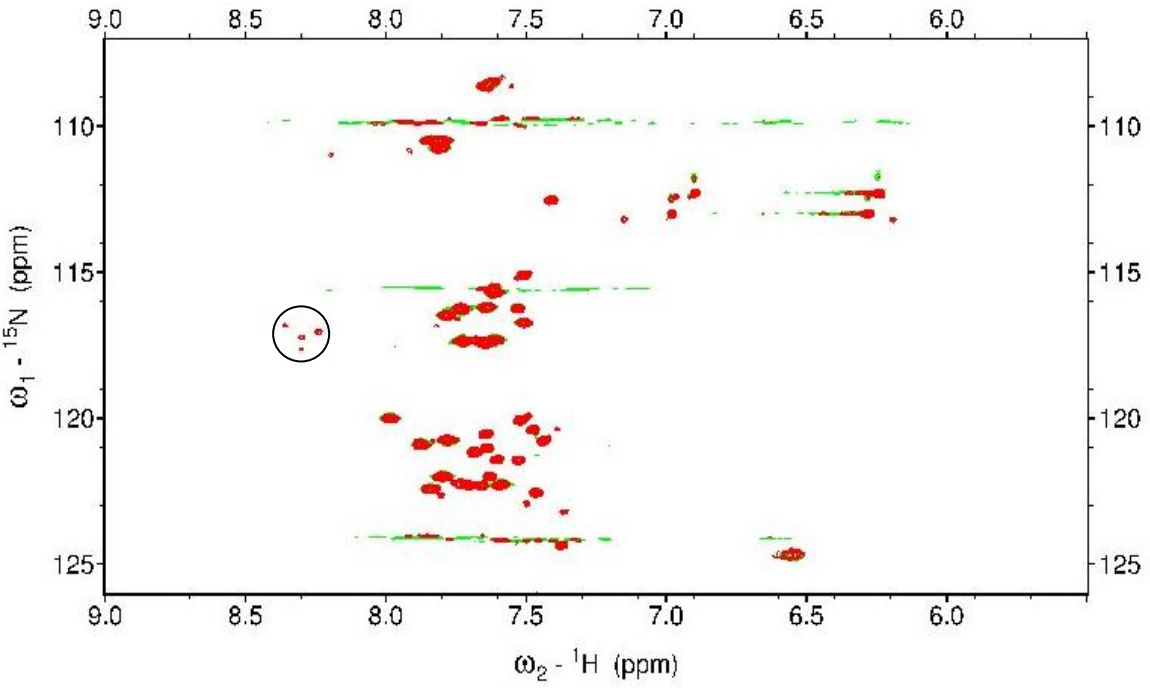


Fig. 23 Overlay of ^{1}H-^{15}N HSQC of phosphorylated and unphosphorylated CIRP-SY

^{13}C ^{15}N labeled CIRP-SY was phosphorylated using In Lysate phosphorylation assay. The unphosphorylated protein is shown in green, the phosphorylated form in red. Additional peaks in the red spectrum, indicating phosphorylation events, are encircled.

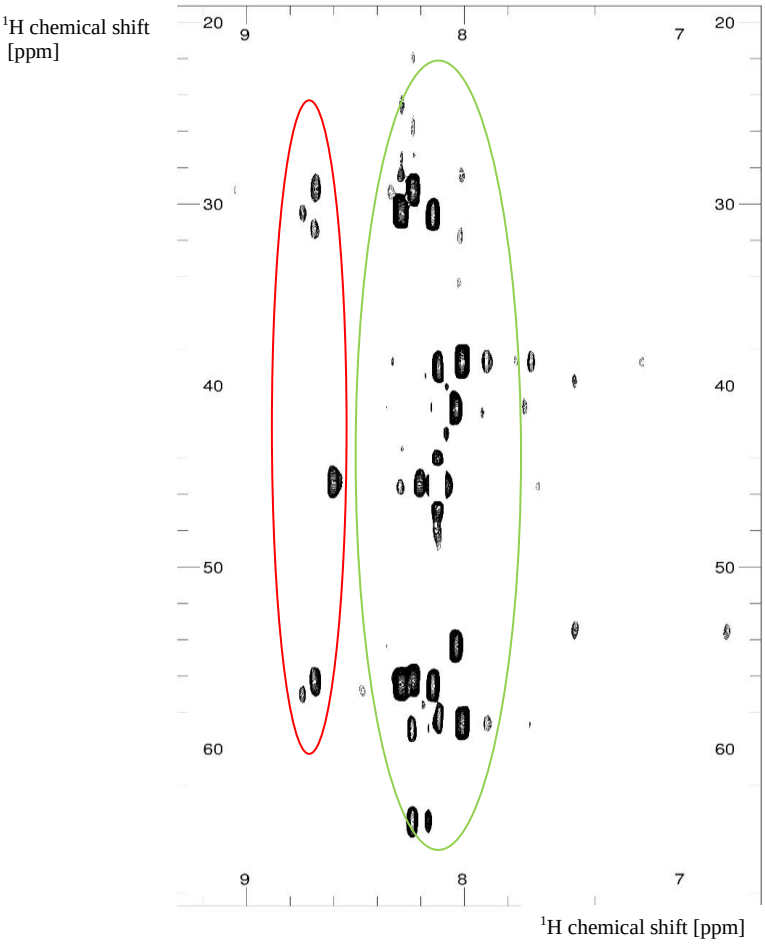


Fig. 24: HN(CO)CACB of $^{13}C^{15}N$ CIRP-SY after In-Lysate Phosphorylation
^{13}C ^{15}N labeled CIRP-SY was phosphorylated using In Lysate phosphorylation assay. Phosphorylated and non-phosphorylated serine residues are encircled in red and green, respectively.

In a HN(CO)CACB experiment, C_α and C_β of the amino acid in position i-1 can be identified. Chemical shifts of corresponding peaks were compared to random coil chemical shifts in literature. [71]

For one serine residue it was clear that the amino acid in position i-1 is a glycine. Since only Serine-159 of CIRP-SY construct in the region of transportion-1 interaction is located following to a glycine residue, it was used for mutagenesis. The remaining chemical shifts of amino acids in position i-1 were suitable for tryptophan, methionine, histidine, leucine, glutamate, glutamine, cysteine or arginine. Screening of the CIRP-SY sequence indicated, that there were only two arginine-serine (145-146, 154-155) and one glutamate-serine (147-148) sequence regions. Accordingly, these sequence regions were used for mutagenesis.

Based on the information of Fig. 24, serine-to-glutamate mutants were designed in order to mimic phosphorylation.

3.2.3. Protein crystals of CIRP-SY/ TNPO1Δlinker complex

Crystals of CIRP-SY/ TNPO1Δlinker complex were obtained after some weeks. The resolution of TNPO1Δlinker /CIRP-SY complex crystals was around 8 Angström. This would be too low to get an atomic resolution structure and identify the binding site of CIRP-SY on the transportin-1 protein.

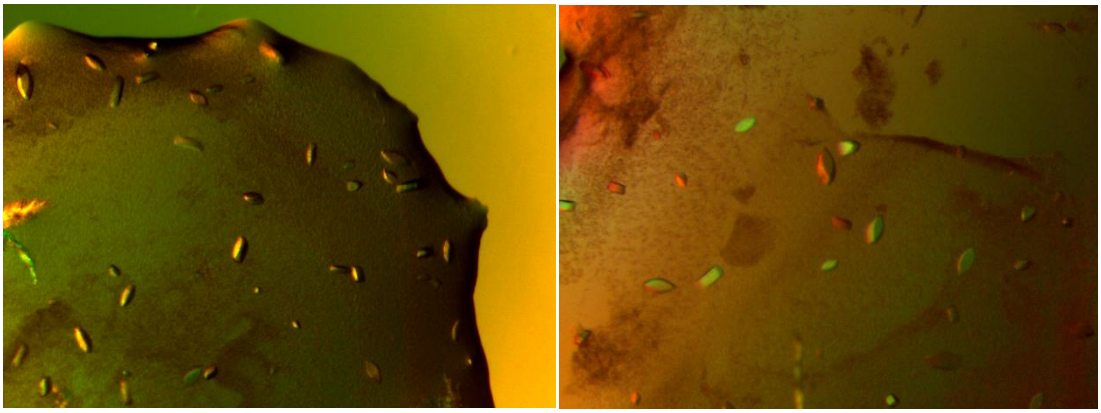


Fig. 25: Protein crystals of TNPO1Δlinker /CIRP-SY complex
8 x magnification

3.3. NMR based metabolomics

3.3.1. Patient cohort of Multiple Sclerosis patients

Biofluid samples (CSF and serum) of patients suffering from Multiple Sclerosis (MS) were measured on NMR 500 MHz spectrometer using the setup for NMR based metabolomics. To analyze the metabolic effects of MS/CIS, PCA and O-PLS-DA was carried out using Matlab software.

In Principal Component Analysis (PCA), only weak clustering was observed (Fig. 26). Outliers were identified using a multivariate generalization of Student's t-test (Hotellings T^2).

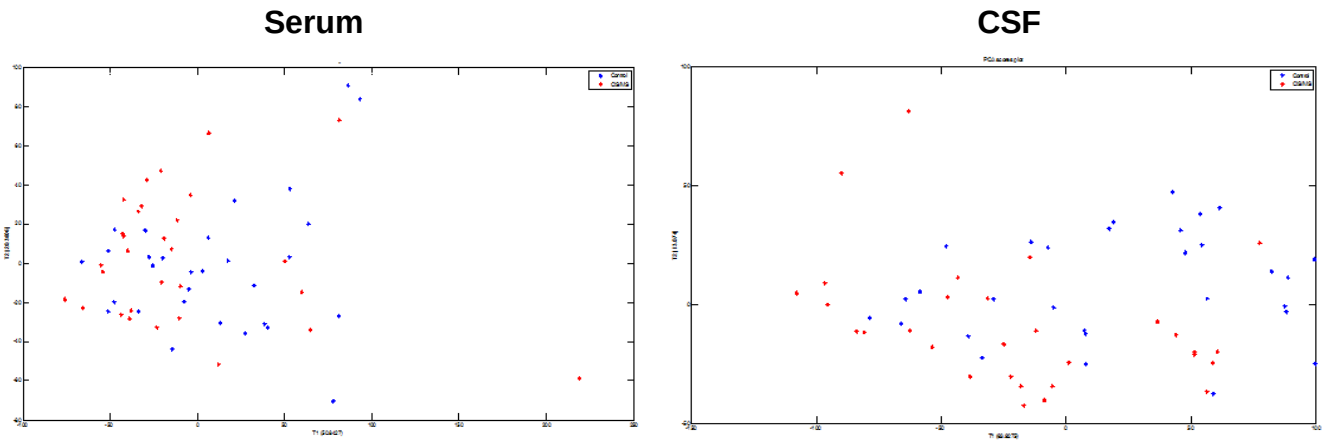


Fig. 26: Principal Component Analysis of serum and CSF

Healthy controls (blue) and patients (red) were compared using PCA. Only weak clustering was observed. Outliers were excluded using Hotellings T^2.

An O-PLS-DA indicated a well-defined clustering of control and patient samples (Fig. 27). Due to an erythrocytes contamination, one CSF sample was excluded in the O-PLS-DA.

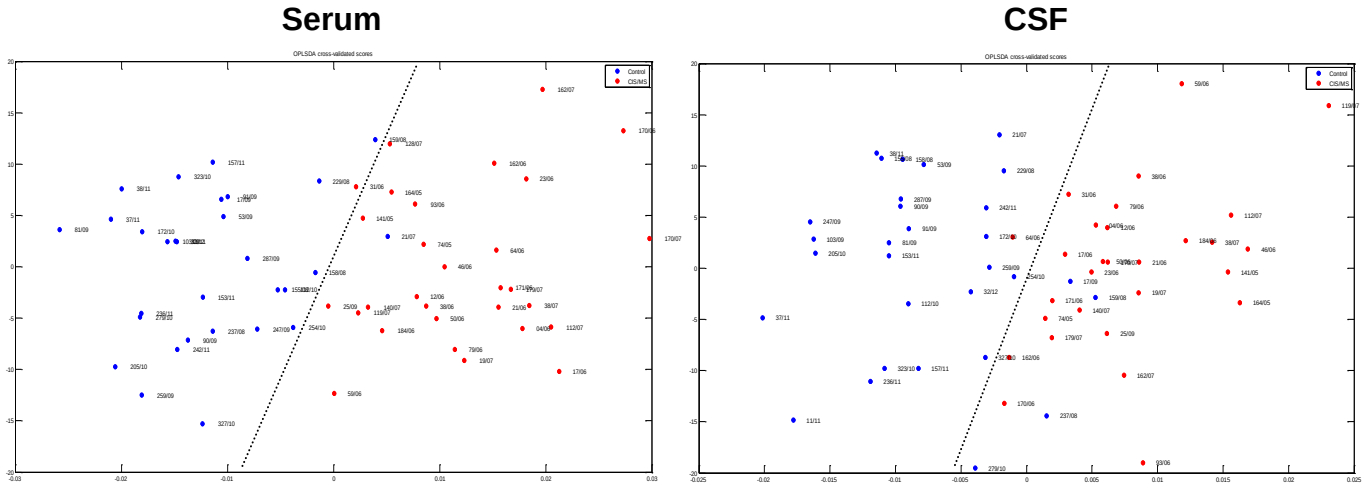


Fig. 27: O-PLS-DA of serum and CSF

O-PLS-DA indicates a clustering of patients (red) and control (blue) samples in serum as well as CSF. Due to an erythrocytes contamination, one CSF sample was excluded in the O-PLS-DA. Component 2 (orthogonal) in dependence of component 1 (correlated).

For O-PLS-DA, a high value of Q_Y^2 identifies statistically significant phenotypes. Back-scaling transformation and projection were used for biomarker visualization. Positive or negative signals indicate positive or negative covariance with healthy controls,

respectively. A color code of spectrum enables identification of biomarker-rich regions based on the correlation matrix.

Formate, acetat, hydroxyl branched-chain amino acids as well as glutamine were identified as biomarkers for distinction of 30 healthy controls and 30 patients suffering from multiple sclerosis or clinically isolated syndrome in serum as well as CSF. In addition, an aberrant sarcosine pattern was observed in serum, whereas in CSF the propylene glycol level was slightly affected (Fig. 28).

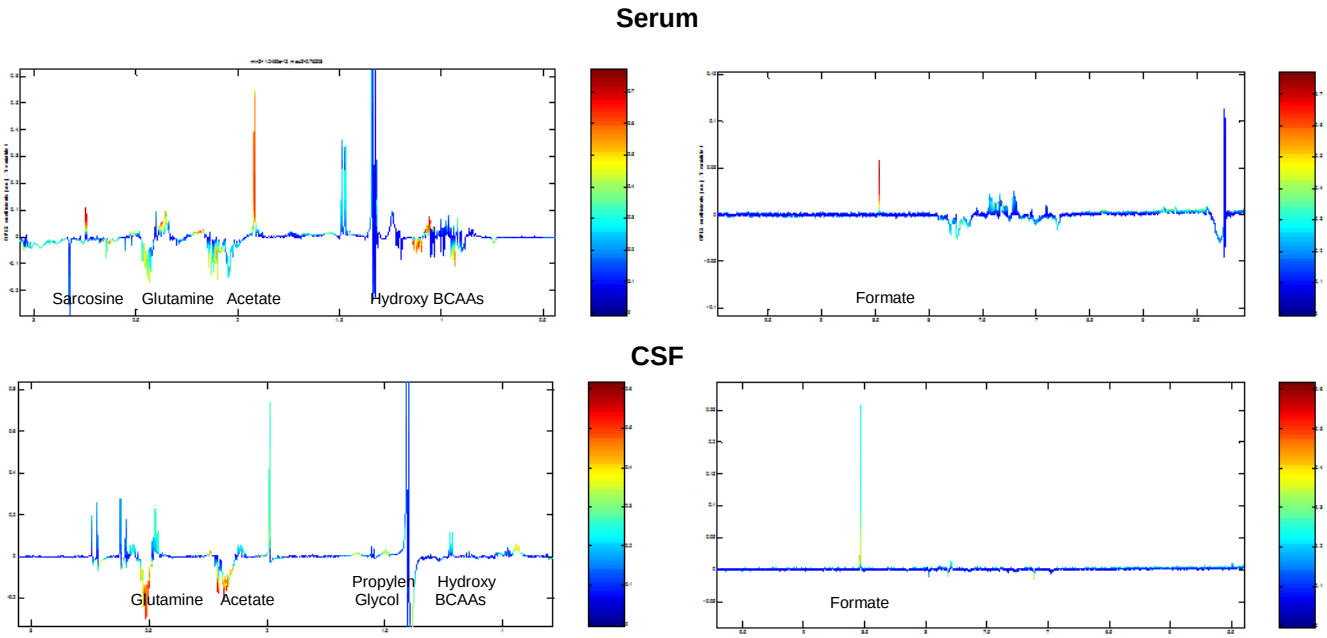


Fig. 28: O-PLS-DA coefficient plots for biomarker identification in Serum and CSF

1D ^{1}H NMR CPMG spectra of serum and CSF were used for O-PLS-DA. Positive or negative signals indicate positive or negative covariance with healthy controls, respectively. Color code enables identification of biomarker rich regions. Covariance vs. ^{1}H chemical shift in ppm.
BCAA, branched-chain amino acid

4. Discussion

4.1. CIRP as a model system for transportin-1 dependent nuclear import

This study indicated a binding of Cold inducible RNA binding protein to transportin-1. On the one hand, binding of the CIRP-RGG motif to transportin-1 was confirmed and on the other hand a second binding site was identified with CIRP-SY (Fig. 21, Fig. 22, Tab. 5).

RGG-motif mediated binding to transportin-1, has already been reported to be an important nuclear import mechanism in the case of transportin-1 binding of FUS.[4,6] These findings implicate that binding of the RGG motif may be a general import mechanism of proteins of the RGG/RG proteome. Arginine methylation of this RGG motif may also alter the affinity to transportin-1 of CIRP. Since the same sequence motif of FUS and CIRP binds to transportin-1, the binding site on transportin-1 may also be the same. Therefore an arginine methylation of the RGG motif may also weaken the CIRP binding to the ß-karyopherin.

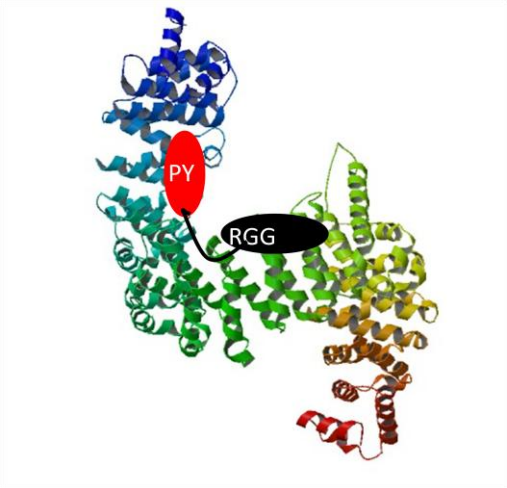


Fig. 29: Transportin-1 binding model based on FUS study
PDB: 2Z5J

It has been reported, that transportin-1 recognizes and binds PY nuclear localization signal and RGG motifs (Fig. 29). Since FUS-PY and FUS-RGG motifs have distinct binding sites on transportin-1 and the distance of binding sites on transportin-1 structure and in the flexible FUS protein are appropriate, a binding of both motifs is detected. This is also

supported by the fact, that there is no reduction of affinity of the full length FUS construct to transportin-1 observed. [4,6]

In CIRP, a C-terminal serine-tyrosine rich region was identified as new nuclear import signal. This region shows also an affinity to transportin-1. However, the affinity of CIRP-SY is reduced compared to CIRP-RGG or FUS-PY (Tab. 5). In addition, NMR data indicate only binding of CIRP-RGG in the full length construct (Fig. 22). These findings reveal an overlap of CIRP-RGG and CIRP-SY binding sites in transportin-1 and therefore a competition of the two binding motifs. Due to the higher affinity of CIRP-RGG, this motif binds in the full length construct (Fig. 30).

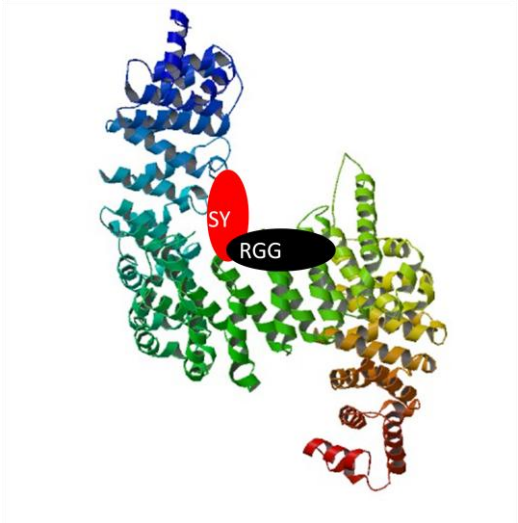


Fig. 30: Transportin-1 binding model based on CIRP data
PDB: 2Z5J

In summary, it can be stated that transportin-1 contains different binding sites for proteins. Affinity to transportin-1 depends on localization of these binding sites. If a protein contains two different transportin-1 binding sites and these binding sites overlap in transportin-1 structure (i.e. CIRP-RGG and CIRP-SY), the affinity of the full length protein is reduced compared to the affinity of the strongest binding motif, due to the competition of the sequence motifs. This results in a reduced nuclear import of the protein.

CIRP is mainly nuclear localized, which implicates that the nuclear import is efficient in healthy, physiological conditions.[72] Since this CIRP-SY region contains many putative phosphorylation sites, this post-translational modification may regulate the binding of the

CIRP-SY region. FUS, a member of the RG/RGG proteome, shows an insufficient nuclear import in case of a mutation in the PY-motif resulting in a reduced affinity and in case of a methylated RGG-motif. Due to this fact, a phosphorylation of CIRP-SY may enhance the transportin-1 affinity and therefore compensate for a reduced binding of CIRP-RGG in case of a methylation event, which is an abundant reaction in physiological conditions.
To further investigate this hypothesis, an atomic resolution structure of CIRP-SY/transportin-1 complex would reveal the exact binding site and ITC data of serine-to-glutamate mutants would give information about the affinity of CIRP-SY to transportin-1.

Taken all together, RGG sequence motifs seem to be favored transportin-1 binding sites of the RGG/RG proteome. The affinity of these RGG motifs is mainly regulated by PRMT1-dependent arginine methylation. In addition the nuclear import may be regulated by additional transportin-1 binding sites including PY nuclear localization signal, but also serine-tyrosine rich regions which may be in addition modulated by post-translational phosphorylation. This post-translational modification could enhance transportin-1 affinity and therefore lead to interplay of SY and RGG-binding to transportin-1 depending on type of modification. A further investigation of the effect of post-translational modifications of these binding sites could give insight in the exact regulation of nuclear import of proteins harboring arginine-glycine rich regions.

4.2. CIRP vs. RBM3

CIRP was used as a model system because it is a protein harboring RG/G motifs in the sequence and has important physiological functions. In addition, this protein is interesting, because it has a high sequence homology to another RG/G protein, the RBM3 (RNA binding motif protein 3, Fig. 31).

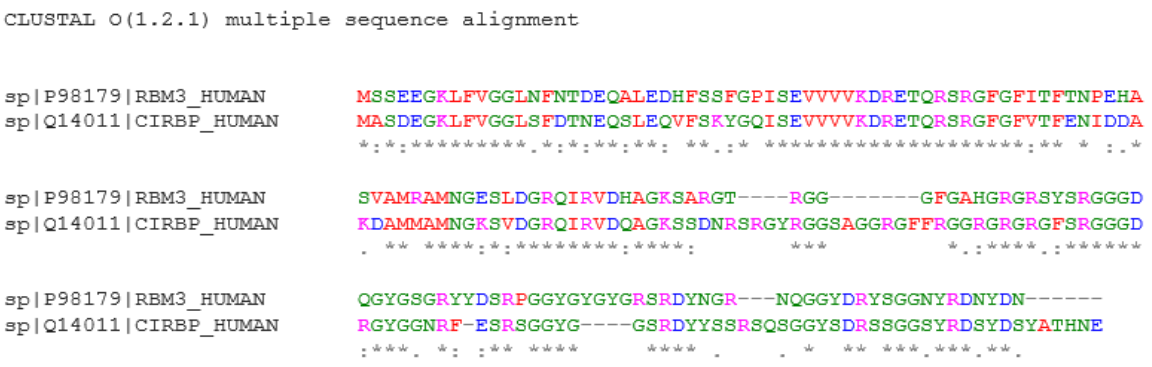


Fig. 31: Clustal o (1.2.1) multiple sequence alignment of human CIRP and RBM3

Both of these proteins are mainly nuclear localized and involved in cellular stress response. Although these proteins are involved in similar processes, they show strongly differing disease outcomes. CIRP is associated with patients suffering from cancer diseases with a poor prognosis, whereas RBM3 is mainly up-regulated in cancer patients that survived after chemotherapy. [72]

A deeper understanding of the effect of post-translational modifications and the nuclear import of these similar proteins would largely extend the knowledge about the involvement in diseases.

4.3. Experimental outlook

4.3.1. SY region in CIRP is a transportin-1 binding motif

The RGG region of CIRP was already identified as a transportin-1 binding motif for the nuclear-cytoplasmic shuttling. [5,48] CIRP contains an additional C-terminal motif which is interacting with the karyopherinß2 transportin-1, the SY-region. The binding of this domain is weaker than RGG binding, but the affinity is still high. In presence of RGG region, in the CIRP full length construct, the SY region does not seem to bind (Fig. 22). In case of a reduced affinity of CIRP-RGG, i.e. due to a mutation or a methylation event, the CIRP-SY may compensate for a reduced binding to transportion-1 and may maintain the inevitable nuclear-cytoplasmic shuttling of CIRP. This could be further analyzed by ITC measurements of methylated CIRP-RGG/CIRP full length binding to transportin-1.

4.3.2. Phosphorylation of CIRP-SY may regulate transportin-1 binding

In Lysate-Phosphorylation assays indicated a phosphorylation of serine residues in CIRP-SY (Fig. 23). Quick Change PCR mutagenesis was used to produce glutamate mutants of the distinct serine residues which mimic phosphorylation events due to a similar chemical environment. In further ITC experiments of transportin-1 and CIRP-SY, these mutants may clarify the role of the post-translational phosphorylation on transportin-1 binding and therefore the role on nuclear import.

4.3.3. Structural characterization of the transportin-1/CIRP-SY complex

The crystallization of CIRP-SY/transportin-1 was successful. Unfortunately, the resolution of the crystals was too low to obtain the structure. This may be due to the high flexibility of transportin-1. However, a screening of a higher amount of crystals may give insight into the structure, since there may be some crystals with a better resolution.

4.4. Metabolomic Profiling of Multiple Sclerosis samples

An untargeted approach of metabolomic profiling of different human biofluids (serum and CSF) was used for the distinction of healthy controls and patients suffering from multiple sclerosis/clinically isolated syndrome. A significant alteration in the metabolome of biofluids was observed in diseased individuals. Acetate, formate, glutamate and branched-chain amino acids were proposed as possible biomarkers for diagnosis of multiple sclerosis in CSF as well as serum. In addition, sarcosine level in serum and propylene glycol level in CSF were altered in patients.

The diagnosis of multiple sclerosis still remains difficult. [54–56,58] The approach of 1D ^{1}H NMR metabolomics may identify a new set of metabolites altered in MS/CIS.

Acetate is a small metabolite capable of passing the blood-brain barrier which is evidence for an increased level in serum as well as CSF. Furthermore acetate is metabolized to glutamate via the intermediate α-ketoglutarate, which may explain the deregulation of acetate and glutamate in the metabolic profile. [73,74] The generation of acetate by ethanol oxidation is associated with the production of reactive oxygen species, therefore with oxidation of fatty acids to aldehydes and in further consequence with the neurotoxicity and neurodegeneration. [75]

N-methyl-D-aspartate receptor (NMDAR) is a receptor on neurons and it is activated by glutamate, but also sarcosine, a metabolite increased in patients' serum, enhances the NMDAR function. A deregulation of this receptor activity may lead to neurotoxic effects. [76]

Studies of De Simone et al. (2013) indicate a neurotoxicity of high concentrations of branched-chain amino acids. [77]

Formate, a product of methanol metabolism, is known to be neurotoxic and increased concentrations were observed in serum and CSF. [78]

In high concentrations, propylene glycol causes neurotoxic effects. This metabolite is converted into pyruvic acid; therefore it is part of the glucose metabolism. [79]

Altogether, deregulated metabolites are involved in amino acid metabolism, but also energy metabolism, which has already been shown by other studies. [61–63]

The NMR-based metabolomics may be a promising tool for diagnosis of multiple sclerosis, a neurodegenerative disease characterized by neurological deficiencies in relapse form. Monitoring of disease by NMR-based metabolomics may help to predict relapses and to apply targeted therapies based on metabolic changes in biofluids. However, these applications need further proof by screening a higher amount of samples and by extending the study to a multicenter approach.

A following ADMA extraction step may enable, in addition to the identification of previously discussed biomarkers, an analysis of involvement of arginine methylation in the development of multiple sclerosis.

Bibliography

1. Paik, W. K. & Kim, S. Enzymatic methylation of protein fractions from calf thymus nuclei. *Biochem. Biophys. Res. Commun.* **29,** 14–20 (1967).
2. Tang, J. *et al.* PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. *J. Biol. Chem.* **275,** 7723–7730 (2000).
3. Lin, W. J., Gary, J. D., Yang, M. C., Clarke, S. & Herschman, H. R. The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J. Biol. Chem.* **271,** 15034–15044 (1996).
4. Dormann, D. *et al.* Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J.* **31,** 4258–4275 (2012).
5. De Leeuw, F. *et al.* The cold-inducible RNA-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor. *Exp. Cell Res.* **313,** 4130–4144 (2007).
6. Suárez-Calvet, M. *et al.* Monomethylated and unmethylated FUS exhibit increased binding to Transportin and distinguish FTLD-FUS from ALS-FUS. *Acta Neuropathol.* (2016). doi:10.1007/s00401-016-1544-2
7. Zou, L. *et al.* Correlation of SRSF1 and PRMT1 expression with clinical status of pediatric acute lymphoblastic leukemia. *J. Hematol. Oncol.* **5,** 42 (2012).
8. Avasarala, S. *et al.* PRMT1 Is a novel regulator of epithelial-mesenchymal-transition in non-small cell lung cancer. *J. Biol. Chem.* **290,** 13479–13489 (2015).
9. Frisone, P. *et al.* SAM68 : signal transduction and RNA metabolism in human cancer. *Biomed Res. Int.* **2015,** 39 (2015).
10. Thandapani, P., O'Connor, T. R., Bailey, T. L. & Richard, S. Defining the RGG/RG Motif. *Mol. Cell* **50,** 613–623 (2013).
11. Kiledjian, M. & Dreyfuss, G. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.* **11,** 2655–2664 (1992).
12. Hadian, K. *et al.* Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV rev protein. *J. Biol. Chem.* **284,** 33384–33391 (2009).
13. Shaw, D. J. *et al.* Identification of a self-association domain in the Ewing's sarcoma protein: A novel function for arginine-glycine-glycine rich motifs? *J. Biochem.* **147,** 885–893 (2010).
14. Souki, S. K. & Sandri-Goldin, R. M. Arginine methylation of the ICP27 RGG box regulates the functional interactions of ICP27 with SRPK1 and Aly/REF during

herpes simplex virus 1 infection. *J. Virol.* **83,** 8970–8975 (2009).

15. Rajyaguru, P., She, M. & Parker, R. Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. *Mol. Cell* **45,** 244–254 (2012).
16. Ishimaru, D. *et al.* Mechanism of regulation of bcl-2 mRNA by nucleolin and A+U-rich element-binding factor 1 (AUF1). *J. Biol. Chem.* **285,** 27182–27191 (2010).
17. Zhang, J., Tsaprailis, G. & Bowden, G. T. Nucleolin stabilizes Bcl-XL messenger RNA in response to UVA irradiation. *Cancer Res.* **68,** 1046–1054 (2008).
18. Chen, J., Guo, K. & Kastan, M. B. Interactions of nucleolin and ribosomal protein L26 (RPL26) in translational control of human p53 mRNA. *J. Biol. Chem.* **287,** 16467–16476 (2012).
19. Cho, E.-C. *et al.* Arginine methylation controls growth regulation by E2F-1. *EMBO J.* **31,** 1785–1797 (2012).
20. Gurunathan, G., Yu, Z., Coulombe, Y., Masson, J.-Y. & Richard, S. Arginine methylation of hnRNPUL1 regulates interaction with NBS1 and recruitment to sites of DNA damage. *Sci. Rep.* **5,** 10475 (2015).
21. Bedford, Mark T., Clarke, S. G. Protein Arginine Methylation in Mammals : Who , What , and Why. *Mol cell.* **33,** 1–13 (2009).
22. Bedford, M. T. & Richard, S. Arginine methylation: An emerging regulator of protein function. *Mol. Cell* **18,** 263–272 (2005).
23. Najbauer, J. & Aswad, D. W. Diversity of methyl acceptor proteins in rat pheochromocytoma (PC12) cells revealed after treatment with adenosine dialdehyde. *J. Biol. Chem.* **265,** 12717–12721 (1990).
24. Lee, H. W., Kim, S. & Paik, W. K. S-adenosylmethionine: protein-arginine methyltransferase. Purification and mechanism of the enzyme. *Biochemistry* **16,** 78–85 (1977).
25. Roeder, S. *et al.* SAM levels, gene expression of SAM synthetase, methionine synthase and ACC oxidase, and ethylene emission from N. suaveolens flowers. *Plant Mol. Biol.* **70,** 535–546 (2009).
26. Côté, J. & Richard, S. Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem.* **280,** 28476–28483 (2005).
27. Carson, N. a. J. *et al.* Homocystinuria: A new inborn error of Metabolism associated with Mental Deficiency. *Arch. Dis. Child.* **38,** 425–436 (1963).
28. Gibson, J. B., Carson, N. A. J. & Neill, D. W. Pathological findings in homocystinuria. *J. Clin. Pathol.* **17,** 427–437 (1964).

29. David, B. Y. & Piper, R. CCXIII . THE INHIBITION OF ESTERASES BY EXCESS SUBSTRATE . (1930).
30. Wang, L. *et al.* Modulation of cystathionine ß-synthase level regulates total serum homocysteine in mice. *Circ. Res.* **94,** 1318–1324 (2004).
31. Esse, R. *et al.* Protein arginine hypomethylation in a mouse model of cystathionine ß-synthase deficiency. *FASEB J.* **28,** 2686–2695 (2014).
32. Pollard, V. W. *et al.* A novel receptor-mediated nuclear protein import pathway. *Cell* **86,** 985–994 (1996).
33. Fridell, R. a, Truant, R., Thorne, L., Benson, R. E. & Cullen, B. R. Nuclear import of hnRNP A1 is mediated by a novel cellular cofactor related to karyopherin-beta. *J. Cell Sci.* **110 (Pt 1,** 1325–1331 (1997).
34. Lee, B. J. *et al.* Rules for Nuclear Localization Sequence Recognition by Karyopherinß2. *Cell* **126,** 543–558 (2006).
35. Zinszner, H., Albalat, R. & Ron, D. A novel effector domain from the RNA-binding protein TLS or EWS is required for oncogenic transformation by CHOP. *Genes Dev.* **8,** 2513–2526 (1994).
36. Zinszner, H., Sok, J., Immanuel, D., Yin, Y. & Ron, D. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J. Cell Sci.* **110,** 1741–1750 (1997).
37. Zinszner, H., Immanuel, D., Yin, Y., Liang, F. X. & Ron, D. A topogenic role for the oncogenic N-terminus of TLS: nucleolar localization when transcription is inhibited. *Oncogene* **14,** 451–61 (1997).
38. Vance, C. *et al.* Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science (80-.).* **323,** 1208–11 (2009).
39. Dormann, D. *et al.* ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *EMBO J.* **29,** 2841–2857 (2010).
40. Kwiatkowski, T. J. *et al.* Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science (80-.).* **323,** 1205–1208 (2009).
41. Neary D., et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 1546–1554 (1998).
42. Langenhove, V. Genetic contribution of FUS to frontotemporal lobar degeneration. *Neurology.* **74,** 366–71 (2010).
43. Sinha, R. *et al.* Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. *Mol. Cell. Biol.* **30,** 2762–2774 (2010).
44. J. Côté, FM. Boisvert, MC. Boulanger, MT. Bedford, S. R. Sam68 RNA Binding

Protein Is an In Vivo Substrate for Protein Arginine N-Methyltransferase 1. *Mol. Biol. Cell* **14,** 274–287 (2003).

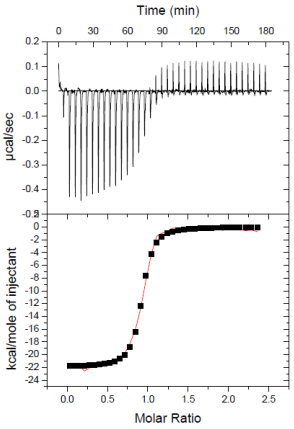
45. Al-Fageeh, M. B. & Smales, C. M. Cold-inducible RNA binding protein (CIRP) expression is modulated by alternative mRNAs. *RNA* **15,** 1164–1176 (2009).
46. Nishiyama, H. *et al.* Decreased expression of cold-inducible RNA-binding protein (CIRP) in male germ cells at elevated temperature. *Am. J. Pathol.* **152,** 289–96 (1998).
47. Nishiyama, H. *et al.* A glycine-rich RNA-binding protein mediating cold-inducible suppression of mammalian cell growth. *J. Cell Biol.* **137,** 899–908 (1997).
48. Aoki, K., Ishii, Y., Matsumoto, K. & Tsujimoto, M. Methylation of Xenopus CIRP2 regulates its arginine- and glycine-rich region-mediated nucleocytoplasmic distribution. *Nucleic Acids Res.* **30,** 5182–5192 (2002).
49. No Title. at <http://www.uniprot.org/uniprot/Q14011>
50. Clarke, C. J. & Haselden, J. N. Metabolic profiling as a tool for understanding mechanisms of toxicity. *Toxicol. Pathol.* **36,** 140–147 (2008).
51. Bingol, K., High, N. & Field, M. Multidimensional approaches to NMR-based metabolomics. *Anal Chem* **86,** 47–57 (2014).
52. Alonso, A., Marsal, S. & Julià, A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front. Bioeng. Biotechnol.* **3,** 23 (2015).
53. Dhar, S. *et al.* Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. *Sci. Rep.* **3,** 1311 (2013).
54. Lublin, F. D. *et al.* Defining the clinical course of multiple sclerosis : The 2013 revisions Defining the clinical course of multiple sclerosis The 2013 revisions. 1–10 (2014). doi:10.1212/WNL.0000000000000560
55. Polman, C. H. *et al.* Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann. Neurol.* **69,** 292–302 (2011).
56. Leary, S. M., Porter, B. & Thompson, a J. Multiple sclerosis: diagnosis and the management of acute relapses. *Postgrad. Med. J.* **81,** 302–308 (2005).
57. Shah, P. Symptomatic management in multiple sclerosis. *Ann. Indian Acad. Neurol.* **18,** S35–S42 (2015).
58. A. Charil, TA Yousry, M Rovaris, F Barkhof, N De Stefano, F Fazekas, DH Miller, X Montalban, JH Simon, C Poman, M. F. MRI and the diagnosis of multiple sclerosis: expanding the concept of ‘no better explanation’. *Lancet Neurol* 841–52 (2006).
59. Gebregiworgis, T. *et al.* Potential of urinary metabolites for diagnosing multiple sclerosis. *ACS Chem. Biol.* **8,** 684–690 (2013).

60. Gebregiworgis, T. *et al.* A Urinary Metabolic Signature for Multiple Sclerosis and Neuromyelitis Optica. *J. Proteome Res.* (2016). doi:10.1021/acs.jproteome.5b01111
61. Tavazzi, B. *et al.* Serum Metabolic Profile in Multiple Sclerosis Patients. *Mult. Scler. Int.* **2011,** 1–8 (2011).
62. Cocco, E. *et al.* 1H-NMR analysis provides a metabolomic profile of patients with multiple sclerosis. *Neurol. - Neuroimmunol. Neuroinflammation* **3 ,** 1–10 (2016).
63. Noga, M. J. *et al.* Metabolomics of cerebrospinal fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis. *Metabolomics* **8,** 253–263 (2012).
64. Mangalam, A. & Poisson, L. Profile of Circulatory Metabolites in a Relapsing-remitting Animal Model of Multiple Sclerosis using Global Metabolomics. *J. Clin. …* 1–25 (2013). doi:10.4172/2155-9899.1000150.Profile
65. Dickens, A. M. *et al.* A type 2 biomarker separates relapsing-remitting from secondary progressive multiple sclerosis. *Neurology* **83,** 1492–1499 (2014).
66. Weber, S. *et al.* PRMT1-mediated arginine methylation of PIAS1 regulates STAT1 signaling. *Genes and Development* **23,** 118–132 (2009).
67. Avasarala, S. *et al.* PRMT1 is a Novel Regulator of Epithelial-Mesenchymal-Transition in Non-Small Cell Lung Cancer. *J. Biol. Chem.* **290,** jbc.M114.636050 (2015).
68. Bicker KL., Obianyo O., Rust HL., T. P. A combinatorial approach to characterize the substrate specificity of protein arginine methyltransferase 1. *Mol Biosyst* **7,** 48–51 (2011).
69. Goulet, I., Gauvin, G., Boisvenue, S. & C??t??, J. Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. *J. Biol. Chem.* **282,** 33009–33021 (2007).
70. Cloarec, O. *et al.* Evaluation of the Orthogonal Projection on Latent Structure Model Limitations Caused by Chemical Shift Variability and Improved Visualization of Biomarker Changes in 1H NMR Spectroscopic Metabonomic Studies. *Anal. Chem.* **77,** 517–526 (2005).
71. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5,** 67–81 (1995).
72. Zhu, X., Bührer, C. & Wellmann, S. Cold-inducible proteins CIRP and RBM3, a unique couple with activities far beyond the cold. *Cell. Mol. Life Sci.* (2016). doi:10.1007/s00018-016-2253-7

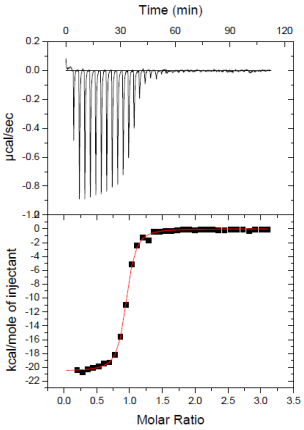
73. Wang, J. *et al.* Metabolic products of [2-(13) C]ethanol in the rat brain after chronic ethanol exposure. *J. Neurochem.* **127,** 353–64 (2013).
74. Fein, G. and Meyerhoff, D. Ethanol in Human Brain by Magnetic Resonance Spectroscopy: Correlation With Blood and Breath Levels, Relaxation and Magnetization Transfer. *Alcohol Clin Exp Res* **8,** 1227–1235 (2000).
75. Hernández, J. A., López-Sánchez, R. C. & Rendón-Ramírez, A. Lipids and Oxidative Stress Associated with Ethanol-Induced Neurological Damage. *Oxid. Med. Cell. Longev.* **2016,** (2016).
76. Zhang, H. X., Lyons-Warren, A. & Thio, L. L. The glycine transport inhibitor sarcosine is an inhibitory glycine receptor agonist. *Neuropharmacology* **57,** 551–555 (2009).
77. De Simone, R. *et al.* Branched-chain amino acids influence the immune properties of microglial cells and their responsiveness to pro-inflammatory signals. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1832,** 650–659 (2013).
78. Dorman, D. C., Bolon, B. & Morgan, K. T. The toxic effects of formate in dissociated primary mouse neural cell cultures. *Toxicol. Appl. Pharmacol.* **122,** 265–72 (1993).
79. Belavy, D., Sunn, N., Lau, Q. & Robertson, T. Absence of neurotoxicity with perineural injection of ultrasound gels: assessment using an animal model. *BMC Anesthesiol.* **13,** 18 (2013).
80. https://www.embl.de/pepcore/pepcore_services/cloning/pdf/pETM-11.pdf.

Supplementary

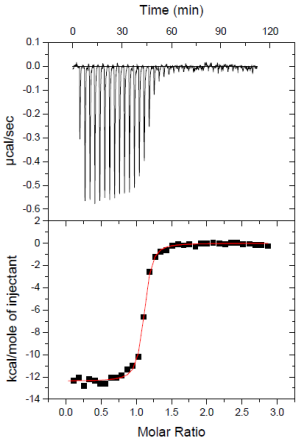
TNPO1 vs FUS-PY



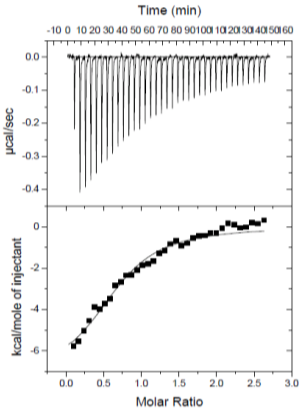
TNPO1Δlinker vs FUS-PY



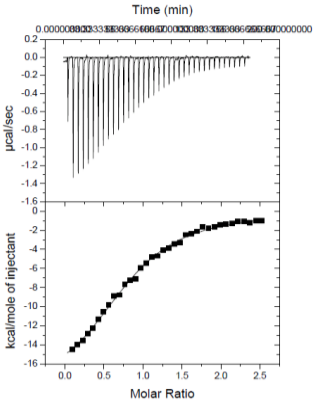
TNPO1/SY vs. FUS-PY



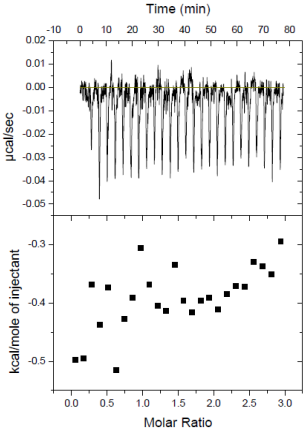
TNPO1 vs CIRP-SY



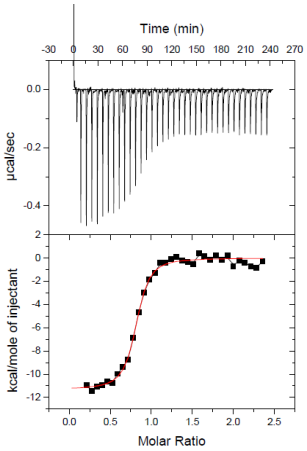
TNPO1Δlinker vs. CIRP-SY



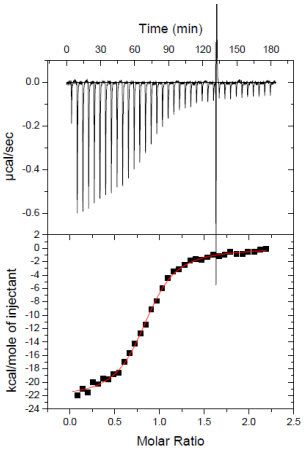
TNPO1/PY vs. CIRP-SY



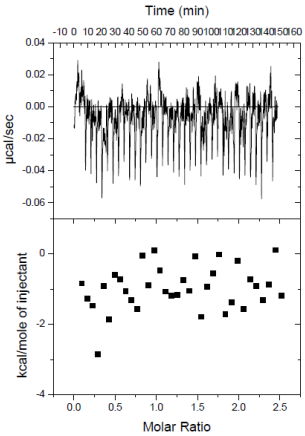
TNPO1 vs CIRP-RGG



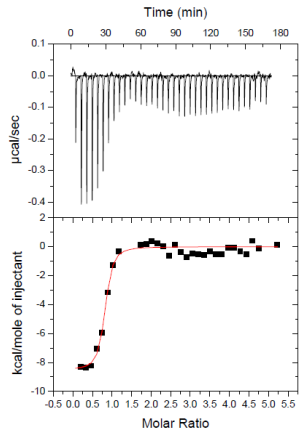
TNPO1Δlinker vs. CIRP-RGG



TNPO1/RGG vs CIRP-SY



TNPO1/SY vs. CIRP-RGG



Suppl.Fig. 1: ITC final figures of Transportin-1
Binding of different CIRP constructs and FUS-PY to transportin-1 with or without flexible loop region is shown.

Suppl. Tab. 1: Thermodynamic parameters of transportin-1 binding

	K	ΔH	ΔS
TNPO1 vs. CIRP-SY	$2.53e^{5} \pm 4.97e^{4}$	-7321 ± 582	-1.14
TNPO1 vs. FUS-PY	$2.08e^{7} \pm 2.18e^{6}$	-21910 ± 154	-40.0
TNPO1 vs. CIRP-RGG	$1.34e^{7} \pm 2.10e^{6}$	-11310 ± 163	-5.33
TNPO1Δlinker vs. CIRP-SY	$1.20e^{5} \pm 6.69e^{3}$	-19420 ± 447	-45.3
TNPO1Δlinker vs. FUS-PY	$1.62e^{7} \pm 1.10e^{6}$	-20560 ± 98.7	-36.0
TNPO1Δlinker vs. CIRP-RGG	$2.75e^{6} \pm 1.56e^{5}$	-22300 ± 181	-45.3
TNPO1/CIRP-SY vs. FUS-PY	$3.02e^{7} \pm 3.37e^{6}$	-12380 ± 67.9	-7.30
TNPO1/FUS-PY vs. CIRP-SY		No binding	
TNPO1/CIRP-SY vs. CIRP-RGG	$1.44e^{7} \pm 4.82e^{6}$	-8537 ± 268	3.65
TNPO1/CIRP-RGG vs. CIRP-SY		No binding	