JYU DISSERTATIONS 346

Chandan Jung Thapa

# Structural Insight of PP2A Inhibitor Proteins and their Interaction with PP2A A- and B56-subunits

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF MATHEMATICS
AND SCIENCE

JYU DISSERTATIONS 346

Chandan Jung Thapa

# Structural Insight of PP2A Inhibitor Proteins and their Interaction with PP2A A- and B56-subunits

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi helmikuun 12. päivänä 2021 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
on February 12 2021 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2021

Editors
Varpu Marjomäki
Department of Biological and Environmental Science, University of Jyväskylä
Ville Korkiakangas
Open Science Centre, University of Jyväskylä



Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-8486-1

ISBN 978-951-39-8486-1 (PDF)
URN:ISBN:978-951-39-8486-1
ISSN 2489-9003

# ABSTRACT

Thapa, Chandan Jung
Structural insight of PP2A inhibitor proteins and their interaction with PP2A A- and B56-subunits
Jyväskylä: University of Jyväskylä, 2021, 53 p.
(JYU Dissertations
ISSN 2489-9003; 346)
ISBN 978-951-39-8486-1 (PDF)
PP2A inhibiittoriproteiinien rakenteet ja vuorovaikutus PP2A A- ja B56-alayksiköiden kanssa
Diss.

Protein phosphatase 2A (PP2A), the principal Serine/Threonine phosphatase, functions as a tumor suppressor. In most of the cancers, the PP2A tumor suppressor activity is inhibited either genetically or by the overexpression of PP2A inhibitor proteins. It is therapeutically tempting idea to reactivate inhibited PP2A by small molecule modulators, For that, a better understanding of the structural and molecular framework of PP2A inhibition is required. In this study, we have characterized the structural properties of PP2A inhibitor proteins ARPP-19, ARPP-16, and ENSA, and their interaction with A- and B56-subunits of PP2A by combining NMR spectroscopy with SAXS and microscale thermophoresis. The results reveal that both ENSA and ARPP proteins are intrinsically disordered, but not completely random coil as they have three regions having the propensity to form transient α-helical structures. Both ARPPs and ENSA were observed to interact with PP2A A-subunit with modest affinity, whereas, the interaction with B56-subunit is weak and transient. When ARPP and ENSA proteins are phosphorylated by Gwl/MAST3 kinase they inhibit PP2A during mitosis. The phosphomimetic mutants resembling phosphorylation of ARPPs showed increased affinity towards both A- and B56- subunits, while the corresponding phosphomimetic mutants of ENSA failed to bind to all other B56 subunits except B56α. Two distinct interaction modes of ARPPs and ENSA with the PP2A A-subunit were identified. In ARPPs, a second transient α-helix including its flanking region forms an A-subunit binding motif, whereas ENSA interacts with A-subunit using an extended region comprising all three transient α-helical regions. Together, these studies suggest that intrinsically disordered ARPPs and ENSA bind to PP2A transiently using preformed structural elements. Altogether, our results provide a crucial step towards the understanding the molecular bases behind PP2A inhibition, which provide a foundation for the development of novel and clinically feasible PP2A targeted therapies.

Keywords: ARPP-16/ARPP-19; ENSA; Intrinsically Disordered proteins; NMR, Protein Phosphatase 2A, SAXS.

*Chandan Jung Thapa, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland*

# TIIVISTELMÄ

Thapa, Chandan Jung
PP2A inhibiittoriproteiinien rakenteet ja vuorovaikutus PP2A A- ja B56-alayksiköiden kanssa
Jyväskylä: Jyväskylän yliopisto, 2021, 53 p.
(JYU Dissertations
ISSN 2489-9003; 346)
ISBN 978-951-39-8486-1 (PDF)
Diss.

Proteiinifosfataasi 2A (PP2A) on merkittävä kasvunrajoittimena toimiva Seriini/Treoniini-fosfataasi. Monissa syöpätyypeissä PP2A:n kasvunrajoitinaktiivisuus on estetty joko geneettisesti tai PP2A:n toiminnan estävien proteiinien avulla. Terapeuttisesti on houkuttelevaa uudelleen aktivoida PP2A estämällä inhibiittori proteiinien toiminta pienmolekyyleillä. Tätä varten on kuitenkin oleellista ymmärtää PP2A:n toiminnan estävien inhibiittoriproteiinien molekulaarinen ja rakenteellinen mekanismi. Tässä tutkimuksessa olemme karakterisoineet PP2A:n inhibiittoriproteiinien ARPP-19, ARPP-16 ja ENSA rakenteellisia ominaisuuksia sekä näiden vuorovaikutusmekanismia PP2A:n A- ja B56-alayksiköiden kanssa yhdistämällä NMR-spektroskopiaa, SAXS:a ja mikroskaalan termoforeesia. Tutkimuksista saadut tulokset osoittavat, että sekä ARPP-että ENSA-proteiinit ovat luontaisesti rakenteettomia proteiineja, mutta näillä proteiineilla on kuitenkin taipumus muodostaa hetkellisesti α-kierteitä. Sekä ARPP-ja ENSA-proteiinien havaittiin sitoutuvan PP2A:n A-alayksikköön kohtalaisen vahvasti, kun taas sitoutuminen B56-alayksiköihin on heikompaa. Gwl/MAST3-kinaasien fosforyloima ARPP-19 ja ARPP-16 estävät PP2A:n toiminnan mitoosin aikana. Gwl/MAST3-kinaasien fosforylointia matkivien mutaatioiden havaittiin kasvattavan ARPP-19- ja ARPP-16 affiniteettia sekä A- että B56-alayksiköihin. Vastaavat mutaatiot ENSA-proteiinissa sen sijaan esti sitoutumisen kaikkiin muihin B56-alayksiköhin paitsi B56α-alayksikköön. ARPP-ja ENSA-proteiinien havaittiin sitoutuvan PP2A A-alayksikköön eri tavoilla. ARPP-proteiini käyttää sitoutumisessa toista α-kierrettä sekä tämän viereisiä aminohappoja. ENSA proteiini puolestaan hyödyntää sitoutuessaan kaikkia kolmea hetkellisesti muodostuvaa α-kierrettä. Nämä tulokset yhdessä osoittavat, että rakenteettomat ARPP-ja ENSA-proteiinit sitoutuvat PP2A-proteiiniin käyttäen ennalta muodostuneita rakenteellisia alueita. Yhteenvetona voidaan todeta, että meidän tutkimusten tulokset antavat oleellista tietoa, joka auttaa kehittämään uusia ja kliinisesti käyttökelpoista PP2A-proteiiniin kohdennettuja lääkehoitoja.

Avainsanat: ARPP-16/ARPP19; ENSA; luontaisesti epäjärjestäytynyneet proteiinit, NMR, Proteiinifosfataasi 2A, SAXS.

*Chandan Jung Thapa, Jyväskylän yliopisto, Bio- ja ympäristötieteiden laitos PL 35, 40014 Jyväskylän yliopisto*

**Author's address** Chandan Jung Thapa
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland
chandan.j.thapa@jyu.fi

**Supervisors** Prof. Perttu Permi
Department of Biological and Environmental Science
Department of Chemistry
NanoScience Center (NSC)
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Ulla Pentikäinen
Institute of Biomedicine
Turku Bioscience
Tykistökatu 6
FI-20520 Turku BioScience
Finland

**Reviewers** Docent Tommi Kajander
Institute of Biotechnology
Viikinkaari 1
P.O. Box 65
FI-00014 University of Helsinki
Finland

Assoc. Prof. Mikko Metsä-Ketelä
Department of Biochemistry
FI-20014 University of Turku
Finland

**Opponent** Prof. Mark S. Johnson
Department of Biochemistry
Tykistökatu 6
FI-20250 Åbo Academy University
Finland

# CONTENTS

## LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original papers, which will be referred to in the text by their Roman numerals:

I. Chandan J. Thapa, Tatu Haataja, Ulla Pentikäinen and Perttu Permi 2020. $^{1}H$, $^{13}C$ and $^{15}N$ NMR chemical shift assignments of cAMP-regulated phosphoprotein-19 and -16 (ARPP-19 and ARPP-16). *Biomolecular NMR Assignments* 14: 227-231.

II. Chandan Thapa, Pekka Roivas, Tatu Haataja Perttu Permi and Ulla Pentikäinen. Intrinsically disordered PP2A inhibitor proteins', ARPP-16/19, interaction mechanism with PP2A. (Submitted manuscript)

III. Chandan Thapa, Pekka Roivas, Tatu Haataja Perttu Permi and Ulla Pentikäinen. Endogenous PP2A inhibitor protein ENSA, structurally related to ARPP-19, differentially interact with PP2A. (Manuscript)

## CONTRIBUTION TO THE WORK

I. I was involved in the entire project. I cloned APRP-16/ARPP-19, optimized their production and purified the $^{15}N$ and $^{15}N$-$^{13}C$ ARPP16 and ARPP-19 for the Nuclear Magnetic Resonance (NMR). Prof. Permi performed NMR measurements. I performed the assignment of mainchain $^{1}H$, $^{13}C$, and $^{15}N$ chemical shifts. I wrote the manuscript together with my supervisors.

II. I cloned ARPP and B56 constructs. I optimized the expression and purified both labelled and unlabelled ARPP constructs. I was involved in SAXS data collection. I designed the NMR titration experiments and the data were collected by Prof. Permi. I performed the analysis of SAXS and NMR data. I participated in designing and performing MST experiments together with my co-author Pekka Roivas. I analyzed all of the MST data. I wrote the manuscript together with my supervisors.

III. I expressed and purified the $^{15}N$ and $^{15}N$-$^{13}C$ ENSA. I participated in SAXS data collection. I participated in designing MST and NMR titration experiments. I performed the analysis of all data and drafted manuscript together with my supervisors.

## ABBREVIATIONS

| | |
|---|---|
| ARPP | cAMP regulated phosphoprotein |
| Cdc25 | Cell division cycle protein 25 |
| Cdk1 | Cyclin-dependent kinase 1 |
| CSP | Chemical shift perturbation |
| ENSA | α-Endosulfine |
| EOM | Ensemble optimization method |
| Gwl | Greatwall kinase |
| HEAT | Huntingtin/elongation/A subunit/TOR motif |
| IDP | Intrinsically disordered protein |
| MAST3 | Microtubule-associated serine/threonine-protein kinase 3 |
| MST | Microscale thermphoresis |
| NMR | Nuclear magnetic resonance |
| PKA | Protein kinase A |
| PP2A | Protein phosphatase 2A |
| PTM | Post translational modification |
| SAXS | Small angle X-ray scattering |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SEC | Size exclusion chromatography |
| SSP | Secondary structure propensity |

# 1 INTRODUCTION

## 1.1 Protein Phosphorylation

Post translational modifications (PTMs) are reversible or irreversible chemical modifications of proteins following protein biosynthesis that enable diversification of proteins' function beyond what is dictated by gene transcripts. The common modifications that occur in proteins include phosphorylation, glycosylation, ubiquitination, nitrosylation, lipidation, acetylation, methylation and proteolysis and have remarkable impact on normal cell biology and pathogenesis. Among the different modifications, phosphorylation is the most common PTM in proteins (Khoury *et al.* 2011). Phosphorylation is defined as the covalent addition of phosphate group to a side chain of a suitable amino acid on a protein (Fig. 1). The phosphate group in phosphorylated amino acid has a large hydrated shell and carries two negative charges, which makes phospho-amino acid a new chemical entity compared to the natural amino acids. These unique properties of phosphorylated amino acids in proteins can be exploited to regulate crucial biological functions like, protein stabilization, protein-protein interaction, cellular signal transduction, and allosteric regulation of proteins and enzymes (Hunter 2012).

Phosphorylation of hydroxyl-containing amino acids (serine, threonine and tyrosine) is predominant in eukaryotes, but phosphorylation of other amino acids like histidine, arginine, lysine, cysteine, aspartate and glutamate also exist. Phosphate ester bonds in pSer, pThr and pTyr are relatively more stable than the phosphorothioate bond in pCys and phosphoramidate bonds in pHis, pLys and pArg (Hardman *et al.* 2019). The quantification of phosphorylated amino acids in chicken cells by Hunter and Sefton suggested a very low abundance of tyrosine phosphorylation. In normal cells, the relative abundance of tyrosine, serine, and threonine phosphorylation was 0.03%, 92%, and 8%, respectively (Hunter and Sefton 1980). Later mass spectrometer-based analysis of 6,600 *in vivo* phosphorylation sites allowed the re-evaluation of the relative abundance of tyrosine, serine, and threonine phosphorylation, which was 1.8%, 86.4%, and

11.8%, respectively (Olsen *et al.* 2006). According to the phosphorylation data recorded in PhosphoSitePlus® (www.phosphosite.org), 15.2 % of Y, 60.1% of S, and 24.7% of T out of 294,097 non-redundant phosphorylation site are phosphorylated (Hornbeck *et al.* 2015).

Reversible protein phosphorylation is a crucial process that occurs in the cells, according to the physiological demands. The coordinated regulation of phosphorylation dependent pathways is required and is maintained by kinases and phosphatases. Alteration on the activities of any one of the enzyme classes has a significant impact on how the cell responds to its environment and can cause different pathological conditions, including cancer. In human, 518 kinases (90 phosphotyrosine kinases (PTKs) and 428 phosphoserine/threonine kinases (PSKs)) mediate protein phosphorylation and regulate most of the cellular signal transduction (Manning *et al.* 2002). On the other hand, protein dephosphorylation is mediated by 254 (241 active) phosphatases, including 112 phosphotyrosine phosphatases (PTPs) and 42 serine/threonine phosphatases (PSTPs) as listed on the Human Dephosphorylation Database (DEPOD) updated on 2019 (Damle and Köhn 2019).

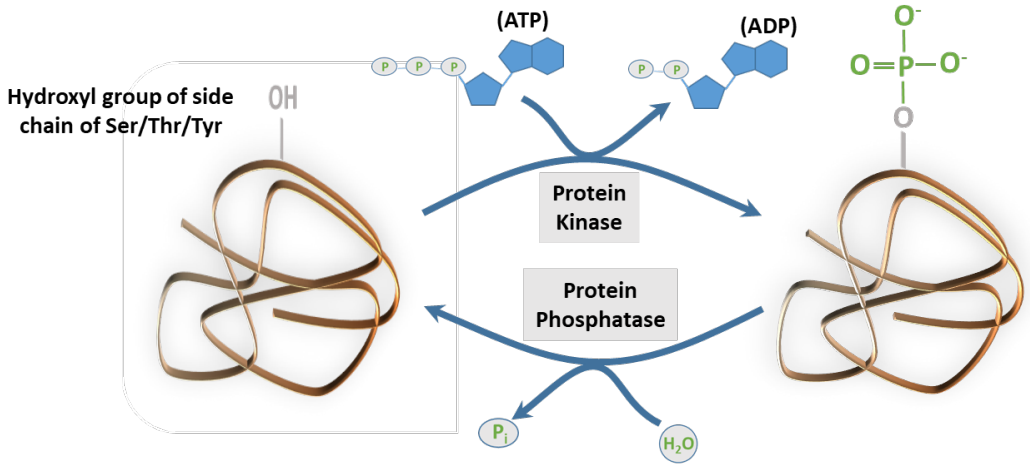


FIGURE 1 Protein phosphorylation. A phosphate group is transferred from adenosine triphosphate (ATP) to a hydroxyl group on Ser, Thr or Tyr residues of the protein by a kinase and can be removed by a phosphatase.

## 1.2 Protein Phosphatase 2A (PP2A)

PP2A is one of the major Serine/Threonine phosphatases and is responsible for a large portion of phosphatase activity in many cells (Mumby and Walter 1993, Wera and Hemmings 1995). It is a heterotrimeric protein and consists of scaffold/structural (A), catalytic (C) and regulatory (B) subunits. The subunit A and C both have two isoforms, whereas, subunit B has 23 isoforms and are divided into four different families (B, B′, B″ and B‴). PP2A is structurally complex and exists in dimeric and trimeric forms. PP2A core enzyme or the

dimeric form is composed of A- and C-subunits, whereas the holoenzyme or the trimeric form consists of core enzyme and one of the alternative regulatory subunit (Mumby and Walter 1993, Kamibayashi *et al.* 1994, Cohen 1997, Janssens and Goris 2001). The assembly of one A, one C and one B subunit can give 92 different heterotrimers (Fig. 2) (Haesen *et al.* 2014). The localization, specificity and regulation of PP2A largely depends on the nature of the B-subunit associated with the holoenzyme. Due to the variability in the PP2A holoenzyme composition, PP2A can act on large number of proteins or substrates and influence most of the cellular mechanisms including cell cycle, cell mobility, cytoskeleton dynamics, signal transduction, cell proliferation and apoptosis (Janssens and Goris 2001, Eichhorn *et al.* 2009, Haesen *et al.* 2014, Kauko and Westermarck 2018). Upon interaction with a protein encoded by DNA tumor virus (SV40 protein small T antigen/polyoma small T antigen) and PP2A inhibitor (okadaic acid), PP2A is functionally inactivated in cancers. This led to the identification of PP2A as a tumor suppressor (Mumby and Walter 1991, Chen *et al.* 2004, Van Hoof and Goris 2004, Janssens *et al.* 2005).

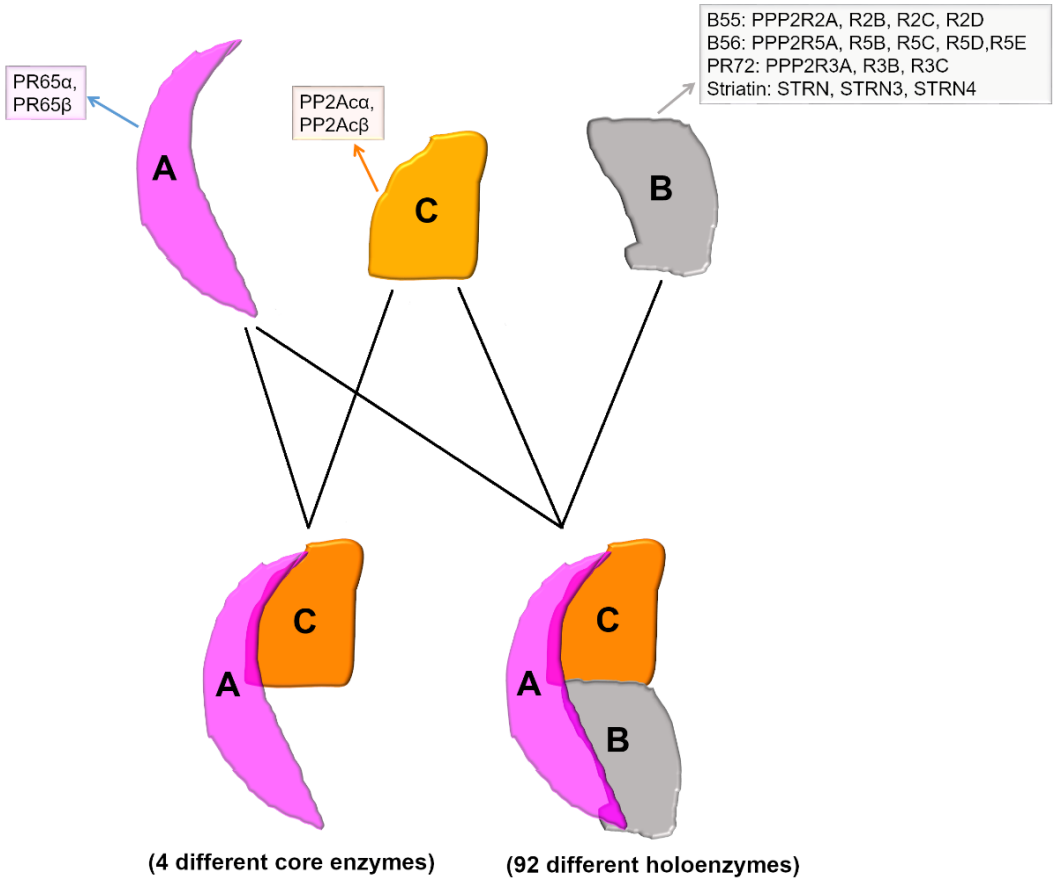


FIGURE 2 Schematic representation of PP2A holoenzyme diversity. Considering the existence of multiple isoforms of each subunits, 4 different AC dimeric core enzyme complex can be formed and 92 different heterotrimeric holoenzymes can be assembled.

### 1.2.1 PP2A subunits and their mutation and aberrant expression in cancers

#### 1.2.1.1 Scaffolding A-subunit (PR65)

PP2A scaffolding A-subunit (PR65) exists in two isoforms, PR65α and PR65β, which are 86% identical. They are encoded by two distinct genes, *PPP2R1A* and *PPP2R1B* (Hemmings *et al.* 1990). In normal tissues, about 90% of the core and/or holoenzyme assemblies of PP2A contains alpha isoform of A-subunit. The expression of PR65α is high in all tissues, whereas PR65β expression is high only in testis and low in most of the tissues (Zhou *et al.* 2003). Despite 86% sequence identity between PR65α and PR65β, PR65β knockout mouse is embryonically lethal (Ruediger *et al.* 2011). It seems the holoenzyme containing PR65α and PR65β are involved in different developmental stages, as indicated by high levels of PR65β during oogenesis and late embryogenesis in *Xenopus laevis*, whereas levels of PR65α expression was low until larval stage, then increased dramatically (Bosch *et al.* 1995). In the holoenzyme assembly, PP2A A-subunit serves as the assembly base on which the catalytic C-subunit can interact and aid the interaction with regulatory subunits and substrates, whereas A-subunit regulates substrate interaction in its dimeric form (Price and Mumby 2000). The structure of PR65 is composed of 15 tandem repeats of 39 amino acids, known as HEAT (huntingtin/elongation/A subunit/TOR) motif (Fig. 3). Each HEAT repeat motif is characterized by a pair of antiparallel helices connected by inter- and intra- repeat loops. These 15 HEAT repeats of PR65 subunit are organized to form extended and curved structure (Hemmings *et al.* 1990, Groves *et al.* 1999). The structural analysis of PR65 reveals that the catalytic subunit binds to HEAT repeat 11-15 of PR65 and the regulatory subunit binds to HEAT repeat 1-10 (Fig. 4) (Xu *et al.* 2006, Cho and Xu 2007, Shi 2009). The isoforms of PR65, PR65α and PR65β, interact differently with B and C subunits. In an *in vitro* binding assay, PR65β did not bind to Bα and B' subunits and binding to B" and C-subunit was also reduced in comparison with PR65α (Seeling *et al.* 1999). Whereas, binding was detected between B'α and PR65β in the *in vivo* experiment and there was no difference in binding to other subunits (Zhou *et al.* 2003).

The aberrant expression and mutation of PR65 has been reported in different types of cancer (Table 1). The reported cancer associated mutations of scaffolding subunit (PR65α) are Glu64 to Asp (E64D) substitution in a lung carcinoma, Glu64 to Gly (E64G) substitution in a breast carcinoma, and Arg418 to Trp (R418W) substitution in a melanoma (Calin *et al.* 2000). The mutations E64D and E64G were defective in holoenzyme formation, primarily with B56γ, that leads to constitutive AKT (RAC-alpha serine/threonine-protein kinase) phosphorylation and cell transformation (Ruediger *et al.* 2001, Chen *et al.* 2005). Also, PR65α mutations occur with high frequency (up to 43.2%) in uterine carcinoma and type II serous endometrial cancer. Some of the high frequency mutations are R182G and R183G/W mutation detected in ovarian clear cell carcinoma (Jones *et al.* 2010), P179R/S256F in serous endometrial carcinoma cell line and P179L in primary endometrial serous cancer (McConechy *et al.* 2011, Nagendra *et al.* 2012). PR65α mutations in endometrial and ovarian cancers are

all heterozygous and clustered in HEAT repeat 5 and 7. A biochemical analysis revealed that most of the PR65α mutants contribute to the cancer development by disturbing the ability to bind the B-subunit and by inducing functional haploinsufficiency (McConechy *et al.* 2011, Haesen *et al.* 2016). In a recent study, it has been reported that the overexpression of PR65α enhanced cell proliferation in endometrial and ovarian cell carcinoma. In addition, W257G mutant of PR65α increased migration of cancer cells (Jeong *et al.* 2016). Interestingly, the level of PR65α subunit was reduced 10-fold in approximately 43% of human gliomas, affecting core and holoenzyme enzyme formation (Colella *et al.* 2001).

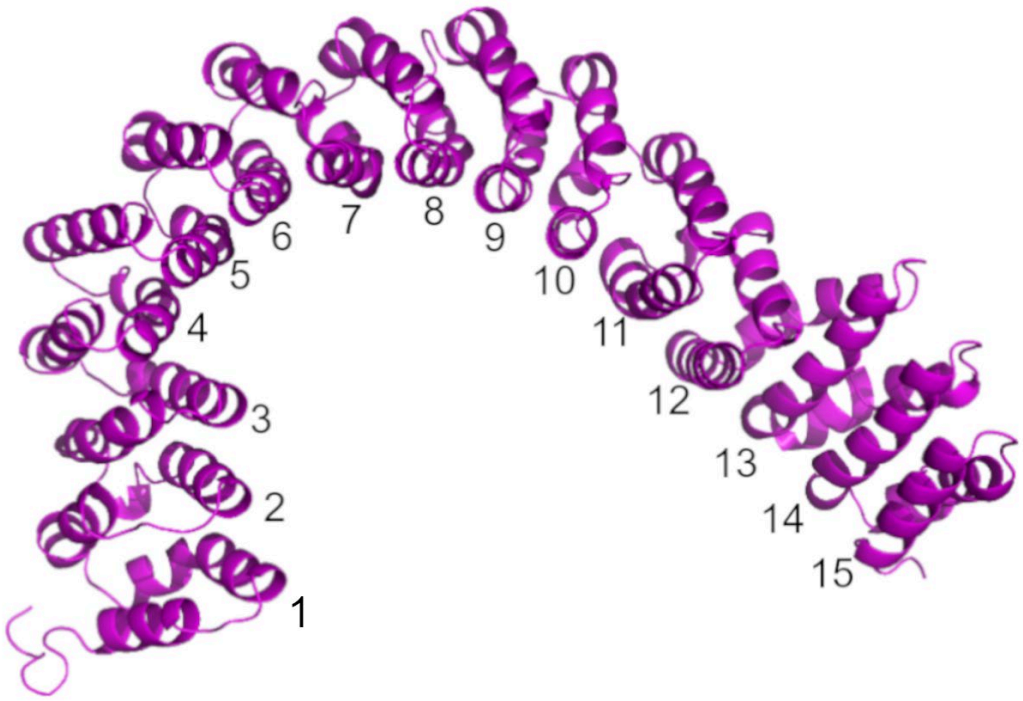


FIGURE 3 Crystal structure of PP2A A-subunit showing the position of consensus HEAT repeat motifs (Protein data bank (PDB) file 1B3U).

Analogous to PR65α, sporadic mutations of PR65β have been found in breast, lung, liver and colon cancers (Table 1) (Wang *et al.* 1998, Takagi *et al.* 2000, Meeusen and Janssens 2017). PR65β mutants lose their ability to form holoenzyme with all B-type subunits tested (B55α, B56γ and PR72) or failed to bind and dephosphorylate the small GTPase (Guanosine Triphosphate hydrolase) RalA at position Ser183 and Ser194, leading to RalA over-activation through hyperphosphorylation (Sablina *et al.* 2007). Alternate splicing of PR65β transcripts (skipping of exons2/3, 3, 9) were observed in B-cell chronic lymphocytic leukemia (B-CLL). The splice variants interfere with the assembly of the functional trimeric PP2A holoenzyme and resulted in the loss of PP2A activity (Kalla *et al.* 2007).

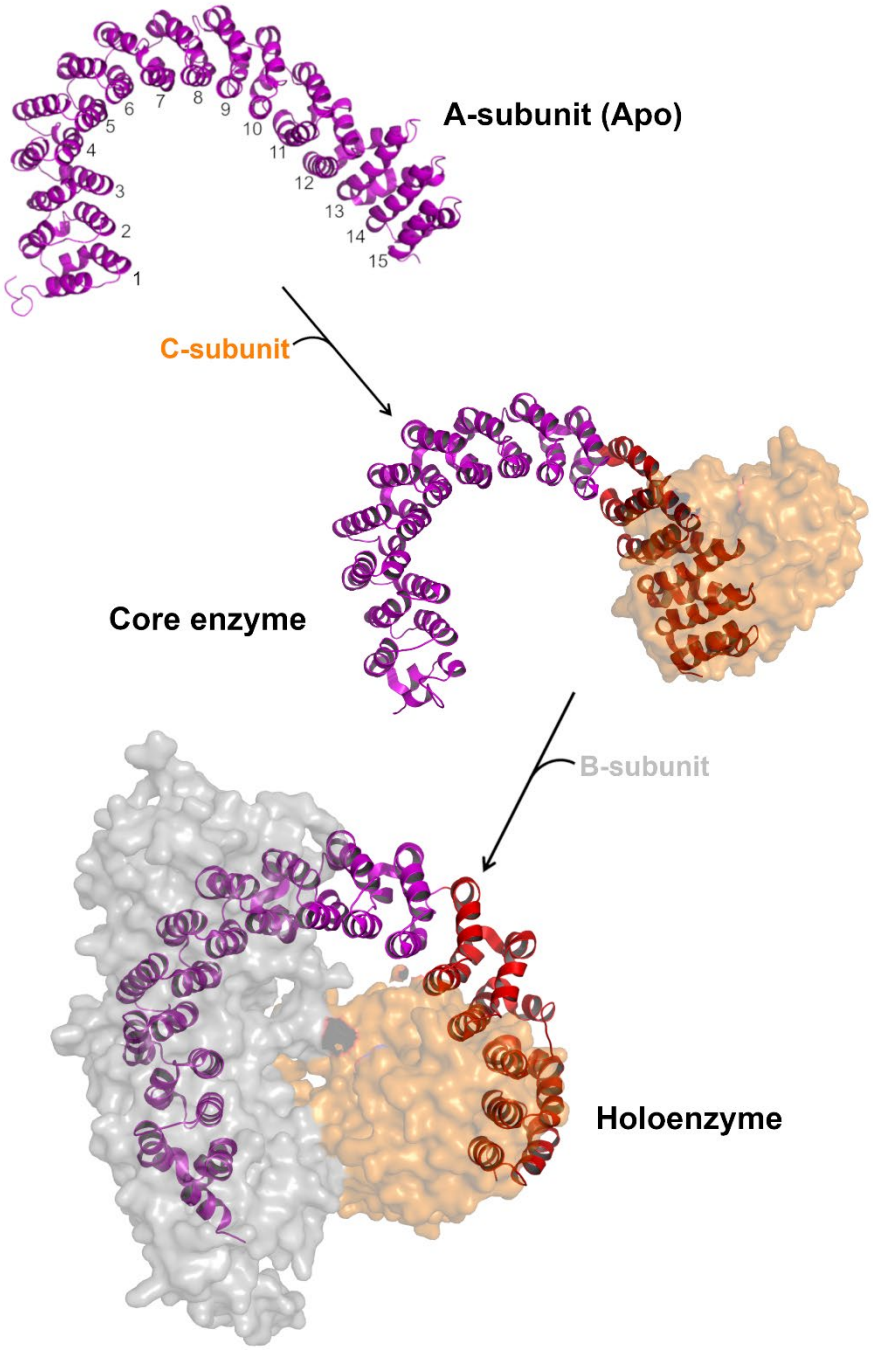


FIGURE 4 Assembly of PP2A holoenzyme containing B56 subunit. In the first step of holoenzyme assembly, the catalytic subunit associates with the HEAT repeat 11-15 of scaffolding subunit (colored red) to form the PP2A core enzyme. Then, a regulatory B56 subunit associates to the PP2A core enzyme to form a holoenzyme. The scaffolding A -subunit (PR65α) is shown with magenta, regulatory B56γ subunit shown with gray and the catalytic Cα-subunit with orange. The coordinates for A-subunit (apo), core enzyme and holoenzyme are taken from 1B3U.pdb, 2IE3.pdb and 2NPP.pdb, respectively.

### 1.2.1.2 Catalytic C-subunit

Similar to the A-subunit, the catalytic C-subunits of PP2A also exist in two isoforms, Cα and Cβ and share 97% sequence identity. They are coded by two distinct genes *PPP2CA* and *PPP2CB*. Both isoforms are expressed ubiquitously and are most abundant in brain and heart tissues. In every cell, Cα is more highly expressed than Cβ (Khew-Goodall and Hemmings 1988). Although Cα and Cβ are almost identical, Cα knockout mice are embryonic lethal and Cβ cannot compensate for the function of Cα (Götz *et al.* 1998, Gu *et al.* 2012). Cα and Cβ have distinct patterns of subcellular localization. Cα is predominantly present at the plasma membrane, colocalized with membrane-associated β-catenin, during early embryonic development, and when the cell starts to differentiate into primitive endoderm, the cytoplasmic distribution of Cα gradually increases. In contrast, Cβ isoform is localized within cytoplasm and nuclei (Götz *et al.* 2000, Baskaran and Velmurugan 2018).

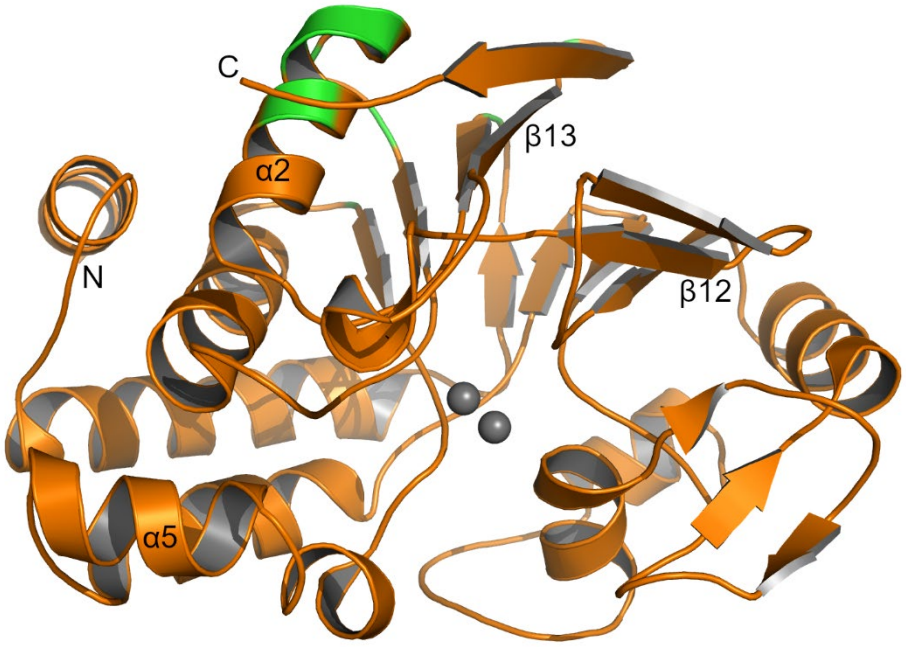


FIGURE 5 Representation of three-dimensional structure of PP2A C-subunit (Protein data bank (PDB) file 2IE3). The manganese ions are depicted as a gray sphere. Position of regions interacting with A-subunit (α2 and C-terminus) and B56 subunit (β12-β13, α5, and C-terminus) are indicated.

The catalytic subunit adopts a globular structure with α/β fold typical of the catalytic subunit of the PPP family of phosphatase and contains two manganese atoms in the active site (Fig. 5). The PP2A core enzyme is held together by 15 intermolecular hydrogen bonds and significant van der Waals contacts between the scaffolding A-subunit and catalytic C-subunit. The interface between A- and C-subunit involves HEAT repeats 11-15 of A-subunit and the region encompassing α2 helix and the carboxyl-terminal fragment of the C-subunit (Xing *et al.* 2006, Xu *et al.* 2006, Cho and Xu 2007). The interface between C- and B56-subunit primarily consist of three discrete areas. In the first area, a

conserved arginine residue (Arg268) in the β12- β13 loop of C-subunit interacts with an acid cluster in the intra-repeat loop 2 of B56. In the second area, intra-repeat loops of B56 HEAT repeats 6–8 interact with residues in the α5 region of C-subunit. In the third area, the carboxyl terminal tail of C-subunit interacts with the residues in B56 HEAT repeat 4-6. The highly conserved C-terminal tail ($^{304}$TPDYEL$^{309}$) of PP2Ac subunit is located at the critical interface between scaffolding PR65 and regulatory B subunit (Xu *et al.* 2006, Cho and Xu 2007). The assembly of B-subunit to the core enzyme is regulated by the methylation and phosphorylation pattern of the C-terminal tail of the PP2Ac. Methylation of L309 is catalyzed by leucine carboxyl methyltransferase (LCMT) and is essential for the binding of B55 subunit but not for other variants like B56 and PR72/B" (Longin *et al.* 2007), whereas phosphatase methylesterase (PME-1) reverses the activity of LCMT adding another facet to the holoenzyme regulation (Ogris *et al.* 1999). Phosphorylation of Tyr307 and unidentified threonine (most probably Thr304) inactivates PP2A (Chen *et al.* 1992, Guo and Damuni 1993). Phosphorylation of Tyr307 and Thr304 inhibits the recruitment of all B55 subunits to the core enzyme. Also, Tyr307 phosphorylation inhibited assembly of B56 α/β/γ/ε subunits to AC core enzymes. However, phosphorylation of Tyr307 had no effect in the holoenzyme assembly with B56δ, PR72 or PR70 subunits (Longin *et al.* 2007). The catalytic subunit (alpha-isoform) of PP2A, PPP2CA, is downregulated in androgen independent prostate cancer cells (Singh *et al.* 2008, Bhardwaj *et al.* 2011) and acute myeloid leukemia (AML) as compared to normal or benign tumor cells (Ramaswamy *et al.* 2015).

#### 1.2.1.3 Regulatory B56-subunit

B56 is one of the family of regulatory subunits and has five different isoforms, α, β, γ, δ, and ε. The members of PP2A B56 family are phosphoproteins and are 80% identical to each other over the central 400-amino acid conserved region but differ in their N- and C-terminals. The expression of B56 isoforms differs intracellularly. B56α, B56β, and B56ε are predominantly expressed in the cytoplasm while B56γ are concentrated in nucleus and B56δ is expressed in both the nucleus and cytoplasm (McCright *et al.* 1996, Cho and Xu 2007). The structure of B56γ contains eighteen alpha helical units that are stacked against each other to form an elongated, superhelical structure, similar to HEAT repeats (Fig. 6). The molecular surface of B56γ contains few protrusions and depressions resulting from the uniform packing of HEAT-like repeats. The hydrophobic convex surface of B56γ associates with A-subunit, whereas the acidic, concave face of the molecule is tilted towards the active site of C-subunit. B56γ makes extensive interactions with the intra-repeat loops of A-subunit HEAT repeats 2-8. The interface between B56γ and A-subunit is divided into two areas, a large area and a small area. The large area of the interface is formed between HEAT repeats 4-5 of B56γ and ridge of HEAT repeat 2-5 of A-subunit, whereas the small area is formed by the extended loop in HEAT repeat 2 of B56γ and the ridge of HEAT repeats 7-8 of A-subunit. In general, the A- and B56-subunit interface is characterized by loose interaction which corroborates with the fact that A- and B-subunit cannot form a stable complex (Xu *et al.* 2006, Cho and Xu 2007).

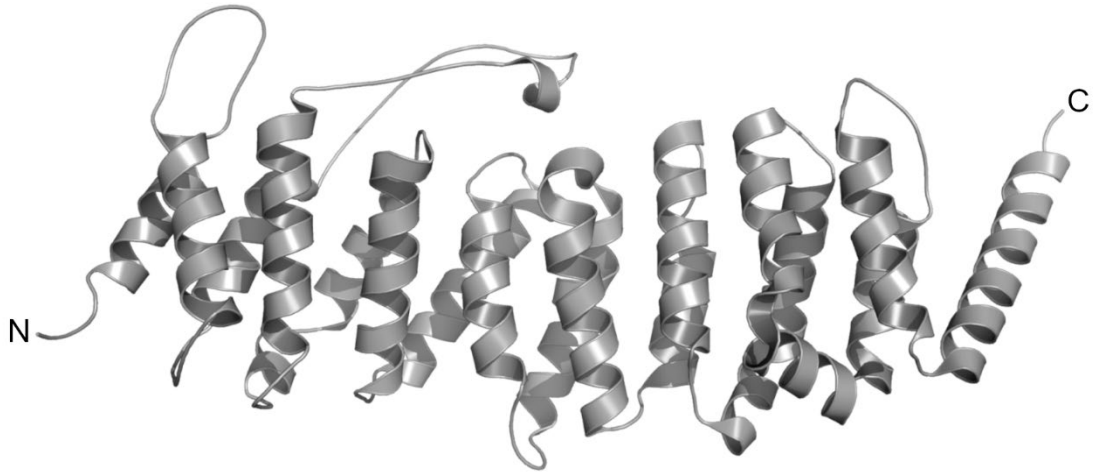


FIGURE 6 The three-dimensional structure of PP2A B56γ (PDB code: 2NPP). The helical units of B56γ are stacked together to form HEAT-like repeat. N- and C-termini are indicated.

It has been reported that the mutations of gene encoding PP2A B56-subunits occur in low frequency as compared to A-subunits in different types of cancer (Table 1). Besides mutation, decreased expression of PP2A B56-subunit and the expression of PP2A inhibitor proteins is more commonly linked to cancer progression (Arroyo and Hahn 2005, Meeusen and Janssens 2017, Kauko and Westermarck 2018). PP2A B56α binds and dephosphorylates Ser62 of oncoprotein c-MYC and promotes its degradation, which eventually induces oncogene induced senescence of the melanoma cells. Whereas, the depletion of B56α in normal melanocytes upregulates c-MYC affecting oncogene induced senescence triggered by members of RAS protein family (Mannava *et al.* 2012). In PP2A B56γ, F395C mutation is found in lung cancer. A cancer derived B56γ mutation, F395C, failed to promote p53 Thr55 dephosphorylation and p21 induction, suggesting that this mutant lost the p53-dependent tumor suppressive role of PP2A (Shouse *et al.* 2010). In addition to mutation, B56γ has been found to be downregulated in primary melanoma (Deichmann *et al.* 2001), whereas it is overexpressed in metastatic melanoma (Ito *et al.* 2000). PP2A-B56γ holoenzyme dephosphorylates serine phosphorylation on paxillin which is required for regulation of focal adhesion assembly. However, the N-terminally truncated isoform of B56γ failed to dephosphorylate paxillin. This results in the increased paxillin phosphorylation which induces paxillin recruitment into focal contacts and tumor cell motility (Ito *et al.* 2000). Loss or down-regulation of B56β and B56γ prolong Ser473 phosphorylation in AKT, contributing to cell transformation (Chen *et al.* 2004, Rocher *et al.* 2007, Sablina *et al.* 2010). Inhibition of tumor suppressor activity of PP2A in AML cells expressing c-KIT mutants (V560G and D816V) is associated with the downregulation of B56α, B56γ and B56δ subunits (Roberts *et al.* 2010). Recent *in vivo* study showed that B56δ knockout mice are susceptible to spontaneous development of hematologic malignancies and hepatocellular carcinoma. It has been reported that the loss of B56δ resulted in activation of oncogenic c-MYC via Ser62 hyperphosphorylation (Lambrecht *et al.*

2018). Likewise, PP2A B56ε is frequently downregulated in AML (Cristóbal *et al.* 2013) and is associated with the destabilization of p53 (Jin *et al.* 2010).

Collectively, the deregulation of PP2A subunits by mutations, downregulation and post-translational modification of subunits in various cancers have highlighted the role of PP2A subunits as a tumor suppressor. In addition to these mechanisms of inactivation, PP2A is frequently inactivated by overexpression of endogenous PP2A inhibitor proteins in cancers (Kauko and Westermarck 2018, Meeusen and Janssens 2018).

TABLE 1 Summary of mutations in PP2A subunits associated with different types of cancer

| PP2A subunits | Mutations | Types of cancer | References |
|---|---|---|---|
| PR65α | E64D | Lung cancer | (Ruediger *et al.* 2001) |
| | E64G | Breast cancer | (Ruediger *et al.* 2001) |
| | R418W | Melanoma | (Ruediger *et al.* 2001) |
| | R182G, R183G, R183W | Ovarian clear cell cancer | (Jones *et al.* 2010) |
| | P179R, S256F, P179L | Serous endometrial cancer | (McConechy *et al.* 2011, Nagendra *et al.* 2012) |
| PR65β | G90D | Lung and Breast cancer | (Wang *et al.* 1998, Esplin *et al.* 2006) |
| | P65S, K343E, D504G | Lung cancer | (Wang *et al.* 1998) |
| | L101P, V448A, V545A | Colon cancer | (Wang *et al.* 1998) |
| B56γ | C39R, E164K | Pooled glandular tumor sample | (Nobumori *et al.* 2012, 2013) |
| | S220N | Uterine leimyosarcoma | (Nobumori *et al.* 2013) |
| | Q256R, L257R | Melanoma | (Nobumori *et al.* 2013) |
| | F395C | Lung cancer | (Shouse *et al.* 2010, Nobumori *et al.* 2013) |
| | A383G | Intestinal cancer | (Shouse *et al.* 2010) |

## 1.3 Endogenous Protein phosphatase 2A inhibitor proteins

The pivotal role of PP2A in cell cycle control necessitates proper regulation of enzyme activity. During normal cell growth, PP2A is activated in some stages of the cell cycle and deactivated in others, emphasizing the presence of regulatory mechanisms for the dynamic control of cellular phosphatase activities. The endogenous inhibitor proteins of PP2A coordinate the regulation of the overall activity of PP2A (Shenolikar 1995). Current evidence shows that PP2A inhibition via non-genomic mechanisms is more frequent in cancer and mainly involve overexpression of PP2A inhibitor proteins (Meeusen and Janssens 2017, Kauko and Westermarck 2018). In recent years, cellular proteins, such as Cancerous Inhibitor of Protein Phosphatase 2A (CIP2A) (Junttila *et al.* 2007), Suvar/Enhancer of zeste/Trithorax (SET) (Mei Li, Anthony Makkinje 1996), Phosphatase MethylEsterase-1 (PME-1) (Ogris *et al.* 1999), Biorientation of chromosomes in cell division 1 (BOD1) (Porter *et al.* 2013), cAMP-Regulated Phosphoprotein-16/19 (ARPP-16/19), and α-Endosulfine (ENSA) (Mochida *et al.* 2010, Gharbi-Ayachi *et al.* 2010, Song *et al.* 2014, Lü *et al.* 2015, Jiang *et al.* 2016, Ye *et al.* 2020) are emerging as a key player in cell proliferation and survival. Among these inhibitors, ARPP-16/ARPP-19, ENSA, and BOD1 are associated with the cell cycle regulation by modulating PP2A, whereas CIP2A, SET, and PME-1 are linked to the proliferative status of cells (Kauko and Westermarck 2018).

Intrigued by the role of ARPP-16/19 and ENSA in the regulation of PP2A complexes and association with human cancer progression and development (Song *et al.* 2014, Lü *et al.* 2015, Jiang *et al.* 2016), we herein characterized the structural properties of ARPP-16/19 and ENSA, and their interaction with different subunits of PP2A.

### 1.3.1 ARPP-16/ARPP-19

cAMP-regulated phosphoprotein-19 (ARPP-19) and its splice variant ARPP-16 are the members of the α-endosulfine family, which are almost identical, with the additional 16 amino acids in the N-terminal of the former. They were first identified as the substrate of cAMP-regulated protein kinase in the neostriatum. ARPP-16 is specifically enriched in basal ganglia, whereas ARPP-19 is present ubiquitously in neuronal and non-neuronal cells (Horiuchi *et al.* 1990, Girault *et al.* 1990, Dulubova *et al.* 2001). The expression of ARPP-19 and ARPP-16 is differentially regulated, as indicated by a high abundance of ARPP-19 in embryonic tissue and its gradual decrease with development, while ARPP-16 starts appearing during the postnatal period (Girault *et al.* 1990). ARPP-16 and ARPP-19 found in different species, including human, porcine, bovine, rodent, and frog, show high sequence similarity to each other. Furthermore, they contain two conserved regions in all species, the first region encompasses Greatwall (Gwl) kinase or Microtubule-associated serine/threonine-protein kinase 3 (MAST3) phosphorylation site and contains a motif of eight residues, KYFDSGDY (Mochida *et al.* 2010, Gharbi-Ayachi *et al.* 2010, Andrade *et al.* 2017),

and the second region includes the protein kinase A (PKA) phosphorylation site (Dulubova *et al.* 2001). Biochemical characterization showed both ARPP-16 and ARPP-19 are highly basic, heat, and acid-stable (Horiuchi *et al.* 1990, Dulubova *et al.* 2001). Upon binding on 3' end of growth-associated protein-43 (*GAP-43*) mRNA, ARPP-19 acts as a transactivator of *GAP-43* mRNA stability that promotes the development and plasticity of the nervous system in a nerve growth factor-dependent manner (Irwin *et al.* 2002). A significant reduction of ARPP-19 expression has been reported in the brain cells of patients with neuronal system diseases, such as Down Syndrome and Alzheimer's disease (Kim *et al.* 2001).

Furthermore, ARPP-16 and ARPP-19 play important roles in cell cycle regulation. The activation of Gwl kinase phosphorylates ARPP-19, which in turn binds and inhibits PP2A, promoting mitotic entry and maintaining mitotic state (Mochida *et al.* 2010, Gharbi-Ayachi *et al.* 2010). The phosphorylation of ARPP-16 at Ser46 by MAST3 kinase, a homolog of MASTL/Gwl kinase, inhibits PP2A in the mammalian brain (Andrade *et al.* 2017). Phosphorylation of ARPPs by MAST3 kinase interacts and selectively inhibits PP2A holoenzyme containing B55α and B56δ (Andrade *et al.* 2017). A recent study revealed that the phosphorylation at Ser46 (Ser62) and Ser88 (Ser104) of ARPP-16 (ARPP-19) by MAST3 and PKA, respectively, have mutual antagonist relation. The prior phosphorylation of ARPP-16 by PKA slowed down the phosphorylation of ARPP16 by MAST3 kinase and vice versa (Andrade *et al.* 2017, Musante *et al.* 2017). In addition, PKA can phosphorylate MAST3 directly and inhibit its activity (Musante *et al.* 2017). Several studies have reported the involvement of ARPP19 in tumorigenesis and progression. ARPP-19 is overexpressed in hepatocellular carcinoma (HCC) (Song *et al.* 2014), human glioma (Jiang *et al.* 2016), and AML (Mäkelä *et al.* 2019). Besides, overexpression of ARPP-19 mediates tamoxifen-resistance in breast cancer and papillary thyroid carcinoma (PTC). The depletion of ARPP-19 via the up-regulation of miR-320a and miR-26a restores the tamoxifen sensitivity in tamoxifen resistance breast cancer and PTC, respectively (Lü *et al.* 2015, Gong *et al.* 2018). Likewise, overexpression of ARPP-19 inhibits the tumor-suppressive role of miR-802 in human laryngeal cancer (Ye *et al.* 2020). Another recent study has shown that the upregulation of ARPP-19 promoted the Herceptin resistance in human epidermal growth factor receptor 2 (HER2)-positive gastric cancer (Gao *et al.* 2020). Collectively, cancers related to elevated expression of ARPP-19 are aggressive with poor relapse-free survival and overall survival rates (Jiang *et al.* 2016, Mäkelä *et al.* 2019, Ye *et al.* 2020, Gao *et al.* 2020). Therefore, it is important to expand our understanding of the underlying molecular mechanism of ARPP-19 in cancer progression.

### 1.3.2 α- Endosulfine (ENSA)

Human ENSA is closely related to cAMP-regulated phosphoprotein -19 (ARPP-19), but is encoded by different genes (Peyrollier *et al.* 1996). It was initially identified as an endogenous ligand of sulfonylurea receptor associated with ATP-dependent potassium (K-ATP) channel in pancreatic beta cells regulating

insulin secretion (Virsolvy-vergine *et al.* 1992, Heron *et al.* 1998, Bataille *et al.* 1999). ENSA is expressed in wide array of tissues, expressing K-ATP channels (Inagaki *et al.* 1995, Sakura *et al.* 1995), with the highest levels occurring in the lung and brain and lower levels in the pancreas (Heron *et al.* 1998). It has been reported that human ENSA blocks K-ATP channel and facilitate the release of neurotransmitter, acetylcholine, from the striatal cholinergic interneuron (Kim and Lubec 2001a, b) and insulin from pancreatic beta-cells (Heron *et al.* 1998). ENSA, like other antidiabetic sulfonylureas, can bind and close the K-ATP channels. This induces membrane depolarization and subsequent generation of an action potential that leads to the voltage-gated $Ca^{2+}$ influx. This wave of $Ca^{2+}$ triggers insulin secretion (Heron *et al.* 1998, Bataille *et al.* 1999). The secretion of neurotransmitters from striatal cholinergic interneuron is similar to the insulin secretion from the pancreatic β-cells (Amoroso *et al.* 1989, Lee *et al.* 1997, Kim and Lubec 2001a, b). The decreased expression of ENSA is associated with the neurological conditions like Down syndrome (Kim and Lubec 2001a), Alzheimer's disease (Kim and Lubec 2001b), and synucleinopathies (Ysselstein *et al.* 2017).

It has also been reported that the phosphorylation of ENSA at Ser67 by Gwl kinase inhibits PP2A-B55 and allows mitotic entry in *Xenopus* egg extracts (Mochida *et al.* 2010). At the same time, the ablation of ARPP-19 from *Xenopus* egg extract, but not of ENSA, prevented mitotic progression (Gharbi-Ayachi *et al.* 2010). In addition to the mitotic progression, the Gwl-ENSA-PP2A pathway is important for S phase progression. The depletion of ENSA from the human cell promotes extension of the S phase by decreasing the number of replication forks. This phenotype is the result of PP2A-B55 dependent dephosphorylation of Cdk phosphorylated Treslin, a protein responsible for the firing of replication origins (Charrasse *et al.* 2017). This suggests that PP2A inhibitor ENSA regulates cell cycle progression differently from ARPP-19 (Hached *et al.* 2019).

## 1.4 Role of ARPPs and ENSA in the regulation of $G_2$/M transition

Mitotic entry and progression are regulated by the balanced action of cyclin B-Cdk1 kinase and PP2A phosphatase (Fig. 7) (Lorca and Castro 2013). Normal mitotic entry requires the phosphorylation of mitotic substrate by protein kinase cyclin B-Cdk1 (Schwartz and Shah 2005, Song *et al.* 2014). During interphase, Wee1 and Myt1 kinases act together to negatively regulate newly formed cyclin B-Cdk1 by phosphorylating Thr14 and Tyr15 (Gould and Nurse 1989, O'Farrell 2001, Kishimoto 2015, Vigneron *et al.* 2016). During G2, the level of Cdc25 increases and reverses the Cdk1 phosphorylation by Wee1/Myt1 (Moreno *et al.* 1990, García-Blanco *et al.* 2019).

At G2/M transition, the balance between Wee1/Myt1 and Cdc25 shifts towards the removal of inhibitory phosphorylation in Cdk1 by Cdc25. This results in the activation of a small population of cyclin B-Cdk1, thereby, activating the Cdk1 autoactivation loop. It has also been reported that the cyclin B-Cdk1 is initially activated by a putative initial activator, however, which molecules are the initial activator of eukaryotic cyclin B-Cdk1 remains unclear (Okumura *et al.* 2002, Kishimoto 2015). On the other hand, the initial activator of cyclin B-Cdk1 complex in starfish oocyte is well characterized and found to be AKT/Protein kinase B (PKB) (Okumura *et al.* 2002, Kishimoto 2015). The Cdk1 auto-activation loop activates a larger population of cyclin B-Cdk1 by phosphorylating Wee1/Myt1 and Cdc25, which leads to the inactivation of Wee1/Myt1 and activation of Cdc25 (Solomon *et al.* 1990, O'Farrell 2001, Kishimoto 2015, Mochida *et al.* 2016). PP2A antagonizes the cyclin B-Cdk1 catalyzed phosphorylation of Wee1/Myt1 and Cdc25 (Vigneron *et al.* 2009, Lorca and Castro 2013). Therefore, the activity of PP2A should be inhibited to achieve the full activation of the cyclin B-Cdk1. The Gwl kinase, one of the members of cyclin B-Cdk1 auto-activation loop, maintains mitosis by promoting the inactivation of PP2A via the Gwl-endosulfine pathway (Vigneron *et al.* 2009, Mochida *et al.* 2010, Gharbi-Ayachi *et al.* 2010).

A small starter amount active cyclin B-Cdk21 can activate Gwl. The activation of Gwl kinase phosphorylates Ser62/Ser67 on ARPP-19/ENSA, which in turn binds and inhibits PP2A, promoting mitotic entry and regulating proliferation (Mochida *et al.* 2010, Gharbi-Ayachi *et al.* 2010). Also, ARPP-19 is phosphorylated at Ser23 in human ARPP-19 by cyclin B-Cdk1, a different conserved site than the Gwl phosphorylation site (Okumura *et al.* 2014). ARPP-19 phosphorylated by cyclin B-Cdk1 can suppress the activity of PP2A partially which could be sufficient for the activation of cyclin B-Cdk1 auto-activation loop (Okumura *et al.* 2014). This explains the importance of ARPP-19 in the initiation and maintenance of mitosis.

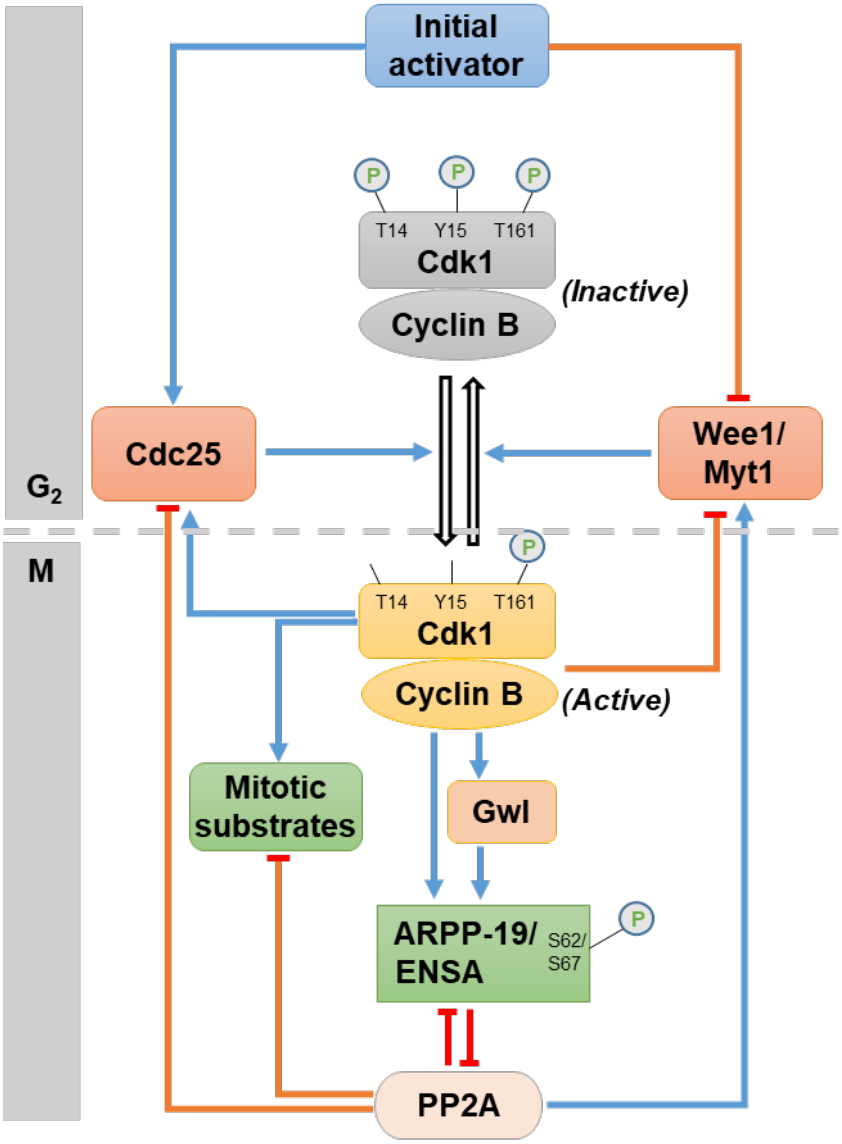


FIGURE 7 Illustration of activation of Cyclin B-Cdc2 and role of ARPP19 during mitotic entry At the G2/M phase transition, an initial activator alters the balance between Wee1/Myt1 and Cdc25 to stimulate the activation of a small number of cyclin B-Cdk1 (initial activation). The initial activation of cyclin B-cdk1 then leads to the activation of the autoactivation loop to generate large amounts of active cyclin B-Cdk1. Gwl is activated The threshold amount of activated cyclin B-cdk1 activates Gwl that, in turn, phosphorylates ARPP-19/ENSA. The phosphorylated ARPP-19/ENSA binds and inhibits PP2A.

# 2 AIMS OF THE WORK

PP2A inhibition is important for human cell transformation and cancer progression. Therefore, it is essential to understand the PP2A-inhibitory mechanism to develop new and better therapeutics for cancer. Though the potential PP2A inhibitor protein ARPP-19/ENSA has been identified, details about their structure and ARPPs/ENSA interactions are unknown. To shed light on these questions, we set the following aims in our study:

- To carry out structural characterization of ARPP-16, ARPP-19 and ENSA,
- To understand how ARPPs and ENSA interact with PP2A A- and different B56-subunits, and
- To identify the region in ARPPs and ENSA involved in binding with PP2A A- and B56-subunits.

# 3 SUMMARY OF THE METHODS

| Method | Publication |
|---|---|
| Ligation independent cloning | I, II, III |
| Restriction digestion and ligation dependent cloning | I,II |
| Site-directed mutagenesis | II, III |
| Bacterial Protein expression | I, II, III |
| Protein Purification | I, II, III |
| MicroScale thermophoresis (MST) based interaction study | II, III |
| SAXS data collection and analysis | II, III |
| NMR data collection | I, II, III |
| NMR data analysis; assignment of chemical shifts | I, II, III |
| NMR data analysis; spin relaxation rates and RDSM | II, III |
| Calculation of structural ensemble based on NMR parameters | II, III |
| NMR titration experiments | II, II |

# 4 RESULTS

## 4.1 cAMP regulated phosphoproteins (ARPP-16/19 and ENSA) are intrinsically disordered

The structural characteristics of cAMP-regulated phosphoproteins (ARPP-19, ARPP-16, and ENSA) were investigated using bioinformatics tools like PONDR (Xue *et al.* 2010) and IUPred2A (Mészáros *et al.* 2018, Erdős and Dosztányi 2020). Both of these tools predicted that ARPPs and ENSA are disordered (I; Fig. 1B, II; Fig. 2A, III; Fig. 1B and E), whereas N-terminal end of ENSA was more ordered than other parts. IUPred2A also predicted the presence of two disordered binding regions in ARPP-16 and three in ARPP-19 (II; Fig. 2A), whereas ENSA has a single extended disordered binding region (III; Fig. 1E). Size exclusion chromatography (SEC) separates proteins according to their hydrodynamic radii rather than their molecular masses. ARPP-16, ARPP-19, and ENSA migrate with the volume characteristics of globular protein with a molecular mass of 16 kDa, 20 kDa, and 18 kDa, respectively (Fig. 8C). Since the molecular weight of ARPP-16, ARPP-19, and ENSA are only 10.7 kDa, 12.6 kDa, and 13.4 kDa, respectively, their anomalous migration in SEC suggests that ARPPs and ENSA have more extended conformation than their globular counterparts. Another feature that ARPP proteins and ENSA share with intrinsically disordered proteins (IDPs) are an unusual migration on SDS-PAGE (Fig. 8D).

Further structural characterization was carried out using NMR spectroscopy, which offers unprecedented opportunities for the structural and dynamic studies of IDPs. The $^{15}$N-HSQC spectrum of ARPPs and ENSA displayed a poorly dispersed $^{15}$N, $^{1}$H correlations, extensive signal overlap, and variable peak intensities, a characteristic feature of disordered proteins (Fig. 9, I; Fig. 1C and D, II; Fig. 2B, III, Fig. 1C). Nearly complete backbone resonance assignment was obtained for all three using a set of HN-detected and HA-detected triple resonance experiments (I; Fig. 2 and 3). The residual secondary structure was characterized by $^{15}$N relaxation data and secondary structure propensity (SSP) calculation. The secondary structure propensities were estimated for ARPPs and ENSA with the

SSP software using a chemical shift of $^{13}C\alpha$, $^{13}C\beta$, and $^{1}H\alpha$. Three transiently structured helices were identified in all three proteins from secondary chemical shifts and it was supported by $^{15}N$ relaxation data (II; Fig. 2A, III; Fig. 1E). Next, we adopted a reduced spectral density mapping approach to interpreting the $^{15}N$ relaxation data. The dynamics measured at *J(0), J($\omega_N$), and J(0.87$\omega_H$)* frequencies indicate increased rigidity for residues $^{9}$EKAEEAKLKARY$^{20}$, $^{32}$FLRKRLQK$^{39}$ and $^{41}$QKYFDSGDYNMAKAKMK$^{57}$ of ARPP-16 and residues $^{25}$EKAEEAKLKARY$^{36}$, $^{48}$FL$^{49}$ and $^{57}$QKYFDSGDYNMAKAKMKNK$^{75}$ of ARPP-19, that correspond to the transient α-helical regions obtained from SSP calculation (II; Fig. 3). Similarly, reduced spectral density mapping of ENSA suggests restricted backbone motion of the regions containing residues $^{31}$RAEEAKLKAK$^{40}$ and $^{53}$FLMKRLQKGQKYFDSGDYNMAKAKMKN$^{79}$ (III; Fig. 2).

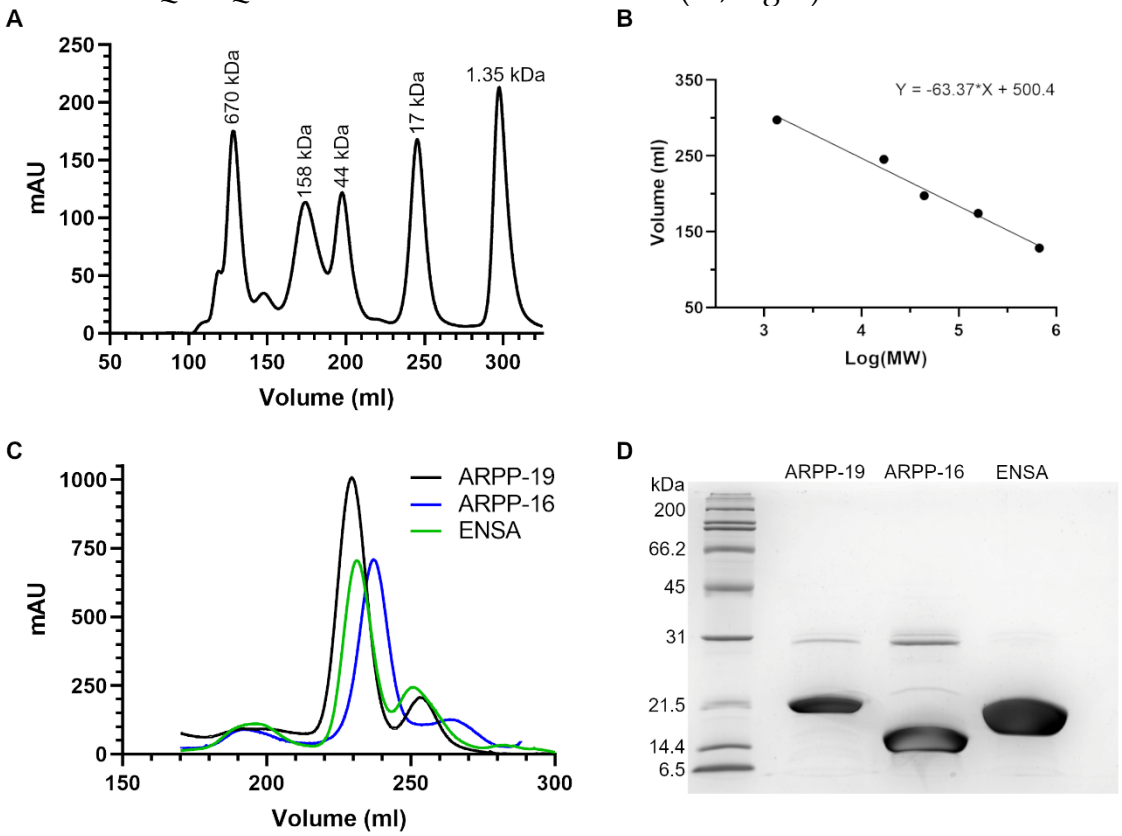


FIGURE 8 Characterization of the ARPP-16, ARPP-19 and ENSA. **A-** SEC profile of Gel filtration standard (Bio-Rad laboratories, Inc.) using Superdex 200 16/60 column. **B-** A calibration curve of protein standard generated by plotting the log of molecular weight of the standard proteins against the retention volumes (ml). The gel filtration standard include bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa) and Vitamin $B_{12}$ (1.35 kDa). The calibration curve is used to estimate the molecular weight of a protein based on the elution volume. **C-** SEC profile of ARPP-19 (black), ARPP-16 (blue) and ENSA (red). The ARPP-19, ARPP-16 and ENSA eluted from the SEC column with the retention volume of 228 ml, 234.1 ml, and 231.4 ml, respectively. **D-** SDS-PAGE gel image of purified ARPP-16, ARPP-19 and ENSA shows their anomalous migration. 5 µg purified proteins was loaded on 12% SDS-PAGE gel.

Additional information on the structural properties of ARPPs and ENSA was obtained from small angle X-ray scattering (SAXS) experiments. For all the proteins, we observed a large radius of gyration ($R_g$) and large $D_{max}$ with extended distance distribution p(r) functions (II; Fig. S7, III; Fig. S2). The Kratky representation of these data showed a monotonic rise in the curve as expected for an intrinsically disordered protein (II; Fig. 2C, III; Fig. 1D). Furthermore, the broad distribution of $R_g$ and $D_{max}$ values obtained from SAXS data analysis using ensemble optimization method (EOM) software also supported the flexible nature of the protein (II; Fig. 4C, III; Fig. 3).

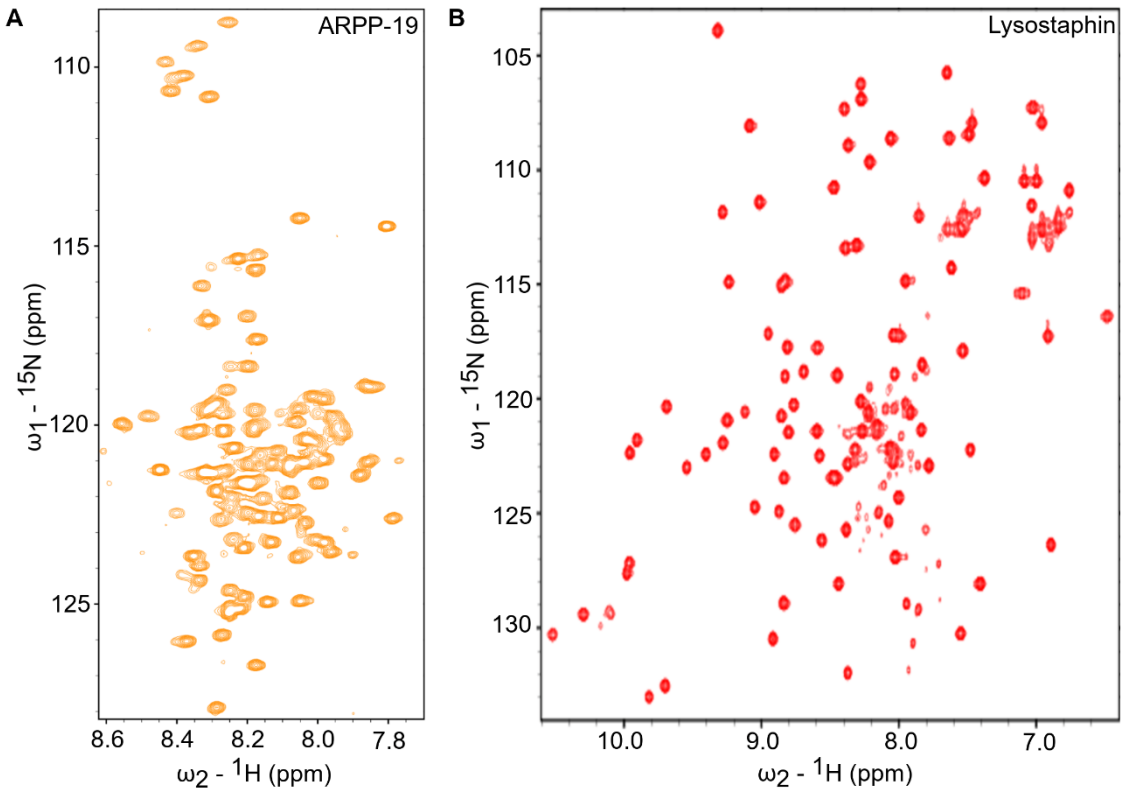


Figure 9 $^{1}H$-$^{15}N$ HSQC spectra of **A**- ARPP-19 and **B**- Lystostaphin, illustrating the decrease in resonance dispersion in the $^{1}H$ dimension of intrisically disordered protein, ARPP-19, in comparison to the ordered protein, Lysostaphin.

As it is not possible to characterize IDPs by single representative structure, an ensemble of structure that best fits the experimental data is frequently used to characterize them. Among different experimental techniques, NMR spectroscopy has become the technique of choice to access atomic resolution information of IDPs under conditions that are close to the physiological condition. We calculated ensemble of structures of ARPP proteins and ENSA using ENSEMBLE software suite. The ENSEMBLE program combines SAXS experimental data and different NMR observables like chemical shifts, residual dipolar coupling (RDC), paramagnetic relaxation enhancement (PRE), nuclear overhauser effect (NOE), and $^{15}N$ $R_2$ relaxation rate, into the same refinement protocol to derive the ensemble of IDPs. The ENSEMBLE program uses the following three steps to

determine the structural ensemble of IDPs: (i) generation of a large set of initial conformational pool (ii) back-calculation of different parameters, such as NMR chemical shifts and SAXS scattering, from individual conformation (iii) selection of the ensemble that best fits the experimental data with the one back-calculated. In the calculation, NMR observables, such as chemical shifts, $^{15}N$ $R_2$ relaxation data, and SAXS data were used (II: Fig 4A, III; Fig. 3).

## 4.2 ARPP proteins and ENSA interact to PP2A A-subunit with intermediate affinity

Although many studies have identified ARPP-16, ARPP-19 and, ENSA as important PP2A inhibitors during a cell cycle, it is still unclear how they interact with PP2A. It has been reported that ARPP proteins interact directly with PP2A A-subunit (Andrade *et al.* 2017). To understand the PP2A-ARPPs and, PP2A-ENSA interaction, we initially measured binding affinities of WT ARPPs, WT ENSA and their phosphomimicking mutants to PP2A A-subunit (PR65α) using microscale thermophoresis (MST). The binding of ARPPs to PP2A A-subunit was further characterized by NMR spectroscopy to determine PP2A A-subunit binding region in ARPPs.

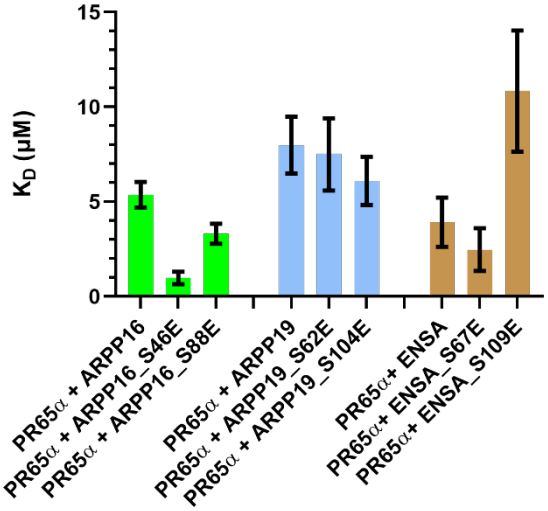


FIGURE 10 Histogram representation of equilibrium dissociation constant for the binding of ARPP-16, ARPP19 and ENSA with PP2A A-subunit obtained from MST. The binding affinity measurement were performed using fluorescently labelled A-subunit, that served as target, and unlabelled ARPP-16, ARPP-19 and ENSA, that served as ligand. Each bar represents the mean $K_D$ ± standard error of mean of 3 independent measurement. All $K_D$ values can be found in (Table 2). In the figure, smaller bars indicates higher affinity and bigger bars indicates lower affinity.

In our experimental setting, the $K_D$ of ARPP-16 wild type to PR65α was 5.4 ± 0.7 μM and that of ARPP-19 wild type and PR65α was 7.9 ± 1.5 μM. The binding affinity of ENSA wild type to PR65α was similar to that of ARPPs with $K_D$ 3.9 ± 1.3 μM. As phosphorylation of ARPP proteins and ENSA by Gwl or MAST3 kinase is essential for the inhibition of PP2A (Gharbi-Ayachi *et al.* 2010, Dupre *et al.* 2013, Juanes *et al.* 2013), we performed the interaction between PP2A A-subunit and the phosphomimicking mutants of ARPP proteins and ENSA, corresponding to GWL or MAST3 kinase phosphorylation and PKA phosphorylation. Interestingly, the binding affinity of phophomimicking mutants of ARPP proteins and ENSA to PP2A A-subunit was similar to that of their respective wild types (Fig. 10, Table 2). This result suggests that the binding of ARPPs and ENSA to PP2A A-subunit could be independent of the phosphorylation.

## 4.3 ARPPs and ENSA shows weak affinity towards PP2A B56-subunits compared to PP2A A-subunit

The PP2A holoenzyme complex, containing regulatory B56-subunit, functions as a tumor suppressor (Chen *et al.* 2004, Li *et al.* 2007, Nobumori *et al.* 2013). The Gwl kinase phosphorylated ARPP-16 can inhibit the tumor suppressor activity of PP2A holoenzyme containing B56-subunit (Andrade *et al.* 2017). Therefore, we attempted to quantify how strongly ARPPs and ENSA interact with different B56 isoforms. The evaluation of MST signals shows that the affinity of ARPPs and ENSA for B56-subunits is very low as compared to A-subunit. Although all three wild types interacted with B56-isoforms used in our study, the strength of interaction is not similar. We observed that the affinity of ARPP-19 and ENSA is stronger for B56α and B56δ as compared to other isoforms, while ARPP-16 binds strongest to B56α (Fig. 11A and 11C). As mentioned earlier, Gwl/MAST3 kinase phosphorylation of ARPPs and ENSA is crucial for the functional inactivation PP2A. Therefore, we performed interaction studies between different PP2A B56-isoforms and phosphomimicking mutants of ARPP proteins and ENSA. The phosphomimicking mutants of ARPP-16 and ARPP-19, corresponding to Gwl/MAST3 kinase phosphorylation, showed increased binding affinity towards all the B56 isoforms, except for ARPP-16 S46E binding to B56α. We observed the similar result for the phosphomimicking mutants corresponding to the PKA phosphorylation except, ARPP16 S88E failed to bind with B56ε (Fig. 11D).

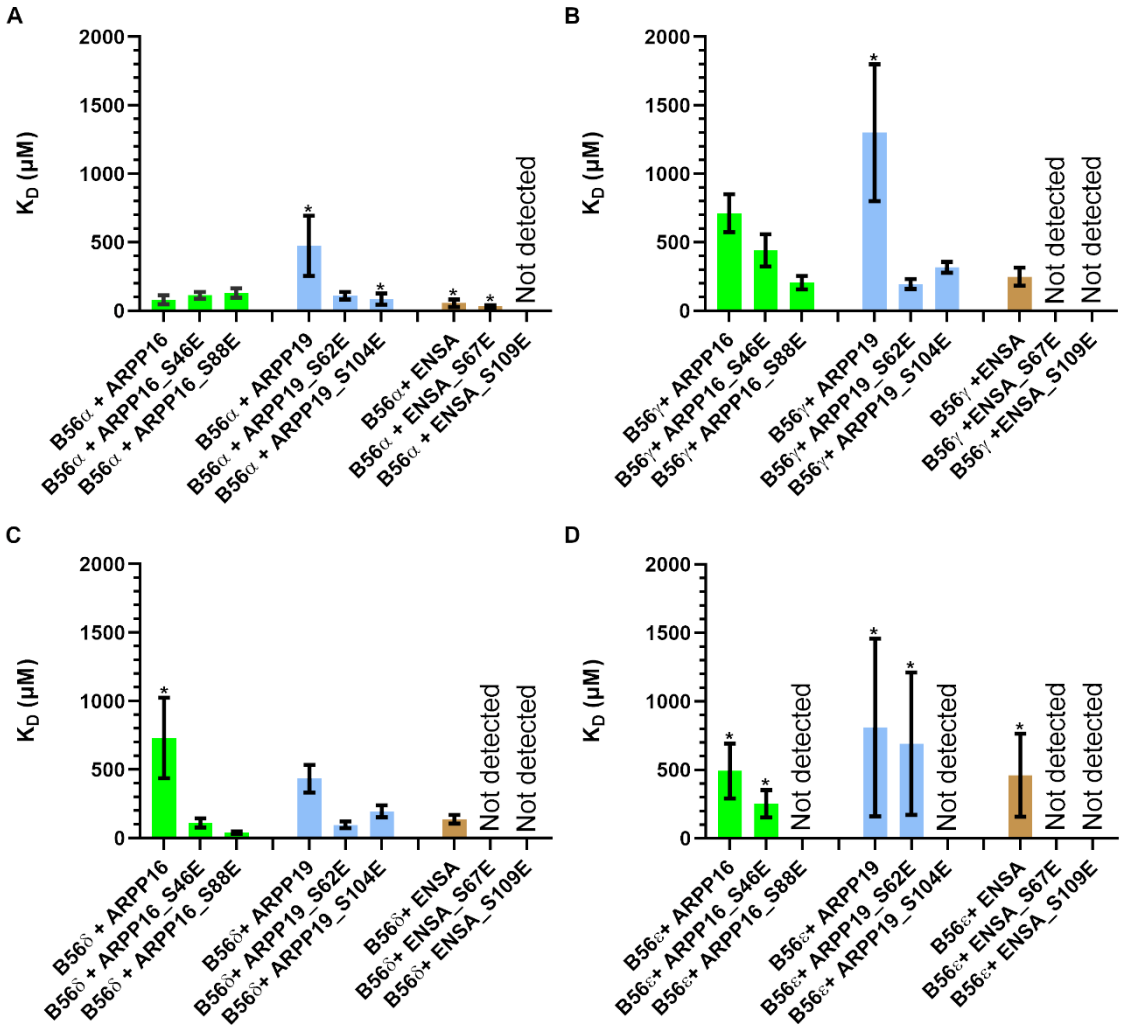


FIGURE 11 Histogram representation of equilibrium dissociation constant for the binding of ARPP-16, ARPP19 and ENSA with different PP2A B56-subunits obtained from MST. The binding affinity measurement were performed using fluorescently labelled B56-subunits that served as target, and unlabelled ARPP-16, ARPP-19 and ENSA, that served as ligand. Each bar represents the mean $K_D$ ± standard error of mean of 3 independent measurement. All $K_D$ values can be found in (Table 2). In the figure, smaller bars indicates higher affinity and bigger bars indicates lower affinity. The sign '*' above the error bar indicates non-reliable error estimate.

Interestingly, wild type ENSA binds to B56-subunits with higher affinity compared to the wild type ARPP proteins, except for the B56ε. The interaction of wild type ENSA with B56ε is weak and transient, which is similar to what we observed for wild type ARPPs. Also, the binding affinity of ENSA S67E to B56α isoform was higher than that of ENSA WT, however, it did not bind to other B56 isoforms. Additionally, all the ENSA phosphomimicking mutants, corresponding to PKA phosphorylation, failed to bind with any of the B56 isoforms (Fig. 11, Table 2). To summarize, the specificity of ARPP proteins and ENSA is different to the different B56 isoforms.

TABLE 2 Binding affinity of ARPP-16, ARPP-19, and ENSA with PP2A A- and different B56 subunit.

| PP2A subunit | ARPP-16 | $K_D$ ± SEM (μM) | ARPP-19 | $K_D$ ± SEM (μM) | ENSA | $K_D$ ± SEM (μM) |
|---|---|---|---|---|---|---|
| A-subunit | WT | 5.4 ± 0.7 | WT | 7.9 ± 1.5 | WT | 3.9 ± 1.3 |
| | S46E | 1.0 ± 0.3 | S62E | 7.5 ± 1.9 | S67E | 2.5 ± 1.1 |
| | S88E | 3.3 ± 0.5 | S104E | 6.1 ± 1.3 | S109E | 10.8 ± 3.2 |
| B56α | WT | 80 ± 32 | WT | > 470 | WT | 55.1 ± 27.4 |
| | S46E | 110 ± 26 | S62E | 110 ± 27 | S67E | 32 .1 ± 8.1 |
| | S88E | 130 ± 34 | S104E | 85 ± 40 | S109E | N.D. |
| B56γ | WT | 710 ± 138 | WT | >1300 | WT | 250 ± 65 |
| | S46E | 440 ± 120 | S62E | 195 ± 35 | S67E | N.D. |
| | S88E | 200 ± 50 | S104E | 320 ± 40 | S109E | N.D. |
| B56δ | WT | >700 | WT | 430 ± 100 | WT | 140 ±30 |
| | S46E | 110 ± 33 | S62E | 95 ± 25 | S67E | N.D. |
| | S88E | 40.7 ± 9.2 | S104E | 195 ± 40 | S109E | N.D. |
| B56ε | WT | >500 | WT | >800 | WT | >450 |
| | S46E | 250 ± 100 | S62E | >700 | S67E | N.D. |
| | S88E | N.D. | S104E | N.D. | S109E | N.D. |

## 4.4 ARPPs and ENSA binds differently to PP2A A-subunit

To map the binding site of PP2A A-subunit on the structure of ARPP-16/19 and ENSA, a NMR titration experiment was performed by adding the PP2A A-subunit to the $^{15}$N-labeled ARPP-16/19 and ENSA. The principle of a NMR titration is to observe binding induced chemical shift perturbation (CSP) in the spectrum of the protein upon gradual addition of binding partner. The chemical exchange rate, i.e., the exchange between free and bound states, influences CSP on the NMR spectra during the protein-protein interaction. If the binding is weaker, the exchange rate between free and bound states is large compared to their chemical shift difference. This exchange regime is known as a fast exchange and gives a single cross peak reflecting the population-weighted average chemical shift of the free and bound form. The cross peak will move smoothly from its position in the free state to the bound state upon titration with the binding partner. On the other hand, strong binding gives rise to slow exchange in which the exchange rate between bound and free states is significantly smaller than the difference in their chemical shift. Therefore, two cross-peaks are observed corresponding to the chemical shifts of free and bound states. As the binding partner is added, the signal of free protein slowly disappears, and that of bound form appears. In the intermediate regime, the exchange rate is comparable in magnitude to their chemical shift difference. It gives rise to a single NMR peak at the chemical shift between a free and bound form with the increased linewidth due to exchange broadening (line broadening). A line

broadening may cause the protein resonance peak to become weaker and finally disappear.

At first, the addition of equimolar concentration of PP2A A-subunit led to a line broadening of several ARPP-19 residues, namely F48, L49, Q54, K55, G56, K58, and F60. Upon the addition of PP2A A-subunit at concentration higher than equimolar concentration, NMR signals from L39 disappeared and chemical shift changes were observed for ARPP-19 residues between L39 and S62. This chemical shift perturbations map the A-subunit binding region in ARPP-19 to be a linear motif of residues L39-S62 (II; Fig. 8C and 8D). This is consistent with the general protein-protein binding mode of IDPs (Perkins *et al.* 2010, Mollica *et al.* 2016). Most of the residues did not experience chemical shift perturbation, indicating that PP2A A-subunit binding does not change the overall conformation of ARPP-19. We observed similar phenomenon when ARPP-16 was titrated with the PP2A A-subunit. The ARPP-16 binds to PP2A A-subunit with the linear motif of 24 residues (L23-S46), which is corresponding to the PP2A A-subunit binding region at ARPP-19 (II; Fig. 8A and 8B).

Next, the interaction of ENSA with the PP2A A-subunit was mapped using NMR spectroscopy. The following residues in $^{15}$N-labelled ENSA manifest significant changes in $^{1}$H, $^{15}$N chemical shifts upon addition of equimolar concentration of PP2A A-subunit: R31, K36, L37, A39, K40, Y41, Q59, Y64, D66, G68, and M72. The severe line broadening was observed for A35, F53, L54, M55, Q62, F65, S67, N71 and A75 residues. Upon addition of PP2A A-subunit at concentrations twice than the ENSA concentration, chemical shift changes were observed for S43, S51, G61, G68, D69 and K76. The close inspection of the CSP map revealed that the ENSA binds to PP2A A-subunit via two linear motifs, (i) residues A30-G45 (binding site B1) and S51-K76 (binding site B2) (III; Fig. 5). For this reason, we consider that the ARPPs and ENSA use different mode of binding to PP2A A-subunit.

# 5 DISCUSSION

The amino acid composition of a protein has a significant influence on the secondary and tertiary structure of the protein. The sequence analysis of ARPP proteins and ENSA revealed a high abundance of disorder promoting amino acids like Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro and a low abundance of order promoting amino acids like Ile, Leu, Val, Cys, Asn, Trp, Tyr and Phe, characteristic feature of intrinsically disordered proteins (III; Fig 1A) (Uversky *et al.* 2000, Romero *et al.* 2001, Dunker *et al.* 2001). The data obtained from SDS-PAGE, size exclusion chromatography, small angle X-ray scattering, and NMR spectroscopy also suggest ARPPs and ENSA are IDPs.

We observed the aberrant elution profile of ARPPs and ENSA in size exclusion chromatography (Fig. 8C). In size exclusion chromatography (SEC), the elution volume ($V_e$) of a protein can be used to estimate the apparent molecular weight of the unknown protein because the $V_e$ has a linear relationship with the logarithm of the molecular weight of proteins. In addition to the molecular weight, the migration of protein in the SEC column is influenced by its shape or hydrodynamic radius. Therefore, IDPs have significantly higher $V_e$ than that of globular protein of the same molecular weight. A general rule of thumb regarding the migration of IDPs in the SEC column is that molten globule-type IDPs and pre-molten globule or random coil-type IDP elute with an apparent molecular weight that is 2-3 and 4-6 times larger than the actual molecular weight, respectively (Uversky 2002, Csizmók *et al.* 2006, Tompa 2010).

SDS-PAGE is one of the most frequently used technique to visualize the proteins and provide accurate information on protein's molecular weight. The apparent molecular weight of the ARPP-16, ARPP-19, and ENSA as observed in SDS-PAGE were significantly higher than their theoretical molecular weight (Fig. 8D). IDPs are well-known for their unusual SDS-PAGE mobility because they show to have larger apparent molecular weight than their theoretical molecular weight. This emanates from the unusual amino acid composition of IDPs, which reduces the binding of SDS to the protein and migrates slowly through the gel than the globular protein of comparable molecular mass (Tompa 2010).

A comprehensive characterization of the structural and dynamical behavior of IDPs is a prerequisite for understanding the role of unfolded state structure and their conformational dynamics in the cellular functions. However, it has been a challenge to characterize IDPs because of their conformational polydispersity and very fast dynamics. Here we combine the two most powerful biophysical methods, NMR and SAXS, for the detailed structural characterization of ARPPs and ENSA.

We obtained smooth and featureless scattering curves for ARPPs and ENSA that are typical scattering curves of IDPs (II; Fig. S2 and S3, III; Fig. S1) (Bernadó 2010, Receveur-Brechot and Durand 2012). Often Kratky plot is used as an approach to distinguish between ordered and disordered proteins. Instead of exhibiting a bell-shaped curve of globular proteins, the Kratky plot of ARPPs and ENSA rise to a plateau reminiscent of the flexible polypeptide chain (II; Fig. 2C, III; Fig 1D) (Bernadó and Svergun 2012, Receveur-Brechot and Durand 2012). Also, the distance distribution function, *p(r)*, of ARPPs and ENSA displayed asymmetric *p(r)* function with the extended tail, a characteristic feature of IDPs, compared to the highly symmetric bell-shaped *p(r)* function of globular proteins (II; Fig. S7, III; Fig. S2) (Receveur-Brechot and Durand 2012). The EOM provides an explicit description of the size distribution and structural ensemble of IDPs (Tria *et al.* 2015). The EOM-optimized conformers of wild type ARPPs and ENSA exhibit a bimodal $R_g$ distribution comprising both more compact and flexible structure (II; Fig. 4C, III; Fig. 3). However, the EOM-derived conformers of ARPPs' phosphomimetic mutants displayed less variability and favored more compact conformation. Despite of more compact conformers of phosphomimetic mutants, they are still flexible with the broad $R_g$ distribution comparable to that of the pool (II; Fig. 5C and 5D).

NMR is the most powerful and suitable technique for studying structural propensities and dynamics of IDPs. IDPs are devoid of the propensity to form stable secondary and tertiary structure, yet it is primed to form transient secondary structures (Wright and Dyson 1999, Dunker *et al.* 2001, Dyson and Wright 2002a, Fuxreiter *et al.* 2004). The chemical shift is the primary observable in NMR studies, and given its sensitivity to the local chemical environment of the nucleus, it is the sensitive reporter of the secondary structure content. The structural propensities within a protein can be extracted from the secondary chemical shifts, that is, a characteristic deviation of chemical shifts of certain nuclei from their expected random coil values (Wishart and Sykes 1994, Schwarzinger *et al.* 2001). In the case of α-helical propensities, the chemical shift of $^{13}C\alpha$ and $^{13}CO$ are shifted downfield, whereas β-strand or extended propensities are observed upfield from the corresponding random coil shifts. Similar but opposite behavior will be observed for $^{13}C\beta$, $^{15}N$, $^{1}H\alpha$, and $^{1}HN$ chemical shifts, i.e., chemical shifts are shifted upfield in the case of α-helical propensity, whereas β-structure propensities are shifted downfield from their respective random coil shifts (Dyson and Wright 2002a, 2004, Delaforge *et al.* 2017). Both ARPPs and ENSA displayed similar structural propensities and identified three regions with the alpha-helical propensities (II; Fig. 2A, III; Fig. 1E). SSP score with positive value indicates α-helical propensities, whereas, a

negative value indicates β-strand propensities, where residual SSP score of +1 and -1 reflects fully formed α- and β-structure respectively, while any positive (negative) value in between -1 to +1 provides the quantitative measure of the percentages of the conformer in the disordered state ensemble with α-helical (β-sheet) geometry (Marsh *et al.* 2006). This result is in good accordance with $^{15}N$ relaxation measurement (T1, T2, and heteronuclear NOE (HET-NOE) (II; Fig. S4, III; Fig S3). The interpretation of relaxation data is done using reduced spectral density mapping (II; Fig. 3, III; Fig. 2) (Farrow *et al.* 1994, Lefevre *et al.* 1996). Reduced spectral density mapping provides information about the overall backbone motion of the protein. For the region with transient secondary structure $J_{(0)}$ values are relatively higher, while for flexible regions the values are smaller. Similarly, the high frequency spectral density $J_{(0.87\omega H)}$ value is small for the region with transient secondary structure and higher for the regions with fast internal motion. The spectral density $J_{(\omega N)}$ exhibits relatively invariable values reflecting insensitivity to variation in backbone dynamics. Based on these observations, ARPPs and ENSA cannot be characterized by a single representative ensemble approach. Therefore, the generation of the dynamic structural ensemble is essential to elucidate the structural properties of ARPPs and ENSA. In this study, we used the ENSEMBLE program to generate a structural ensemble of ARPPs and ENSA using the experimental restraints from the NMR and SAXS experiments. ENSEMBLE program generates a large initial pool of conformers by molecular simulation and then selects the conformer using an algorithm that chooses the most relevant subset of conformers which best fits the experimental data.

The frequency of occurrence of IDPs is higher in eukaryotes in comparison to the archeal and bacterial proteome. Up to 33.0% of the proteins in eukaryotic proteomes are predicted to contain an intrinsically disordered region (IDR) that are usually ≥50 residues long (Dunker *et al.* 2001, Ward *et al.* 2004). The extended and flexible nature of IDPs is believed to represent a major functional advantage in the regulation of protein function and protein-protein interaction. According to our interaction study using MST, ARPPs and ENSA bind to the PP2A A-subunit with a modest affinity (low μM) which corroborates with the previous study by Andrade et al (Andrade *et al.* 2017). The intrinsically disordered proteins favor transient interactions, that is, high specificity with low affinity, because of diminishing conformational entropy upon binding (Mészáros *et al.* 2007, Zhou 2012, Mollica *et al.* 2016). IDPs can form a dynamic "fuzzy" complex, where the protein interchanges between various conformations of a dynamic ensemble in the bound state (Tompa and Fuxreiter 2008).

Further titration study with NMR spectroscopy showed a persistent line broadening of few residues in the NMR spectra of the ARPPs complexed with the PP2A A-subunit due to the dynamic behavior of the complex. We then identified the PP2A A-subunit binding site at ARPP-16 and ARPP-19 as a short sequential recognition element formed by the second transient α-helix and the flanking flexible region on both sides (Fig. 12). The interaction of IDPs to their partner is often mediated by a short linear motif and pre-formed structural elements (Fuxreiter *et al.* 2004, Mészáros *et al.* 2007, Perkins *et al.* 2010). There is a

consensus that IDPs undergo induced folding upon binding to their structured partner (Dyson and Wright 2002b, Mészáros *et al.* 2007, Wright and Dyson 2015). The major conformational rearrangements in ARPPs were not observed, but there is clear evidence for the presence of local conformational changes at the PP2A A-subunit binding region. The binding induced CSPs and line broadening of residues at the binding site could be explained either as reduced local molecular tumbling or conformational heterogeneity.

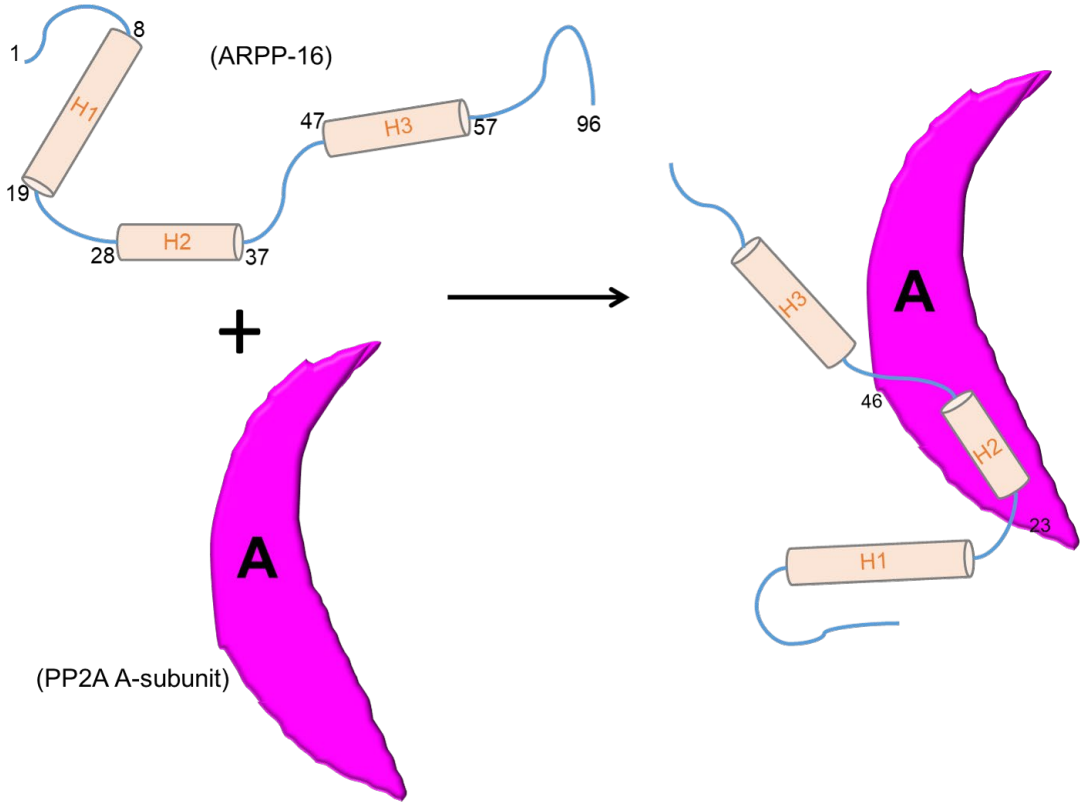


FIGURE 12 A schematic diagram showing possible mechanism of ARPP-16 binding with PP2A A-subunit. The three pre-structured transient helices with slightly different length were experimentally observed from SSP and relaxation studies of ARPP-16. NMR titration experiment mapped linear motif of residues 23-46 as a region that interact with PP2A A-subunit. ARPP-19 binds to PP2A A-subunit with the region corresponding to the PP2A A-subunit binding region in ARPP-16.

Interestingly, we could also map the PP2A A-subunit binding site in ENSA that consists of two distinct protein segments, (i) residues 30 to 45 (binding site 1, B1) and (ii) residues 51 to 76 (binding site 2, B2). The IDPs contain multiple linear motifs that are predicted to bind the same or different binding partners (Dyson and Wright 2002b, Wright and Dyson 2015). The binding of ENSA and PP2A A-subunit seems to fit suitably to the proposed "dock and coalesce" mechanism (Zhou 2012), where a linear motif of ENSA (B2) first docks to a sub-site on PP2A A-subunit and the next linear motif (B1) coalesce around its respective subsite on PP2A A-subunit (Fig. 13).

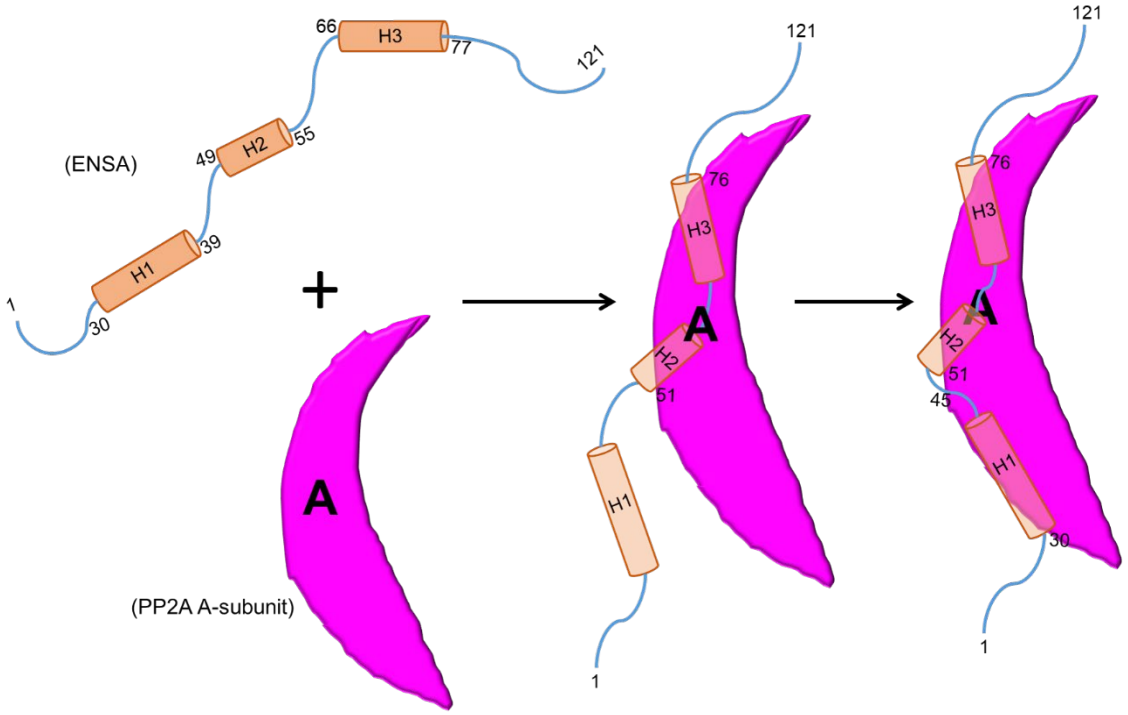


FIGURE 13 A schematic diagram showing possible mechanism of ENSA binding with PP2A A-subunit. The three pre-structured transient helices were experimentally observed from SSP and relaxation studies of ENSA. According to NMR titration experiment, ENSA binds to PP2A A-subunit using an extended binding region and comprise two binding site B1 (residues 30-45) and B2 (residues 51-76). We hypothesize that ENSA bind to PP2A A-subunit using "dock and coalesce" mechanism. In the first step a segment of ENSA docks to the PP2A A-subunit surface. Then, the docked segment of ENSA will guide the interaction of the remaining segment.

Though ARPP-19 and ENSA are highly homologous proteins, they bind to PP2A A-subunit differentially. One reason is that the N-terminal of ENSA is more flexible than the than of ARPPS due to the enrichment of disordered promoting residues, especially glutamate. The increased flexibility allows the protein to wrap around the surface of the target, forming an intricate interaction surface. There are examples of differential binding modes of highly homologous proteins to similar binding partners (Zhang *et al.* 2012, Iosub Amir *et al.* 2015). Our results suggest that ARPP-19 and ENSA can have distinct and specific interaction that regulates different biological processes. This is consistent with the difference in activity between ARPP-19 and ENSA reported previously, where ARPP-19 promotes mitotic entry while ENSA regulates S-phase progression (Charrasse *et al.* 2017, Hached *et al.* 2019).

Furthermore, we have characterized the interaction of ARPPs and ENSA with B56 subunits. The interaction between ARPP-16/19/ENSA and B56 is even weaker than the binding to A-subunit as was observed in the MST binding affinity measurements. The strength of binding of B56 subunits to ARPP wild type and ARPPs phosphomimicking mutants were compared. The resulting $K_D$ was found to be lower when binding to the phosphomimicking mutants of

ARPPs than binding to wild types. This could mean that the phosphorylation will increase the binding affinity of ARPPs to all B56 subunits. As already mentioned, IDPs can interact with multiple binding partners through several conserved sequence motifs, we hypothesize ARPPs bind to PP2A holoenzyme via a two-step mechanism. At first, ARPP binds to the PP2A A-subunit, which serves as the base of assembly, then the specific binding to PP2A B56-subunit is attained through the phosphorylated ARPP. In cells, the concentration of IDPs is maintained low by reduced transcription rate and increased degradation. The high abundance of IDPs can lead to undesirable promiscuous interaction and perturb the balance of a signaling network leading to different pathological conditions (Gsponer *et al.* 2008). Likewise, overexpression of ARPP-19 is reported in different cancer types (Song *et al.* 2014, Jiang *et al.* 2016, Mäkelä *et al.* 2019, Gao *et al.* 2020). Based on our results, MAST3 phosphorylation mimicking mutants of ARPPs bind stronger to B56δ and B56α than other B56 subunits under study. This is in agreement with the earlier study by Andrade et al, which reports that the MAST3 phosphorylated ARPP-16 selectively inhibits B55 and B56δ subunits (Andrade *et al.* 2017, Leslie and Nairn 2019).

The interaction study of ENSA and B56 subunits shows that the wild type ENSA is a stronger binder as compared to the wild type ARPPs. Our results differ from the previous study by Mochida et al which reported that ENSA does not bind with B56 subunits (Mochida *et al.* 2010). Their study was performed in the *Xenopus* interphase egg extract but the recent study has reported that PP2A B56 complex carries out a crucial function in M-phase (Vallardi *et al.* 2019).

A limitation of this work is that the effect of phosphorylation on binding was investigated using phosphomimicking strategy. Although phosphomimicking mutations, Ser to Asp/Glu, are frequently used to probe the structural and functional influence of protein phosphorylation, the Ser to Asp/Glu phosphomimicking mutations cannot fully take over the effect of phosphorylation on the structure and function of a protein. The charge and bulkiness/geometry of the phosphate group is different relative to the carboxyl group of Glu, therefore, phosphomimicking Ser to Glu might not be sufficient to observe the effect of phosphorylation (Chen and Cole 2015).

# 6 CONCLUSION

Based on the results from this study, the following conclusions can be drawn:

1. The structural characterization of ARPP-16, ARPP-19, and ENSA by the synergistic use of NMR and SAXS revealed the intrinsically disordered nature of these proteins. The ARPPs and ENSA, although being intrinsically disordered, have a propensity to form transient secondary structures. The transient secondary structure of the ARPPs and ENSA were characterized based on the secondary chemical shifts which revealed the presence of three short transiently helical regions.

2. Our MST based interaction study demonstrates that ARPPs and ENSA bind PP2A A-subunit with similar modest affinities. In addition, the interaction between ARPPs/ENSA and PP2A B56 is weak and transient as observed frequently in signal transductions.

3. We have characterized the interaction of ARPPs/ENSA with A-subunit of PP2A using NMR spectroscopy. Although ARPPs and ENSA have high sequence similarity and bind to PP2A A-subunit with similar affinities, yet they bind to PP2A A-subunit via two distinct binding mechanisms. The NMR titration of ARPPs with PP2A A-subunit showed that they bind to A-subunit via a linear motif comprising the second transient α-helical region and the random coil regions flanking on both sides. Meanwhile, the binding of ENSA to PP2A A-subunit is extended and consists of two binding interfaces that comprise all three transient α-helical regions.

*Acknowledgements*

This thesis work was carried out at the Department of Biological and Environmental Sciences and Nano Science Center, University of Jyväskylä, Finland with the funding support from University of Jyväskylä Graduate School (JYUGS).

First of all, I would like to express my sincere gratitude to my supervisor Dr. Ulla Pentikäinen for providing me an incredible opportunity to expand my academic cognizance in the bewitching world of proteins. I would like to thank you for giving me an appreciable amount of freedom in the research project, continuous guidance, support and motivation to learn new techniques throughout the project. I would like to express my gratitude and appreciation to my second supervisor Professor Perttu Permi whose expertise was invaluable in formulating the research methodology. His patient support, encouragement and insightful feedback have made this an exciting experience. I feel extremely grateful and lucky to have the opportunity to work with and learn from both of you.

Besides my supervisors, my greatest gratitude goes to my follow-up group members; Professor Jari Ylänne and Professor Maija Nissinen for their encouragement and insightful comments. Your invaluable suggestions were very helpful in achieving the goal of the project.

Next I would like to thank the pre-examiners of my thesis, Docent Tommi Kajander, University of Helsinki, Finland and Associate Professor Mikko Metsä-Ketalä, University of Turku, Finland. I appreciate your time and invaluable comments on my thesis, which helped to considerably improve the quality of the thesis. I would also like to thank Prof. Varpu Marjomäki, the scientific editor of the Jyväskylä Studies in Biological and Environmental Science, for her helpful comments.

I wish to extend my warmest thanks to all the past and present members (Tatu, Linda, Sami, and Pekka) of the UP group. Special thanks to Tatu, I am extremely grateful for your support and guidance in my academic endeavors. Thanks for making the entire time at JYU memorable, be it a coffee break or fishing during the lunch break. You are such an inspiration - maybe I should not attempt to express the inexpressible.

Thanks to Petri for your technical assistance in the lab and also for the memorable time during the coffee and lunch breaks - lunch time fishing with you and Tatu is a perfect way of escaping from the stress. You have a great sense of humor, a willingness to tease, and know how to be edgy and wacky.

Thanks to Lina and Kati - for the friendship that we shared. With all our discussions, I think we learned to find a practical solution to impractical problems like failure of well-optimized experiments, having results that are mostly negative or meaningless, and doubting ourselves as a researcher. Thanks for stopping by my office to check in on me. I feel very lucky to have you as co-workers. I would also like to thank Kati for her time and energy in organizing a get-together and other fun-filled activities.

I would also like to thank Laura, Alli, Rahul, Dhanik, Sailee, Mira, Visa and all the people in the department. I've always believed that the people you work with are what makes the job worthwhile, and you all proved me right. Thanks for the incredible memories and for making the workplace a fun place to be.

Thanks to Jahangir, Tanvir, Abhi, Pooja, and Sudarshan for all the wonderful memories that we shared in and outside the University of Oulu.
I want to thank my Nepali friends in Oulu Bikram, Shiv, Khem, Sajeen, Subodh, and my cousin Saujanya. Thanks a lot for adding so many good memories and times of happiness to my life.

Thanks to Roshan, Kreetika, and Reena - I thank each of you for your beautiful way of being and giving me so much of love. I am so grateful to be in your very guardianship.

At last, but most importantly, my parents and family members, who always supported me in doing what I wanted to do, even in moving so far away from them to chase my dream. Thank you for instilling a strong passion for learning in me, and for doing everything possible to put me on the path to greatness. I owe my success to you and I hope that you will always be as happy as I feel today.

## REFERENCES


Amoroso S., Schmid-antomarchi H., Fosset M. & Lazdunskit M. 1989. Glucose , Sulfonylureas , and Neurotransmitter. *Science* 247: 852–854.

Andrade E.C., Musante V., Horiuchi A., Matsuzaki H., Brody A.H., Wu T., Greengard P., Taylor J.R. & Nairn A.C. 2017. ARPP-16 Is a Striatal-Enriched Inhibitor of Protein Phosphatase 2A Regulated by Microtubule-Associated Serine/Threonine Kinase 3 (Mast 3 Kinase). *The Journal of Neuroscience* 37: 2709–2722.

Arroyo J.D. & Hahn W.C. 2005. Involvement of PP2A in viral and cellular transformation. *Oncogene* 24: 7746–7755.

Baskaran R. & Velmurugan B.K. 2018. Protein phosphatase 2A as therapeutic targets in various disease models. *Life Sciences* 210: 40–46.

Bataille D., Héron L., Virsolvy A., Peyrollier K., LeCam A., Gros L. & Blache P. 1999. A-Endosulfine, a New Entity in the Control of Insulin Secretion. *Cellular and Molecular Life Sciences* 56: 78–84.

Bernadó P. 2010. Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. *European Biophysics Journal* 39: 769–780.

Bernadó P. & Svergun D.I. 2012. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Molecular BioSystems* 8: 151–167.

Bhardwaj A., Singh S., Srivastava S.K., Honkanen R.E., Reed E. & Singh A.P. 2011. Modulation of protein phosphatase 2A (PP2A) activity alters androgen-independent growth of prostate cancer cells: therapeutic implications. *MolecularCancer Therapeutics*. 10: 720–731.

Bosch M., Cayla X., Hoof C. Van, Hemmings B.A., Ozon R., Merlevede W. & Goris J. 1995. The PR55 and PR65 Subunits of Protein Phosphatase 2A from Xenopus laevis Molecular Cloning and Developmental Regulation of Expression. *European Journal of Biochemistry* 230: 1037–1045.

Calin G.A., Iasio M.G. Di, Caprini E., Vorechovsky I., Natali P.G., Sozzi G., Croce C.M., Barbanti-Brodano G., Russo G. & Negrini M. 2000. Low frequency of alterations of the α (PPP2R1A) and β (PPP2R1B) isoforms of the subunit A of the serine-threonine phosphatase 2A in human neoplasms. *Oncogene* 19: 1191–1195.

Charrasse S., Raynaud P., Burgess A., Schwob E., Vera J., Castro A., Gharbi-Ayachi A., Lorca T. & Hached K. 2017. Ensa controls S-phase length by modulating Treslin levels. *Nature Communications* 8: 1–14.

Chen W., Arroyo J.D., Timmons J.C., Possemato R. & Hahn W.C. 2005. Cancer-associated PP2A Aα subunits induce functional haploinsufficiency and tumorigenicity. *Cancer Research* 65: 8183–8192.

Chen Z. & Cole P.A. 2015. Synthetic approaches to protein phosphorylation. *Current Opinion in Chemical Biology* 28: 115–122.

Chen J., Martin B.L. & Brautigan D.L. 1992. Regulation of protein serine-threonine phosphatase type-2A by tyrosine phosphorylation. *Science* 257: 1261–1264.

Chen W., Possemato R., Campbell K.T., Plattner C.A., Pallas D.C. & Hahn W.C. 2004. Identification of specific PP2A complexes involved in human cell transformation. *Cancer Cell* 5: 127–136.

Cho U.S. & Xu W. 2007. Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme. *Nature* 445: 53–57.

Cohen P.T.W. 1997. Novel protein serine/threonine phosphatases: Variety is the spice of life. *Trends in Biochemical Sciences* 22: 245–251.

Colella S., Ohgaki H., Ruediger R., Yang F., Nakamura M., Fujisawa H., Kleihues P. & Walter G. 2001. Reduced expression of the Aα subunit of protein phosphatase 2A in human gliomas in the absence of mutations in the Aα and Aβ subunit genes. *International Journal of Cancer* 93: 798–804.

Cristóbal I., Cirauqui C., Castello-Cros R., Garcia-Orti L., Calasanz M.J. & Odero M.D. 2013. Downregulation of PPP2R5E is a common event in acute myeloid leukemia that affects the oncogenic potential of leukemic cells. *Haematologica* 98: 103–104.

Csizmók V., Szollosi E., Friedrich P. & Tompa P. 2006. A novel two-dimensional electrophoresis technique for the identification of intrinsically unstructured proteins. *Molecular and Cellular Proteomics* 5: 265–273.

Damle N.P. & Köhn M. 2019. The human DEPhOsphorylation Database DEPOD: 2019 update. *Database : the journal of biological databases and curation* 2019: 1–7.

Deichmann M., Polychronidis M., Wacker J., Thome M. & Näher H. 2001. The protein phosphatase 2A subunit Bγ gene is identified to be differentially expressed in malignant melanomas by subtractive suppression hybridization. *Melanoma Research* 11: 577–585.

Delaforge E., Cordeiro T.N., Bernadó P. & Sibille N. 2017. *Conformational Characterization of Intrinsically Disordered Proteins and Its Biological Significance* Webb G.A. (ed.). Springer International Publishing, Cham.

Dulubova I., Horiuchi A., Snyder G.L., Girault J.A., Czernik A.J., Shao L., Ramabhadran R., Greengard P. & Nairn A.C. 2001. ARPP-16/ARPP-19: A highly conserved family of cAMP-regulated phosphoproteins. *Journal of Neurochemistry* 77: 229–238.

Dunker A.K., Lawson J.D., Brown C.J., Williams R.M., Romero P., Oh J.S., Oldfield C.J., Campen A.M., Ratliff C.M., Hipps K.W., Ausio J., Nissen M.S., Reeves R., Kang C.H., Kissinger C.R., Bailey R.W., Griswold M.D., Chiu W., Garner E.C. & Obradovic Z. 2001. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* 19: 26–59.

Dupre A., Buffin E., Roustan C., Nairn A.C., Jessus C. & Haccard O. 2013. The phosphorylation of ARPP19 by Greatwall renders the auto-amplification of MPF independently of PKA in Xenopus oocytes. *Journal of Cell Science* 126: 3916–3926.

Dyson H.J. & Wright P.E. 2002a. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. In: *Advances*, pp. 311–340.

Dyson H.J. & Wright P.E. 2002b. Coupling of folding and binding for unstructured proteins. : 54–60.

Dyson H.J. & Wright P.E. 2004. Unfolded proteins and protein folding studied by NMR. *Chemical Reviews* 104: 3607–3622.

Eichhorn P.J.A., Creyghton M.P. & Bernards R. 2009. Protein phosphatase 2A regulatory subunits and cancer. *Biochimica et Biophysica Acta - Reviews on Cancer* 1795: 1–15.

Erdős G. & Dosztányi Z. 2020. Analyzing Protein Disorder with IUPred2A. *Current protocols in bioinformatics* 70: e99.

Esplin E.D., Ramos P., Martinez B., Tomlinson G.E., Mumby M.C. & Evans G.A. 2006. The glycine 90 to aspartate alteration in the Aβ subunit of PP2A (PPP2R1B) associates with breast cancer and causes a deficit in protein function. *Genes, Chromosomes and Cancer* 45: 182–190.

Farrow N.A., Muhandiram R., Singer A.U., Pascal S.M., Kay C.M., Gish G., Shoelson S.E., Pawson T., Forman-Kay J.D. & Kay L.E. 1994. Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. *Biochemistry* 33: 5984–6003.

Fuxreiter M., Simon I., Friedrich P. & Tompa P. 2004. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *Journal of Molecular Biology* 338: 1015–1026.

Gao X., Lu C., Chen C., Sun K., Liang Q., Shuai J., Wang X. & Xu Y. 2020. ARPP-19 Mediates Herceptin Resistance via Regulation of CD44 in Gastric Cancer. *OncoTargets and Therapy* Volume 13: 6629–6643.

García-Blanco N., Vázquez-Bolado A. & Moreno S. 2019. Greatwall-endosulfine: A molecular switch that regulates PP2A/B55 protein phosphatase activity in dividing and quiescent cells. *International Journal of Molecular Sciences* 20.

Gharbi-Ayachi A., Labbe J.-C., Burgess A., Vigneron S., Strub J.-M., Brioudes E., Van-Dorsselaer A., Castro A. & Lorca T. 2010. The Substrate of Greatwall Kinase, Arpp19, Controls Mitosis by Inhibiting Protein Phosphatase 2A. *Science* 330: 1673–1677.

Girault J., Horiuchi A., Gustafson E., Rosen N. & Greengard P. 1990. Differential expression of ARPP-16 and ARPP-19, two highly related cAMP- regulated phosphoproteins, one of which is specifically associated with dopamine-innervated brain regions. *The Journal of Neuroscience* 10: 1124–1133.

Gong Y., Wu W., Zou X., Liu F., Wei T. & Zhu J. 2018. MiR-26a inhibits thyroid cancer cell proliferation by targeting ARPP19. *American journal of cancer research* 8: 1030–1039.

Götz J., Probst A., Ehler E., Hemmings B. & Kues W. 1998. Delayed embryonic lethality in mice lacking protein phosphatase 2A catalytic subunit Cα. *Proceedings of the National Academy of Sciences of the United States of America* 95: 12370–12375.

Götz J., Probst A., Mistl C., Nitsch R.M. & Ehler E. 2000. Distinct role of protein phosphatase 2A subunit Cα in the regulation of E-cadherin and β-catenin during development. *Mechanisms of Development* 93: 83–93.

Gould K.L. & Nurse P. 1989. Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. *Nature* 342: 39–45.

Groves M.R., Hanlon N., Turowski P., Hemmings B.A. & Barford D. 1999. The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* 96: 99–110.

Gsponer J., Futschik M.E., Teichmann S.A. & Babu M.M. 2008. Tight Regulation of Unstructured Proteins: From Transcript Synthesis to Protein Degradation. *Science* 322: 1365–1368.

Gu P., Qi X., Zhou Y., Wang Y. & Gao X. 2012. Generation of Ppp2Ca and Ppp2Cb conditional null alleles in mouse. *Genesis* 50: 429–436.

Guo H. & Damuni Z. 1993. Autophosphorylation-activated protein kinase phosphorylates and inactivates protein phosphatase 2A. *Proceedings of the National Academy of Sciences of the United States of America* 90: 2500–2504.

Hached K., Goguet P., Charrasse S., Vigneron S., Sacristan M.P., Lorca T. & Castro A. 2019. ENSA and ARPP19 differentially control cell cycle progression and development. *The Journal of cell biology* 218: 541–558.

Haesen D., Asbagh L.A., Derua R., Hubert A., Schrauwen S., Hoorne Y., Amant F., Waelkens E., Sablina A. & Janssens V. 2016. Recurrent PPP2R1A mutations in uterine cancer act through a dominant-negative mechanism to promote malignant cell growth. *Cancer Research* 76: 5719–5731.

Haesen D., Sents W., Lemaire K., Hoorne Y. & Janssens V. 2014. The basic biology of PP2A in hematologic cells and malignancies. *Frontiers in Oncology* 4: 1–11.

Hardman G., Perkins S., Brownridge P.J., Clarke C.J., Byrne D.P., Campbell A.E., Kalyuzhnyy A., Myall A., Eyers P.A., Jones A.R. & Eyers C.E. 2019. Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation. *The EMBO Journal* 38: 1–22.

Hemmings B.A., Adams-Pearson C., Maurer F., Müller P., Goris J., Merlevede W., Hofsteenge J. & Stone S.R. 1990. α- and βForms of the 65-kDa Subunit of Protein Phosphatase 2A Have a Similar 39 Amino Acid Repeating Structure. *Biochemistry* 29: 3166–3173.

Heron L., Virsolvy A., Peyrollier K., Gribble F.M., Cam A. Le, Ashcroft F.M. & Bataille D. 1998. Human α-endosulfine, a possible regulator of sulfonylurea-sensitive KATP channel: Molecular cloning, expression and biological properties. *Proceedings of the National Academy of Sciences of the United States of America* 95: 8387–8391.

Hoof C. Van & Goris J. 2004. PP2A fulfills its promises as tumor suppressor: Which subunits are important? *Cancer Cell* 5: 105–106.

Horiuchi A., Williams K.R., Kurihara T., Nairn A.C. & Greengard P. 1990. Purification and cDNA cloning of ARPP-16, a cAMP-regulated phosphoprotein enriched in Basal Ganglia, and of a related phosphoprotein, ARPP-19. *Journal of Biological Chemistry* 265: 9476–9484.

Hornbeck P. V., Zhang B., Murray B., Kornhauser J.M., Latham V. & Skrzypek E. 2015. PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research* 43: D512–D520.

Hunter T. 2012. Why nature chose phosphate to modify proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367: 2513–2516.

Hunter T. & Sefton B.M. 1980. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proceedings of the National Academy of Sciences* 77: 1311–1315.

Inagaki N., Gonoi T., Clement J.P., Namba N., Inazawa J., Gonzalez G., Aguilar-Bryan L., Seino S. & Bryan J. 1995. Reconstitution of I(KATP): An Inward Rectifier Subunit Plus the Sulfonylurea Receptor. *Science* 270: 1166–1170.

Iosub Amir A., Rosmalen M. Van, Mayer G., Lebendiker M., Danieli T. & Friedler A. 2015. Highly homologous proteins exert opposite biological activities by using different interaction interfaces. *Scientific Reports* 5: 1–11.

Irwin N., Chao S., Nairn A.C., Goritchenko L., Benowitz L.I., Greengard P. & Horiuchi A. 2002. Nerve growth factor controls GAP-43 mRNA stability via the phosphoprotein ARPP-19. *Proceedings of the National Academy of Sciences* 99: 12427–12431.

Ito A., Kataoka T.R., Watanabe M., Nishiyama K., Mazaki Y., Sabe H., Kitamura Y. & Nojima H. 2000. A truncated isoform of the PP2A B56 subunit promotes cell motility through paxillin phosphorylation. *EMBO Journal* 19: 562–571.

Janssens V. & Goris J. 2001. Protein phosphatase 2A: A highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochemical Journal* 353: 417–439.

Janssens V., Goris J. & Hoof C. Van. 2005. PP2A: The expected tumor suppressor. *Current Opinion in Genetics and Development* 15: 34–41.

Jeong A.L., Han S., Lee S., Su Park J., Lu Y., Yu S., Li J., Chun K.H., Mills G.B. & Yang Y. 2016. Patient derived mutation W257G of PPP2R1A enhances cancer cell migration through SRC-JNK-c-Jun pathway. *Scientific Reports* 6: 1–12.

Jiang T., Zhao B., Li X. & Wan J. 2016. ARPP-19 promotes proliferation and metastasis of human glioma. *NeuroReport* 27: 960–966.

Jin Z., Wallace L., Harper S.Q. & Yang J. 2010. PP2A:B56ε, a substrate of caspase-3, regulates p53-dependent and p53-independent apoptosis during development. *Journal of Biological Chemistry* 285: 34493–34502.

Jones S., Wang T.-L., Shih I.-M., Mao T.-L., Nakayama K., Roden R., Glas R., Slamon D., Diaz L.A., Vogelstein B., Kinzler K.W., Velculescu V.E. & Papadopoulos N. 2010. Frequent Mutations of Chromatin Remodeling Gene ARID1A in Ovarian Clear Cell Carcinoma. *Science* 330: 228–231.

Juanes M.A., Khoueiry R., Kupka T., Castro A., Mudrak I., Ogris E., Lorca T. & Piatti S. 2013. Budding Yeast Greatwall and Endosulfines Control Activity and Spatial Regulation of PP2ACdc55 for Timely Mitotic Progression. *PLoS Genetics* 9: 1–14.

Junttila M.R., Puustinen P., Niemelä M., Ahola R., Arnold H., Böttzauw T., Ala-aho R., Nielsen C., Ivaska J., Taya Y., Lu S.L., Lin S., Chan E.K.L., Wang X.J., Grènman R., Kast J., Kallunki T., Sears R., Kähäri V.M. & Westermarck J. 2007. CIP2A Inhibits PP2A in Human Malignancies. *Cell* 130: 51–62.

Kalla C., Scheuermann M.O., Kube I., Schlotter M., Mertens D., Döhner H., Stilgenbauer S. & Lichter P. 2007. Analysis of 11q22-q23 deletion target genes in B-cell chronic lymphocytic leukaemia: Evidence for a pathogenic role of NPAT, CUL5, and PPP2R1B. *European Journal of Cancer* 43: 1328–1335.

Kamibayashi C., Estes R., Lickteigs L., Yango S., Cram C. & Marc C. 1994. Comparison of Heterotrimeric Protein Phosphatase 2A Containing. *Biochemistry*: 20139–20148.

Kauko O. & Westermarck J. 2018. Non-genomic mechanisms of protein phosphatase 2A (PP2A) regulation in cancer. *International Journal of Biochemistry and Cell Biology* 96: 157–164.

Khew-Goodall Y. & Hemmings B.A. 1988. Tissue-specific expression of mRNAs encoding α- and β-catalytic subunits of protein phosphatase 2A. *FEBS Letters* 238: 265–268.

Khoury G.A., Baliban R.C. & Floudas C.A. 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1: 1–5.

Kim S.H. & Lubec G. 2001a. Decreased alpha-endosulfine, an endogenous regulator of ATP-sensitive potassium channels, in brains from adult Down syndrome patients. In: *Protein Expression in Down Syndrome Brain*, Springer Vienna, Vienna, pp. 1–9.

Kim S.H. & Lubec G. 2001b. Brain α-endosulfine is manifold decreased in brains from patients with Alzheimer's disease: A tentative marker and drug target? *Neuroscience Letters* 310: 77–80.

Kim S.H., Nairn A.C., Cairns N. & Lubec G. 2001. Decreased levels of ARPP-19 and PKA in brains of Down syndrome and Alzheimer's disease. *Journal of Neural Transmission, Supplement*: 263–272.

Kishimoto T. 2015. Entry into mitosis: a solution to the decades-long enigma of MPF. *Chromosoma* 124: 417–428.

Lambrecht C., Libbrecht L., Sagaert X., Pauwels P., Hoorne Y., Crowther J., Louis J. V., Sents W., Sablina A. & Janssens V. 2018. Loss of protein phosphatase 2A regulatory subunit B56δ promotes spontaneous tumorigenesis in vivo. *Oncogene* 37: 544–552.

Lee K., Brownhill V. & Richardson P.J. 1997. Antidiabetic Sulphonylureas Stimulate Acetylcholine Release from Striatal Cholinergic Interneurones Through Inhibition of $K_{ATP}$ Channel Activity. *Journal of Neurochemistry* 69: 1774–1776.

Lefevre J.F., Dayie K.T., Peng J.W. & Wagner G. 1996. Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. *Biochemistry* 35: 2674–2686.

Leslie S.N. & Nairn A.C. 2019. cAMP regulation of protein phosphatases PP1 and PP2A in brain. *Biochimica et Biophysica Acta - Molecular Cell Research* 1866: 64–73.

Li H.H., Cai X., Shouse G.P., Piluso L.G. & Liu X. 2007. A specific PP2A regulatory subunit, B56γ, mediates DNA damage-induced dephosphorylation of p53 at Thr55. *EMBO Journal* 26: 402–411.

Longin S., Zwaenepoel K., Louis J. V., Dilworth S., Goris J. & Janssens V. 2007. Selection of protein phosphatase 2A regulatory subunits is mediated by the C terminus of the catalytic subunit. *Journal of Biological Chemistry* 282: 26971–26980.

Lorca T. & Castro A. 2013. The Greatwall kinase: a new pathway in the control of the cell cycle. *Oncogene* 32: 537–543.

Lü M., Ding K., Zhang G., Yin M., Yao G., Tian H., Lian J., Liu L., Liang M., Zhu T. & Sun F. 2015. MicroRNA-320a sensitizes tamoxifen-resistant breast cancer cells to tamoxifen by targeting ARPP-19 and ERRγ. *Scientific reports* 5: 8735.

Mäkelä E., Löyttyniemi E., Salmenniemi U., Kauko O., Varila T., Kairisto V., Itälä-Remes M. & Westermarck J. 2019. Arpp19 Promotes Myc and Cip2a Expression and Associates with Patient Relapse in Acute Myeloid Leukemia. *Cancers* 11: 1774.

Mannava S., Omilian A.R., Wawrzyniak J.A., Fink E.E., Zhuang D., Miecznikowski J.C., Marshall J.R., Soengas M.S., Sears R.C., Morrison C.D. & Nikiforov M.A. 2012. PP2A-B56α controls oncogene-induced senescence in normal and tumor human melanocytic cells. *Oncogene* 31: 1484–1492.

Manning G., Whyte D.B., Martinez R., Hunter T. & Sudarsanam S. 2002. The protein kinase complement of the human genome. *Science* 298: 1912–1934.

Marsh J.A., Singh V.K., Jia Z. & Forman-Kay J.D. 2006. Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Science* 15: 2795–2804.

McConechy M.K., Anglesio M.S., Kalloger S.E., Yang W., Senz J., Chow C., Heravi-Moussavi A., Morin G.B., Mes-Masson A.M., Bowtell D., Chenevix-Trench G., DeFazio A., Gertig D., Green A., Webb P., Carey M.S., McAlpine J.N., Kwon J.S., Prentice L.M., Boyd N., Shah S.P., Gilks C.B. & Huntsman D.G. 2011. Subtype-specific mutation of PPP2R1A in endometrial and ovarian carcinomas. *Journal of Pathology* 223: 567–573.

McCright B., Rivers A.M., Audlin S. & Virshup D.M. 1996. The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to both nucleus and cytoplasm. *Journal of Biological Chemistry* 271: 22081–22089.

Meeusen B. & Janssens V. 2017. Tumor suppressive protein phosphatases in human cancer: emerging targets for therapeutic intervention and tumor stratification. *The International Journal of Biochemistry & Cell Biology*: 1–37.

Meeusen B. & Janssens V. 2018. Tumor suppressive protein phosphatases in human cancer: Emerging targets for therapeutic intervention and tumor stratification. *International Journal of Biochemistry and Cell Biology* 96: 98–134.

Mei Li, Anthony Makkinje Z.D. 1996. The Myeloid Leukemia- associated Protein SET Is a Potent Inhibitor of Protein Phosphatase 2A. *The Journal of Biological Chemistry* 271: 11059–11062.

Mészáros B., Erdös G. & Dosztányi Z. 2018. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Research* 46: W329–W337.

Mészáros B., Tompa P., Simon I. & Dosztányi Z. 2007. Molecular Principles of the Interactions of Disordered Proteins. *Journal of Molecular Biology* 372: 549–561.

Mochida S., Maslen S.L., Skehel M. & Hunt T. 2010. Greatwall Phosphorylates an Inhibitor of Protein Phosphatase 2A That Is Essential for Mitosis. *Science* 330: 1670–1673.

Mochida S., Rata S., Hino H., Nagai T. & Novák B. 2016. Two Bistable Switches Govern M Phase Entry. *Current Biology* 26: 3361–3367.

Mollica L., Bessa L.M., Hanoulle X., Jensen M.R., Blackledge M. & Schneider R. 2016. Binding mechanisms of intrinsically disordered proteins: Theory, simulation, and experiment. *Frontiers in Molecular Biosciences* 3: 1–18.

Moreno S., Nurse P. & Russell P. 1990. Regulation of mitosis by cyclic accumulation of p80cdc25 mitotic inducer in fission yeast. *Nature* 344: 549–552.

Mumby M.C. & Walter G. 1991. Protein phosphatases and DNA tumor viruses: Transformation through the back door? *Molecular Biology of the Cell* 2: 589–598.

Mumby M.C. & Walter G. 1993. Protein serine/threonine phosphatases: structure, regulation, and functions in cell growth. *Physiological Reviews* 73: 673–699.

Musante V., Li L., Kanyo J., Lam T.T., Colangelo C.M., Cheng S.K., Brody A.H., Greengard P., Novère N. Le & Nairn A.C. 2017. Reciprocal regulation of ARPP-16 by PKA and MAST3 kinases provides a cAMP-regulated switch in protein phosphatase 2A inhibition. *eLife* 6: 1–32.

Nagendra D.C., Burke J., Maxwell G.L. & Risinger J.I. 2012. PPP2R1A mutations are common in the serous type of endometrial cancer. *Molecular Carcinogenesis* 51: 826–831.

Nobumori Y., Shouse G.P., Fan L. & Liu X. 2012. HEAT repeat 1 motif is required for B56γ-containing protein phosphatase 2A (B56γ-PP2A) holoenzyme assembly and tumor-suppressive function. *Journal of Biological Chemistry* 287: 11030–11036.

Nobumori Y., Shouse G.P., Wu Y., Lee K.J., Shen B. & Liu X. 2013. B56γ tumor-associated mutations provide new mechanisms for B56γ-PP2A tumor suppressor activity. *Molecular Cancer Research* 11: 995–1003.

O'Farrell P.H. 2001. Triggering the all-or-nothing switch into mitosis. *Trends in Cell Biology* 11: 512–519.

Ogris E., Du X., Nelson K.C., Mak E.K., Yu X.X., Lane W.S. & Pallas D.C. 1999. A protein phosphatase methylaesterase (PME-1) is one of several novel proteins stably associateing with two inactive mutants of protein phosphatase 2A. *Journal of Biological Chemistry* 274: 14382–14391.

Okumura E., Fukuhara T., Yoshida H., Hanada S.I., Kozutsumi R., Mori M., Tachibana K. & Kishimoto T. 2002. Akt inhibits Myt1 in the signalling pathway that leads to meiotic G2/M-phase transition. *Nature Cell Biology* 4: 111–116.

Okumura E., Morita A., Wakai M., Mochida S., Hara M. & Kishimoto T. 2014. Cyclin B-Cdk1 inhibits protein phosphatase PP2A-B55 via a greatwall kinase-independent mechanism. *Journal of Cell Biology* 204: 881–889.

Olsen J. V., Blagoev B., Gnad F., Macek B., Kumar C., Mortensen P. & Mann M. 2006. Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127: 635–648.

Perkins J.R., Diboun I., Dessailly B.H., Lees J.G. & Orengo C. 2010. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* 18: 1233–1243.

Peyrollier K., Heron L., Virsolvy-Vergine A., Cam A. Le & Bataille D. 1996. A Endosulfine Is a Novel Molecule, Structually Related To a Family of

Phosphoproteins. *Biochemical and Biophysical Research Communications* 223: 583–586.

Porter I.M., Schleicher K., Porter M. & Swedlow J.R. 2013. Bod1 regulates protein phosphatase 2A at mitotic kinetochores. *Nature Communications* 4: 1–9.

Price N.E. & Mumby M.C. 2000. Effects of regulatory subunits on the kinetics of protein phosphatase 2A. *Biochemistry* 39: 11312–11318.

Ramaswamy K., Spitzer B. & Kentsis A. 2015. Therapeutic re-activation of protein phosphatase 2A in acute myeloid leukemia. *Frontiers in Oncology* 5: 1–5.

Receveur-Brechot V. & Durand D. 2012. How random are intrinsically disordered proteins? A small angle scattering perspective. *Current protein & peptide science* 13: 55–75.

Roberts K.G., Smith A.M., McDougall F., Carpenter H., Horan M., Neviani P., Powell J.A., Thomas D., Guthridge M.A., Perrotti D., Sim A.T.R., Ashman L.K. & Verrills N.M. 2010. Essential requirement for PP2A inhibition by the oncogenic receptor c-KIT suggests PP2A reactivation as a strategy to treat c-KIT+ cancers. *Cancer research* 70: 5438–5447.

Rocher G., Letourneux C., Lenormand P. & Porteu F. 2007. Inhibition of B56-containing protein phosphatase 2As by the early response gene IEX-1 leads to control of Akt activity. *Journal of Biological Chemistry* 282: 5468–5477.

Romero P., Obradovic Z., Li X., Garner E.C., Brown C.J. & Dunker A.K. 2001. Sequence complexity of disordered protein. *Proteins: Structure, Function, and Genetics* 42: 38–48.

Ruediger R., Pham H.T. & Walter G. 2001. Disruption of Protein phosphatase 2A subunit interaction in human cancers with mutations in the Aa subunit gene. *Oncogene* 20: 1892–1899.

Ruediger R., Ruiz J. & Walter G. 2011. Human Cancer-Associated Mutations in the A Subunit of Protein Phosphatase 2A Increase Lung Cancer Incidence in A Knock-In and Knockout Mice. *Molecular and Cellular Biology* 31: 3832–3844.

Sablina A.A., Chen W., Arroyo J.D., Corral L., Hector M., Bulmer S.E., DeCaprio J.A. & Hahn W.C. 2007. The Tumor Suppressor PP2A Aβ Regulates the RalA GTPase. *Cell* 129: 969–982.

Sablina A.A., Hector M., Colpaert N. & Hahn W.C. 2010. Identification of PP2A complexes and pathways involved in cell transformation. *Cancer Research* 70: 10474–10484.

Sakura H., Ämmälä C., Smith P.A., Gribble F.M. & Ashcroft F.M. 1995. Cloning and functional expression of the cDNA encoding a novel ATP-sensitive potassium channel subunit expressed in pancreatic β-cells, brain, heart and skeletal muscle. *FEBS Letters* 377: 338–344.

Schwartz G.K. & Shah M.A. 2005. Targeting the cell cycle: A new approach to cancer therapy. *Journal of Clinical Oncology* 23: 9408–9421.

Schwarzinger S., Kroon G.J.A., Foss T.R., Chung J., Wright P.E. & Dyson H.J. 2001. Sequence-dependent correction of random coil NMR chemical shifts. *Journal of the American Chemical Society* 123: 2970–2978.

Seeling J.M., Miller J.R., Gil R., Moon R.T., White R. & Virshup D.M. 1999. Regulation of beta-catenin signaling by the B56 subunit of protein phosphatase 2A. *Science (New York, N.Y.)* 283: 2089–2091.

Shenolikar S. 1995. Protein phosphatase regulation by endogenous inhibitors. *Seminars in Cancer Biology* 6: 219–227.

Shi Y. 2009. Serine/Threonine Phosphatases: Mechanism through Structure. *Cell* 139: 468–484.

Shouse G.P., Nobumori Y. & Liu X. 2010. A B56γ mutation in lung cancer disrupts the p53-dependent tumor-suppressor function of protein phosphatase 2A. *Oncogene* 29: 3933–3941.

Singh A.P., Bafna S., Chaudhary K., Venkatraman G., Smith L., Eudy J.D., Johansson S.L., Lin M.F. & Batra S.K. 2008. Genome-wide expression profiling reveals transcriptomic variation and perturbed gene networks in androgen-dependent and androgen-independent prostate cancer cells. *Cancer Letters* 259: 28–38.

Solomon M.J., Glotzer M., Lee T.H., Philippe M. & Kirschner M.W. 1990. Cyclin activation of p34cdc2. *Cell* 63: 1013–1024.

Song H., Pan J., Liu Y., Wen H., Wang L., Cui J., Liu Y., Hu B., Yao Z. & Ji G. 2014. Increased ARPP-19 expression is associated with hepatocellular carcinoma. *International Journal of Molecular Sciences* 16: 178–192.

Takagi Y., Futamura M., Yamaguchi K., Aoki S., Takahashi T. & Saji S. 2000. Alterations of the PPP2R1B gene located at 11q23 in human colorectal cancers. *Gut* 47: 268–271.

Tompa P. 2010. *Structure and function of intrinsically disordered proteins*.

Tompa P. & Fuxreiter M. 2008. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends in Biochemical Sciences* 33: 2–8.

Tria G., Mertens H.D.T., Kachala M. & Svergun D.I. 2015. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* 2: 207–217.

Uversky V.N. 2002. Natively unfolded proteins: A point where biology waits for physics. *Protein Science* 11: 739–756.

Uversky V.N., Gillespie J.R. & Fink A.L. 2000. Why are 'natively unfolded' proteins unstructured under physiologic conditions? *Proteins: Structure, Function and Genetics* 41: 415–427.

Vallardi G., Allan L.A., Crozier L. & Saurin A.T. 2019. Division of labour between pp2a-b56 isoforms at the centromere and kinetochore. *eLife* 8.

Vigneron S., Brioudes E., Burgess A., Labbé J.C., Lorca T. & Castro A. 2009. Greatwall maintains mitosis through regulation of PP2A. *EMBO Journal* 28: 2786–2793.

Vigneron S., Robert P., Hached K., Sundermann L., Charrasse S., Labbé J.C., Castro A. & Lorca T. 2016. The master greatwall kinase, a critical regulator of mitosis and meiosis. *International Journal of Developmental Biology* 60: 245–254.

Virsolvy-vergine A., Leray H., Kuroki S.H.O., Lupo B., Dufour M. & Bataille D. 1992. Endosulfine , an endogenous peptidic ligand for the sulfonylurea receptor : Purification and partial characterization from ovine brain. *Proceedings of the National Academy of Sciences of the United States of America* 89: 6629–6633.

Wang S.S., Esplin E.D., Li J.L., Huang L., Gazdar A., Minna J. & Evans G.A. 1998. Alterations of the PPP2R1B gene in human lung and colon cancer. *Science* 282: 284–287.

Ward J.J., Sodhi J.S., McGuffin L.J., Buxton B.F. & Jones D.T. 2004. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology* 337: 635–645.

Wera S. & Hemmings B.A. 1995. Serine/threonine protein phosphatases. *Biochemical Journal* 311: 17–29.

Wishart D.S. & Sykes B.D. 1994. Chemical shifts as a tool for structure determination. *Methods in Enzymology* 239: 363–392.

Wright P.E. & Dyson H.J. 1999. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* 293: 321–331.

Wright P.E. & Dyson H.J. 2015. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* 16: 18–29.

Xing Y., Xu Y., Chen Y., Jeffrey P.D., Chao Y., Lin Z., Li Z., Strack S., Stock J.B. & Shi Y. 2006. Structure of Protein Phosphatase 2A Core Enzyme Bound to Tumor-Inducing Toxins. *Cell* 127: 341–353.

Xu Y., Xing Y., Chen Y., Chao Y., Lin Z., Fan E., Yu J.W., Strack S., Jeffrey P.D. & Shi Y. 2006. Structure of the Protein Phosphatase 2A Holoenzyme. *Cell* 127: 1239–1251.

Xue B., Dunbrack R.L., Williams R.W., Dunker A.K. & Uversky V.N. 2010. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1804: 996–1010.

Ye H., Jin Q., Wang X. & Li Y. 2020. MicroRNA-802 inhibits cell proliferation and induces apoptosis in human laryngeal cancer by targeting cAMP-regulated phosphoprotein 19. *Cancer Management and Research* 12: 419–430.

Ysselstein D., Dehay B., Costantino I.M., McCabe G.P., Frosch M.P., George J.M., Bezard E. & Rochet J.C. 2017. Endosulfine-alpha inhibits membrane-induced α-synuclein aggregation and protects against α-synuclein neurotoxicity. *Heritage Science* 5: 1–15.

Zhang Y.R., Zhao Y.Q. & Huang J.F. 2012. Retinoid-binding proteins: Similar protein architectures bind similar ligands via completely different ways. *PLoS ONE* 7: 1–8.

Zhou H.-X. 2012. Intrinsic disorder: signaling via highly specific but short-lived association. *Trends in Biochemical Sciences* 37: 43–48.

Zhou J., Pham H.T., Ruediger R. & Walter G. 2003. Characterization of the Aα and Aβ subunit isoforms of protein phosphatase 2A: Differences in expression, subunit interaction, and evolution. *Biochemical Journal* 369: 387–398.

# ORIGINAL PAPERS

# I

## $^{1}H$, $^{13}C$ AND $^{15}N$ NMR CHEMICAL SHIFT ASSIGNMENTS OF cAMP-REGULATED PHOSPHOPROTEIN-19 AND -16 (ARPP-19 AND ARPP-16)

by

Chandan J. Thapa, Tatu Haataja, Ulla Pentikäinen & Perttu Permi, 2020

Biomolecular NMR Assignments 14: 227–231.



https://doi.org/10.1007/s12104-020-09951-w

Biomolecular NMR Assignments
https://doi.org/10.1007/s12104-020-09951-w

ARTICLE

# $^{1}H$, $^{13}C$ and $^{15}N$ NMR chemical shift assignments of cAMP-regulated phosphoprotein-19 and -16 (ARPP-19 and ARPP-16)

**Chandan J. Thapa[1,2,3] · Tatu Haataja[1] · Ulla Pentikäinen[2,3] · Perttu Permi[1,4]**

Received: 23 March 2020 / Accepted: 21 May 2020


**Abstract**
Protein Phosphatase 2A, PP2A, the principal Serine/threonine phosphatase, has major roles in broad range of signaling pathways that include regulation of cell cycle, cell proliferation and neuronal signaling. The loss of function of PP2A is linked with many human diseases, like cancer and neurodegenerative disorders. Protein phosphatase 2A (PP2A) functions as tumor suppressor and its tumor suppressor activity is inhibited by the overexpression of PP2A inhibitor proteins in most of the cancers. ARPP-19/ARPP-16 has been identified as one of the potential PP2A inhibitor proteins. Here, we report the resonance assignment of backbone $^{1}H$, $^{13}C$ and $^{15}N$ atoms of human ARPP-19 and ARPP-16 proteins. These chemical shift values can provide valuable information for the further study of the dynamics and interaction of ARPP-proteins to PP2A using NMR spectroscopy.

**Keywords** Assignments · cAMP-regulated phosphoprotein-19 · HA-detection, intrinsically disordered protein · NMR spectroscopy

## Biological context

cAMP regulated phosphoprotein-19 (ARPP-19), and its splice variant ARPP-16, were originally identified as substrates of the cAMP dependent protein kinase in neostriatum. ARPP-16 is highly expressed in neuronal cells, whereas, ARPP-19 is expressed ubiquitously (Horiuchi et al. 1990). In neuronal cells, ARPP-19 acts as bridge between nerve growth factor and post-transcriptional regulation of neuronal gene expression and controls the neuronal development and plasticity (Irwin et al. 2002). The reduced expression of ARPP-19 is associated with neurological disorders like Alzheimer's disease and Down's syndrome (Kim et al. 2001).

ARPP-19 and ARPP-16 also plays roles in the regulation of cell cycle. Recent studies have reported the role of ARPP-19 in the development and progression of several human cancer types, such as, breast cancer (Lü et al. 2015), hepatocellular carcinoma (Song et al. 2014) and human glioma (Jiang et al. 2016). Phosphorylation of ARPP-19 by the greatwall kinase (Gwl) promotes mitotic entry and maintenance of mitotic state by inhibiting PP2A (Andrade et al. 2017; Gharbi-Ayachi et al. 2010; Song et al. 2014). ARPPs phosphorylated by the MAST3 (Microtubule Associated Ser/Thr kinase 3) kinase, a homolog of MASTL/Gwl kinase, selectively inhibits tumor suppressor PP2A holoenzyme containing B55α and B56δ (Andrade et al. 2017).

Structural level characterization of ARPPs has remained elusive; thus far there is no published structural data available for either ARPP-19 or ARPP-16. Consequently, the interaction of ARPPs with their different binding partners have only been observed at low resolution. In this paper, we report the backbone assignment of ARPP-19 and ARPP-16 as a first step towards structural level understanding of ARPPs function.

**Electronic supplementary material** The online version of this article (doi:https://doi.org/10.1007/s12104-020-09951-w) contains supplementary material, which is available to authorized users.

✉ Perttu Permi
perttu.permi@jyu.fi

[1] Department of Biological and Environmental Science, University of Jyvaskyla, Jyvaskyla, Finland

[2] Institute of Biomedicine, University of Turku, Turku, Finland

[3] Turku Bioscience, University of Turku and Åbo Akademi, Turku, Finland

[4] Department of Chemistry, Nanoscience Center, University of Jyvaskyla, Jyvaskyla, Finland

Published online: 28 May 2020

## Methods and experiments

### Recombinant protein production and purification

The human cAMP regulated phosphoproteins, ARPP-16 and ARPP-19, were overexpressed in the *Escherichia coli* strain BL21 Gold from a plasmid vector carrying the gene conferring ampicillin resistance. For the production of uniformly $^{13}C$ and $^{15}N$ labelled ARPPs, cells were grown in standard M9 minimal media supplemented with 1 g/l of $^{15}NH_4Cl$ and 2 g/l $^{13}C$-D-glucose as the only nitrogen and carbon source, respectively. The proteins were produced as Glutathione S-transferase (GST) fusion proteins. *Escherichia coli* BL21 Gold cells with ARPP-19/ARPP-16 plasmids were cultured in M9 minimal containing 100 µg/ml ampicillin at 37 °C, shaking the culture at 250 rpm until the OD at 600 nm was 0.6. The cells were cooled down to 25 °C and expression of GST fusion proteins were induced by adding 0.4 mM isopropyl β-D-1-thiogalactose at 25 °C for 20 h, shaking the culture at 250 rpm.

The cells were lysed using EmulsiFlex-C3 homogenizer (Avestin) and subsequently centrifuged at 35,000×*g* for 30 min. The lysates were purified with Protino Glutathione Agarose 4B (Macherey-Nagel) according to the manufacturer's instruction. GST was cleaved by Tobacco Etch Virus (TEV) protease (Invitrogen, Life Technologies) at 4 °C for 16 h and removed from the solution with the Glutathione Agarose. The proteins were further purified by size exclusion chromatography with a HiLoad 26/60 Superdex 200 pg column (GE Healthcare) in 50 mM $NaH_2PO_4$ pH 6.5, 100 mM KCl, 1 mM DTT using an ÄKTA pure chromatography system (GE Healthcare). Finally, the proteins were concentrated with Amicon ultra centrifugal 3K filter device (Millipore).

### NMR spectroscopy

NMR spectra of ARPP-19 and ARPP-16 were acquired at 298 K using a Bruker Avance III HD 800 MHz spectrometer equipped with a 5-mm $^{1}H$, $^{13}C$, $^{15}N$ triple resonance TCI CryoProbe. All NMR spectra were measured in 95/5% 50 mM $NaH_2PO_4$, 100 mM NaCl, 1 mM DTT pH 6.5/$D_2O$ buffer. The concentration of ARPP-19 and ARPP-16 used in the data collection were 0.7 mM and 0.4 mM, respectively. For the sequence specific backbone and partial side chain assignment of ARPPs, we used a set of HN-detected triple resonance experiments i.e. HN(CO)CACB, HNCACB (Yamazaki et al. 1994), HNCO (Muhandiram and Kay 1994), i(HCA)CO(CA)NH (Mäntylahti et al. 2009) as well as HA-detected experiments iHA(CA)NCO, HA(CA)CON (Mäntylahti et al. 2010), and HA(CA)CON(CA)HA (Mäntylahti et al. 2011). NMR data were processed using TopSpin 3.5 software package (Bruker Corporation) and analysed using Sparky 3.13 (Lee et al. 2015).

### Assignment and data deposition

The amino acid sequence of ARPP-19 and ARPP-16 are shown in Fig. 1a. The examination of the amino acid sequence shows the enrichment of hydrophilic (26.2%) and charged (33%) residues as well as prolines (8.9%), basic features of intrinsically disordered proteins (IDPs) (Hazy and Tompa 2009; Uversky 2010). The predictor of natively unfolded protein VL-XT algorithm (Romero et al. 2001) indicates both ARPPs are 50% disordered, and the Uversky plot (Uversky et al. 2000) of mean charge against mean scaled hydropathy positions ARPPs among the set of intrinsically disordered proteins (Fig. 1b). The bioinformatical analyses were experimentally confirmed in $^{1}H$–$^{15}N$ 2D HSQC spectra of ARPPs, that is, the disordered nature of ARPPs manifests itself as low dispersion of chemical shift in the $^{1}H^{N}$ dimension, all amide proton resonances falling between 7.7 and 8.5 ppm (Fig. 1c and d). Analogously, severe clustering of Cα and Cβ shifts in HNCACB/HN(CO)CACB spectra was observed, rendering backbone assignment based on Cα and Cβ shifts inefficient. We then resorted to the CO shift based assignment as a remedy i.e. employing HN-detected HNCO and its intraresidual counterpart i(HCA)CO(CA)NH experiments, and supplementing these data with HA-detected CO–N correlation experiments i.e. iHA(CA)NCO, HA(CA)CON, and (HACA)CON(CA)HA. By using the HA-detected spectra we were able to complete the backbone assignment passing the proline rich region uninterrupted. Representative illustrations of sequential walk utilizing iHA(CA)NCO and (HACA)CON(CA)HA, and iHA(CA)NCO and HA(CA)CON spectra for ARPP-16 and ARPP-19, respectively, are shown in Figs. 2 and 3. In this way, omitting N-terminal residues remaining after TEV cleavage, nearly complete backbone assignment of ARPP-19 and ARPP-16 was obtained. The $^{1}H$, $^{15}N$ and $^{13}C$ resonances of ARPP-16 and ARPP-19 have been deposited into BioMagResBank (http://www.bmrb.wisc.edu/) under accession numbers 27911 and 27912, respectively.

In ARPP-16, 99% of $^{1}H$-$^{15}N$ pairs (86 out of 87), 98% of $^{13}C\alpha$ (94 out of 96), 99% of $^{13}C\beta$ (88 out of 89), 97% of $^{1}H\alpha$ (93 out of 96) and 100% of $^{13}CO$ (96 out of 96) resonances were assigned. The $^{1}H^{N}$ of the first residue, Met1, could not be assigned due to exchange broadening. In addition, Cα and Cβ shifts of Ser30 could not be assigned, because of mutual cancellation of opposite phased Cα and Cβ cross peaks. The absence of $^{1}H\alpha$ resonances of Ser7, Arg36 and Gly40 are attributed to significant line broadening.

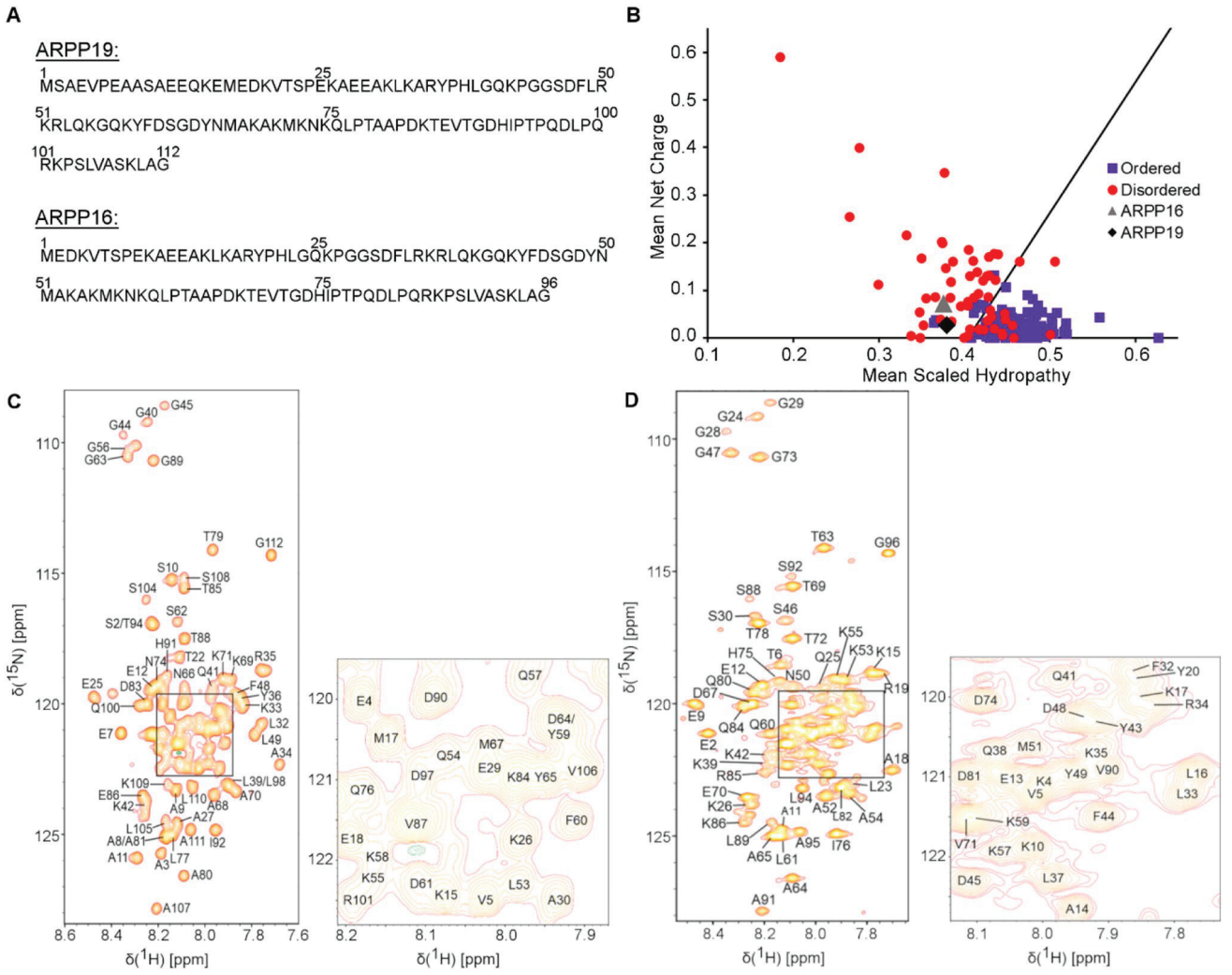


**Fig. 1** **a** The amino acid sequence of human ARPP-19 and ARPP-16. **b** Charge-hydropathy plot of human ARPP-19 and ARPP-16. The sequences of ARPP-19 and ARPP-16 were analyzed at http://www.pondr.com. The position of ARPP-19 and ARPP-16 are shown by green diamond and gray triangle, respectively, in comparison to a set of disordered (red circles) and ordered (blue squares) proteins. The border line drawn between disordered and ordered space is empirically defined by the equation $<H>_b = (<R> + 1.151)/2.785$ (Uversky et al. 2000). **c** and **d** $^{1}$H, $^{15}$N HSQC spectrum of ARPP19 and ARPP16, respectively. The NH resonances are labelled with *one-letter-amino-acid-codes* and residue numbers. Insets show the enlargement of the crowded NH regions in ARPPs

Similarly, 99% of $^{1}$H-$^{15}$N pairs (101 out of 102), 98% of $^{13}$Cα (110 out of 112), 98% of $^{13}$Cβ (103 out of 105), 96% of $^{1}$Hα (108 out of 112) and 100% of $^{13}$CO (112 out of 112) resonances were assigned for ARPP-19. $^{1}H^{N}$ of Ser46 could not be assigned probably due to fast exchange of amide proton with the solvent. Moreover, Cα and Cβ shifts of Glu13 and Lys102 could not be assigned from ARPP-19 as the resonance peaks are most likely overlapped and very weak. The missing $^{1}$Hα resonance of Gln14, which is overlapped with resonance of Lys75, prevents unambiguous assignment. The $^{1}$Hα resonances of Ser23, Arg52 and Gly56, which corresponds to the Ser7, Arg36 and Gly40 residues of ARPP16, are missing due to exchange broadening.

In this paper, we have presented a nearly complete assignment of main-chain $^{1}$H, $^{13}$C, and $^{15}$N chemical shifts in two intrinsically disordered proteins, ARPP-16 and ARPP-19. These assignments allow residue-level characterization of ARPP-16/19 dynamics and interactions in ongoing studies.

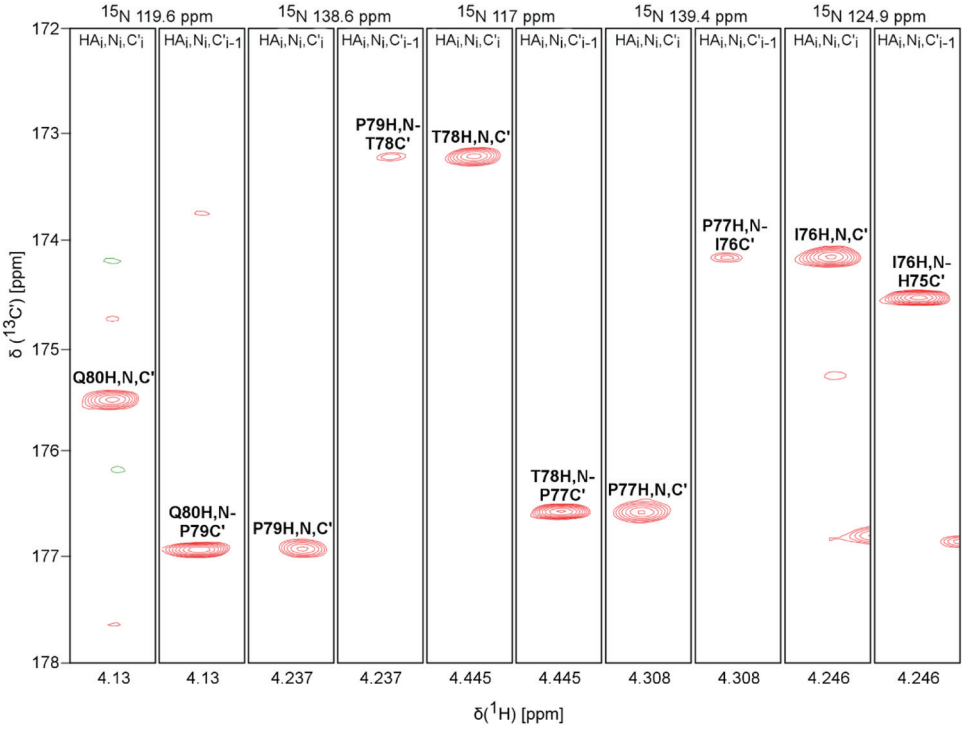


**Fig. 2** Sequential walk through prolines of ARPP-16. Assignment of the residues His75-Ile-Pro-Thr-Pro-Gln80 using iHA(CA)NCO and (HACA)CON(CA)HA spectra

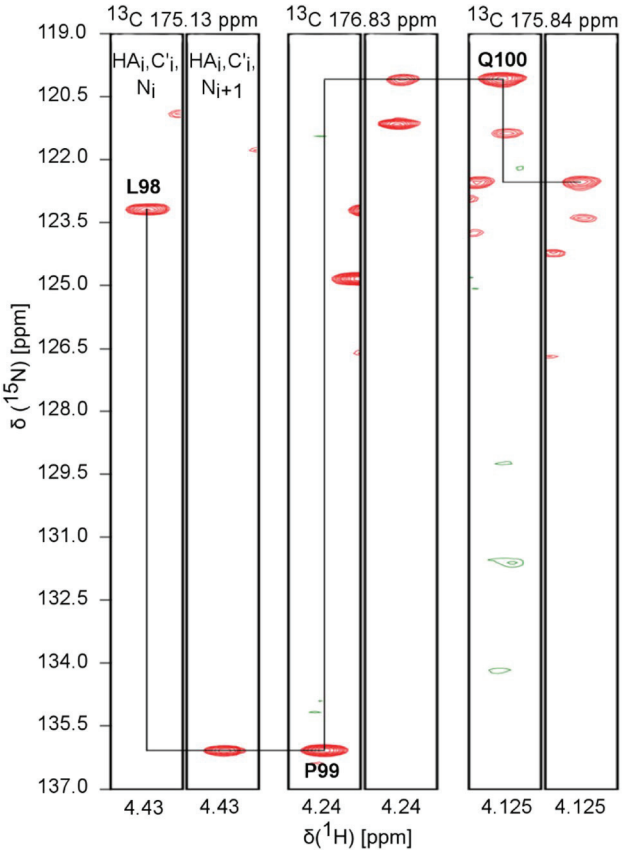


**Fig. 3** Sequential walk through proline of ARPP-19. Assignment of residues Leu98-Pro-Gln100 by employing iHA(CA)NCO and HA(CA)CON spectra

**Acknowledgements** Open access funding provided by University of Jyväskylä (JYU). This work is supported by grants from Academy of Finland (Number 288235 to PP and 28348 to UP). Chandan Thapa is a recipient of doctoral student scholarship from the University of Jyvaskyla Graduate School (JYUGS), Department of Biological and Environmental science, University of Jyvaskyla.



## References

Andrade EC, Musante V, Horiuchi A, Matsuzaki H, Brody AH, Wu T, Greengard P, Taylor JR, Nairn AC (2017) ARPP-16 Is a striatal-enriched inhibitor of protein phosphatase 2A regulated by microtubule-associated serine/threonine kinase 3 (mast 3 kinase). J Neurosci 37:2709–2722. https://doi.org/10.1523/JNEUROSCI.4559-15.2017

Gharbi-Ayachi A, Labbe J-C, Burgess A, Vigneron S, Strub JM, Brioudes E, Van-Dorsselaer A, Castro A, Lorca T (2010) The substrate of GWL ARPP19, controls Mitosis by inhibiting PP2A. Handb Cell Signal 1:1353–1365. https://doi.org/10.1016/B978-0-12-374145-5.00168-6

Hazy E, Tompa P (2009) Limitations of induced folding in molecular recognition by intrinsically disordered proteins. ChemPhysChem 10:1415–1419. https://doi.org/10.1002/cphc.200900205

Horiuchi A, Williams KR, Kurihara T, Nairn AC, Greengard P (1990) Purification and cDNA cloning of ARPP-16, a cAMP-regulated phosphoprotein enriched in Basal Ganglia, and of a related phosphoprotein, ARPP-19. J Biol Chem 265:9476–9484

Irwin N, Chao S, Nairn AC, Goritchenko L, Benowitz LI, Greengard P, Horiuchi A (2002) Nerve growth factor controls GAP-43 mRNA stability via the phosphoprotein ARPP-19. Proc. Natl. Acad. Sci. 99:12427–12431. https://doi.org/10.1073/pnas.152457399

Jiang T, Zhao B, Li X, Wan J (2016) ARPP-19 promotes proliferation and metastasis of human glioma. Neuroreport 27:960–966. https://doi.org/10.1097/WNR.0000000000000638

Kim SH, Nairn AC, Cairns N, Lubec G (2001) Decreased levels of ARPP-19 and PKA in brains of Down syndrome and Alzheimer's disease. J Neural Transm Suppl 61:263–272. https://doi.org/10.1007/978-3-7091-6262-0_21

Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Struct Bioinf 31:1325–1327. https://doi.org/10.1093/bioinformatics/btu830

Lü M, Ding K, Zhang G, Yin M, Yao G, Tian H, Lian J, Liu L, Liang M, Zhu T, Sun F (2015) MicroRNA-320a sensitizes tamoxifen-resistant breast cancer cells to tamoxifen by targeting ARPP-19 and ERRγ. Sci Rep 5:8735. https://doi.org/10.1038/srep08735

Mäntylahti S, Tossavainen H, Hellman M, Permi P (2009) An intra-residual i(HCA)CO(CA)NH experiment for the assignment of main-chain resonances in 15N, 13C labeled proteins. J Biomol NMR 45:301–310. https://doi.org/10.1007/s10858-009-9373-4

Mäntylahti S, Aitio O, Hellman M, Permi P (2010) HA-detected experiments for the backbone assignment of intrinsically disordered proteins. J Biomol NMR 47:171–181. https://doi.org/10.1007/s10858-010-9421-0

Mäntylahti S, Hellman M, Permi P (2011) Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. J Biomol NMR 49:99–109. https://doi.org/10.1007/s10858-011-9470-z

Muhandiram DR, Kay LE (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson Ser B. https://doi.org/10.1006/jmrb.1994.1032

Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK (2001) Sequence complexity of disordered protein-Romero: 2000-proteins: structure, function, and bioinformatics. Wiley Online Library 48:38–48

Song H, Pan J, Liu Y, Wen H, Wang L, Cui J, Liu Y, Hu B, Yao Z, Ji G (2014) Increased ARPP-19 expression is associated with hepatocellular carcinoma. Int J Mol Sci 16:178–192. https://doi.org/10.3390/ijms16010178

Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins Struct Funct Genet 41:415–427.

Uversky VN (2010) The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. J Biomed Biotechnol. https://doi.org/10.1155/2010/568068

Yamazaki T, Muhandiranv DR, Kay LE, Lee W, Arrowsmith CH (1994) A suite of triple resonance NMR experiments for the backbone assignment of 15N, 13C, 2H labeled proteins with high sensitivity. J Am Chem Soc 116:11655–11666. https://doi.org/10.1021/ja00105a005

# II

## INTRINSICALLY DISORDERED PP2A INHIBITOR PROTEINS', ARPP-16/19, INTERACTION MECHANISM WITH PP2A

by

Chandan Thapa, Pekka Roivas, Tatu Haataja, Ulla Pentikäinen & Perttu Permi, 2020

Submitted manuscript

Request a copy from author.

# III

# ENDOGENOUS PP2A INHIBITOR PROTEIN ENSA, STRUCTURALLY RELATED TO ARPP-19, DIFFERENTIALLY INTERACT WITH PP2A

by

Chandan Thapa, Pekka Roivas, Tatu Haataja, Ulla Pentikäinen & Perttu Permi, 2020

Manuscript

Request a copy from author.