

# **Stoichiometric and Structural Investigations of ciliopathy-related protein complex IFT-A**

## **Dissertation**

Der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)

vorgelegt von  
Yasmin Wissinger  
aus Herrenberg

Tübingen  
2016





Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der  
Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	26.01.2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Marius Ueffing
2. Berichterstatter:	Prof. Dr. Robert Feil



*“I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician: he is also a child placed before natural phenomena which impress him like fairy tale.”*

Marie Curie (1867 – 1934)



---

## Table of Content

Table of Content.....	1
Abbreviations .....	7
Summary.....	11
Zusammenfassung.....	13
1 Introduction.....	16
1.1 Cilia.....	16
1.1.1 Structure of a cilium .....	16
1.1.2 Ciliogenesis.....	17
1.1.3 Photoreceptor: a modified sensory cilium.....	18
1.1.4 Intraflagellar transport machinery .....	19
1.1.5 Intraflagellar Transport Protein Complex A - IFT-A.....	21
1.1.6 IFT-A-related ciliopathies .....	21
1.2 Mass spectrometry.....	22
1.2.1 Structural mass spectrometry of protein complexes .....	24
1.2.1.1 Absolute quantification and stoichiometry determination .....	24
1.2.1.2 Chemical crosslinking.....	24
2 Aim of the study.....	27
3 Material and Methods .....	29
3.1 Material .....	29
3.1.1 Equipment.....	29
3.1.2 Consumables and labware.....	30
3.1.3 Chemicals .....	31
3.1.4 Special reagents .....	33
3.1.5 Buffers, solutions and media .....	35
3.1.5.1 <i>E.coli</i> culture .....	35

---

3.1.5.2	Mammalian cell culture .....	35
3.1.5.3	Agarose Gels.....	37
3.1.5.4	SDS-PAGE, Coomassie staining and Western blot analysis.....	37
3.1.5.5	Silverstaining .....	39
3.1.5.6	Affinity purification.....	39
3.1.5.7	Chemical crosslinking .....	40
3.1.6	Kits .....	41
3.1.7	Enzymes.....	41
3.1.8	<i>E.coli</i> strains .....	42
3.1.9	Oligonucleotides .....	42
3.1.10	Plasmids .....	43
3.1.11	Constructs .....	43
3.1.12	Antibodies.....	44
3.1.13	Liquid chromatography and mass spectrometry.....	44
3.1.14	Software and databases .....	44
3.1.14.1	Software .....	44
3.1.14.2	Databases .....	45
3.2	Methods.....	46
3.2.1	Molecular biology.....	46
3.2.1.1	Enzymatic DNA treatment.....	46
3.2.1.2	<i>E.coli</i> culture and plating.....	47
3.2.1.3	Transformation and cryoconservation of <i>E.coli</i> .....	47
3.2.1.4	DNA isolation from <i>E.coli</i> .....	47
3.2.1.5	Polymerase chain reaction - PCR .....	48
3.2.1.6	Cloning of plasmid expression vectors.....	49
3.2.2	Mammalian cell culture .....	51
3.2.2.1	Cryopreservation and thawing of cells .....	51

---

3.2.2.2	Maintenance and growth of cells .....	51
3.2.2.3	Stable isotope labelling of amino acids in cell culture - SILAC.....	51
3.2.2.4	Transfection of mammalian cell lines.....	54
3.2.2.5	Generation and maintenance of stable expression cell lines (Flp-In).....	54
3.2.2.6	Gene editing using the CRISPR/Cas9 system.....	56
3.2.3	Protein chemistry .....	57
3.2.3.1	Protein extraction .....	57
3.2.3.2	Quantification of protein concentration .....	57
3.2.3.3	SDS-PAGE .....	58
3.2.3.4	Western blot analysis and immunostaining.....	59
3.2.4	Analysis of protein-protein interactions.....	60
3.2.4.1	Tandem Affinity Purification (TAP) .....	60
3.2.4.2	One-step FLAG affinity purification.....	62
3.2.4.3	Methanol-chloroform precipitation .....	63
3.2.4.4	In-solution tryptic proteolysis .....	63
3.2.4.5	Desalting via stop-and-go extraction tips.....	64
3.2.4.6	In-gel digestion of silver stained proteins.....	64
3.2.5	Stoichiometric and structural investigations of IFT-A .....	65
3.2.5.1	Absolute quantification of purified IFT-A.....	65
3.2.5.2	Chemical crosslinking of purified protein complex .....	68
3.2.5.3	Size exclusion chromatography of chemically cross-linked peptides .....	70
3.2.5.4	Prefractionation using 3kDa CutOff spin column filtration .....	71
3.2.6	LC-MS/MS analysis.....	71
3.2.6.1	Data-dependent analysis.....	72
3.2.6.2	Targeted mass spectrometry.....	73
3.2.7	Data processing .....	75
3.2.7.1	Data-dependent analysis.....	75

---

3.2.7.2	Targeted mass spectrometry .....	76
4	Results .....	77
4.1	Experimental workflow .....	77
4.2	Generating Flp-In monoclonal cell lines .....	78
4.2.1	Behaviour of generated Flp-In monoclonal cell lines in cell culture .....	78
4.2.2	Validation of integration .....	79
4.2.2.1	PCR-based validation of the proper integration of the transgene .....	79
4.2.2.2	Validation using western blotting in combination with immunostaining .....	80
4.3	Establishing targeted mass spectrometry for absolute quantification of IFT-A.....	81
4.3.1	Selection of representative peptides of IFT-A .....	81
4.3.2	Generation of an equimolar standard mix performing EtEP .....	82
4.3.2.1	Validating tryptic proteolysis .....	83
4.3.2.2	Quantification of standard peptides using isotopically labelled EP .....	85
4.3.3	Adjusted standard mix for absolute quantification .....	89
4.3.4	Parameter settings for targeted mass spectrometry.....	90
4.3.4.1	SRM assay performed on a triple quadrupole .....	90
4.3.4.2	PRM assay performed on an Orbitrap instrument .....	93
4.4	Absolute quantification of purified IFT-A.....	93
4.4.1	Validation of entire purification of IFT-A .....	94
4.4.1.1	One-step affinity purification of IFT-A.....	94
4.4.1.2	Silver staining and MS analysis of TAP eluates .....	95
4.4.2	Differences in complex stoichiometry of IFT-A according to chosen baits .....	96
4.4.2.1	SRM assay using an equimolar standard mix of 14 standard peptides .....	96
4.4.2.2	PRM Assay using an equimolar standard mix of 24 standard peptides .....	100
4.4.3	Reproducibility .....	103
4.5	Induced alterations in IFT-A complex stoichiometry .....	104
4.5.1	Gene editing of IFT-A components using the CRISPR/Cas9 system .....	104



---

4.5.1.1	Validation of generated CRISPR clone mixtures using PCR.....	104
4.5.1.2	Validation of CRISPR/Cas9 edited single clones via Sequencing.....	106
4.5.2	Absolute quantification of IFT-A in CRISPR/Cas9-induced mutant cell clones ..	108
4.5.3	IFT-A complex composition within fibroblasts.....	110
4.5.4	Complex composition of IFT-A during ciliary assembly and disassembly .....	112
4.6	Chemical crosslinking of purified IFT-A .....	115
4.6.1	Optimization of the crosslinking workflow .....	115
4.6.1.1	Presence of dithiotreitol (DTT) during chemical crosslinking .....	115
4.6.1.2	Concentration of the chemical crosslinker DSS.....	117
4.6.1.3	Different methods to enrich cross-linked peptides .....	118
4.6.2	Structural investigation of IFT-A using chemical crosslinking.....	122
4.6.2.1	Identified links within IFT-A after SEC .....	122
4.6.2.2	Identified links within IFT-A using different enrichment methods .....	125
5	Discussion .....	135
5.1	Stoichiometric analysis of IFT-A.....	135
5.2	Structural investigations of IFT-A.....	140
5.3	Modelling IFT-A.....	144
5.4	Perspectives .....	145
6	References .....	147
7	Annex .....	155
7.1	Additional Figures and Tables .....	155
7.2	Figure Index .....	185
7.3	Table Index .....	187
7.4	Acknowledgements / Danksagung .....	189
7.5	Curriculum vitae .....	191



## Abbreviations

<b>aa</b>	Amino acid(s)	<b>ESI</b>	Electrospray ionisation
<b>ABC</b>	Ammonium bicarbonate	<b>FA</b>	Formic acid
<b>ACN</b>	Acetonitrile	<b>FBS</b>	Fetal bovine serum
<b>AP</b>	Affinity purification	<b>FDR</b>	False discovery rate
<b>APS</b>	Ammonium persulfate	<b>FT</b>	Flowthrough
<b>AQUA</b>	Absolute quantification	<b>GOI</b>	Gene of interest
<b>ATP</b>	Adenosine 5'-triphosphate	<b>h</b>	Hour
<b>BBS</b>	Bardet Biedl Syndrome	<b>HAc</b>	Acetic acid
<b>bp</b>	Base pairs	<b>HCD</b>	Higher energy collision induced dissociation
<b>BSA</b>	Bovine serum albumin	<b>HEK293</b>	Human embryonic kidney 293 cells
<b>Cas9</b>	CRISPR associated protein 9	<b>HEK293T</b>	Human embryonic kidney 293T cells
<b>CC</b>	Connecting cilium	<b>HPLC</b>	High-performance liquid chromatography
<b>cDNA</b>	Complementary DNA	<b>HPLC H<sub>2</sub>O</b>	Water (HPLC grade)
<b>CHCl<sub>3</sub></b>	Chloroform	<b>HRP</b>	Horseradish peroxidase
<b>CID</b>	Collision induced dissociation	<b>hTERT</b>	Human telomerase reverse transcriptase
<b>CL</b>	Crosslinking	<b>IFT</b>	Intraflagellar transport
<b>CRISPR</b>	Clustered regularly interspaced palindromic repeats	<b>IFT-A</b>	Intraflagellar transport complex A
<b>(C)</b>	C-terminal	<b>IFT-B</b>	Intraflagellar transport complex B
<b>Da</b>	Dalton	<b>IP</b>	Immunoprecipitation
<b>dATP</b>	2'-deoxyadenosine 5'-triphosphate	<b>IT</b>	Ion trap
<b>ddH<sub>2</sub>O</b>	Ultra-pure water	<b>kb</b>	Kilobase
<b>dH<sub>2</sub>O</b>	Deionized water	<b>kDa</b>	Kilodalton
<b>DMEM</b>	Dulbecco's modified Eagle medium	<b>LB</b>	Luria-Bertani
<b>DMSO</b>	Dimethylsulfoxid	<b>LC</b>	Liquid chromatography
<b>DNA</b>	Deoxyribonucleic acid	<b>LC-MS/MS</b>	Liquid chromatography-tandem mass spectrometry
<b>dNTP</b>	2'-deoxynucleotide 5'-triphosphate	<b>LCA</b>	Leber congenital amaurosis
<b>dsDNA</b>	Double-stranded DNA	<b>LCA5</b>	Lebercilin
<b>DTT</b>	Dithiotreitol	<b>LFQ</b>	Label-free quantification
<b>ECL</b>	Enhanced chemiluminescence	<b>LTQ</b>	Linear trap quadrupole
<b><i>E. coli</i></b>	<i>Escherichia coli</i>	<b>MeOH</b>	Methanol
<b>EDTA</b>	Ethylendiaminetetraacetic acid	<b>min</b>	Minute(s)
<b>e.g.</b>	For example	<b>M<sub>r</sub></b>	Molecular weight
<b>EP</b>	Equalizer Peptide		
<b>EtEP</b>	Equimolarity through equalizer peptide		
<b>EtOH</b>	Ethanol		

---

MRM	Multiple reaction monitoring	TBST	TBS-Tween
MS	Mass spectrometry	TEMED	N,N,N',N'-tetramethylethylene-diamine
msec	millisecond(s)		
MS/MS	Tandem mass spectrometry	TFA	Trifluoroacetic acid
MW	Molecular weight	tMS	Targeted mass spectrometry
m/z	Mass to charge ratio	Tris	Tris(hydroxymethyl)aminomethane
<b>NaCl</b>	Sodium chloride	<b>U</b>	Unit (enzymatic activity)
NHS ester	N-hydroxysuccinimide ester	UV/Vis	Ultraviolet/visible
NP-40	Nonidet P-40		
(N)	N-terminal	v/v	Volume per volume
<b>OD</b>	Optical density	<b>WB</b>	Washing buffer
		wt	Wildtype
<b>p.a.</b>	Pro analysis (reagent-grade)	w/v	Weight per volume
PAGE	Polyacrylamide gel electrophoresis		
PBS	Phosphate-buffered saline		
PCR	polymerase chain reaction		
PEI	Polyethylenimine		
POI	Protein of interest		
PPI	Protein-protein interaction		
PRM	Parallel reaction monitoring		
PVDF	Polyvinylidene difluoride		
<b>RNA</b>	Ribonucleic acid		
RPE	Retinal pigmentary epithelium		
rpm	Round per minute		
RT	Room temperature		
<b>SDS</b>	Sodium dodecyl sulfate		
SDS-PAGE	SDS-polyacrylamide gel electrophoresis		
SF-TAP	Strep-FLAG-tandem affinity purification		
sec	Second(s)		
SILAC	Stable isotope labelling in cell culture		
siRNA	Short interfering RNA		
SRM	Selected reaction monitoring		
ssDNA	Single stranded DNA		
StageTips	Stop-and-go extraction tips		
<b>TAP</b>	Tandem affinity purification		
TBS	Tris-buffered saline		

**Amino acids:**

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V



---

## Summary

Cilia are evolutionary conserved organelles which protrude from almost every polarized eukaryotic cell. These hair-like organelles are vital for human and animal development and physiology being associated with cell cycle and proliferation. Depending on its molecular structure cilia can be classified in motile and immotile cilia. Immotile cilia, also called primary cilia, are characterized by the lack of a central microtubule doublet and generally serve as sensory organelles. Based on the fact that cilia are present in cells and organs of the human body, malfunction of ciliary proteins and defects in cilia lead to a wide range of human diseases which are summarized under the term: ciliopathies. To understand the underlying mechanisms of ciliopathies, it is crucial to study processes within cilia. So far, no biosynthesis machinery is described within a cilium. To transport proteins from the cytoplasm to the tip of a cilium, which is essential for the ciliary assembly and its maintenance, the bidirectional intraflagellar transport (IFT) is necessary. This transport mechanism is driven by motor proteins and two multiprotein complexes IFT-A and IFT-B. IFT-A, consisting of six known complex components, is involved in retrograde transport of protein cargo from the ciliary tip back to the cell body. As described in previous studies, malfunction of IFT-A proteins leads to an accumulation of IFT-B particles in the ciliary tip which results in shortened and bulged cilia. Additionally, mutations within genes encoding IFT-A proteins are described to cause ciliopathies like Sensenbrenner syndrome which is characterized by ectodermal as well as skeletal anomalies. The presented study aims to investigate stoichiometric and structural properties of IFT-A, which is essential for the molecular function of IFT-A during intraflagellar transport and is assistant to unveil its role in IFT-A-related diseases.

As basic prerequisite of this study, Flp-In monoclonal cell lines stably expressing N-terminally Strep-Flag (SF)-tagged baits were generated to circumvent the influence of artificial overexpression on stoichiometry of the protein complex. Three different baits were chosen: IFT122, an integral part of the IFT-A, TULP3 which is described to be associated with the IFT-A and LCA5 which represents an rather labile and transiently bound interaction partner of IFT-A. Using an integral part of the protein complex of interest (Flp-In (N)-SF-IFT122) led to a drastic change in complex composition. Due to the higher affinity of TULP3 to IFT-A compared to LCA5, higher amount of the IFT-A complex could be purified from Flp-In (N)-SF-TULP3 enabling a more robust determination of the stoichiometric and structural investigation.

To determine complex stoichiometry performing absolute quantification, the establishment of a targeted mass spectrometry approach is crucial. Depending on the applied mass spectrometer, two different approaches were set up: Selected Reaction Monitoring (SRM) and Parallel

---

Reaction Monitoring (PRM). Another important step for the absolute quantification is the generation of a standard mix containing known amounts of representative peptides for the proteins of interest. To create an economic equimolar standard mix, the already described “Equimolarity through Equalizer Peptide” (EtEP) method was used. At the end, absolute quantification performing PRM on a Q-Exactive mass spectrometer in combination with an equimolar standard mixture was used to study complex stoichiometry of IFT-A. Data analysis was performed using the software Skyline. This study unveiled naturally occurring compositions of IFT-A which change during different stages of ciliogenesis and cilia disassembly. To investigate the impact of disease causing variants on the composition of IFT-A, CRISPR/Cas9 system was used to generate targeted mutations within genes encoding two IFT-A proteins (IFT43 and WDR35) in Flp-In (N)-SF-TULP3 cells. One of the generated monoclonal cell lines, carrying a mutation in the gene encoding IFT43 (c.541\_542insA/p.T181Nfs\*2), showed significant changes within IFT-A complex composition that could explain the malfunction.

The second part of this study aims to determine binding sites within IFT-A by chemical crosslinking in combination with mass spectrometry that serve as basis for structural models of IFT-A. For chemical crosslinking of purified IFT-A, the homobifunctional amine-reactive crosslinker disuccinimidyl suberate (DSS) was used. This crosslinker contains two identical N-hydroxysuccinimidyl groups which enable the formation of stable amide bonds with primary amines of N-termini and lysine residues of proteins. Based on a defined spacer length of DSS (11Å), distance information of cross-linked peptides can be achieved. To enrich low abundant cross-linked peptides before LC-MS/MS analysis, size exclusion chromatography (SEC) is a common used prefractionation method. However, SEC is time-, labour- and cost-intense. A facilitated method to reduce sample complexity prior to the analysis of cross-linked peptides is portrayed in this study. This economic method is based on preseparation using 3kDa CutOff spin columns. For the identification of linked peptides the computational software pipeline xQuest/xProphet was used. To illustrate identified links, free available software Xinet was used afterwards. Identified crosslinks as well as cross-linked positions within a protein sequence using both preseparation methods were comparable. Applying chemical crosslinking to purified IFT-A, well-known interaction domains within the complex components were confirmed. Furthermore, new crosslinking hotspots which are not described so far were identified in this study.

With the data obtained in this study, the foundation of a structural model of the IFT-A can be generated by combining the determined complex stoichiometry with obtained structural information of the protein complex IFT-A.



---

## Zusammenfassung

Zilien sind evolutionär konservierte Zellfortsätze, die auf fast allen polarisierten, eukaryotischen Zellen vorhanden sind. Diese sogenannten „Antennen der Zelle“ sind unverzichtbar für die Entwicklung von Mensch und Tier, da sie zum einen wichtig für die Zellfortbewegung sind aber auch Signale aus der Umwelt aufnehmen können. Basierend auf ihrer molekularen Struktur können Zilien in zwei Subtypen eingeteilt werden: bewegliche und unbewegliche Zilien. Während bewegliche Zilien ein zentrales Mikrotubuli-Dublett besitzen, fehlt dieses in unbeweglichen Zilien, die auch primäre Zilien genannt werden. Diese primären Zilien dienen in der Regel als Mechano- und Chemosensoren. Aufgrund des Vorkommens von Zilien im gesamten menschlichen Körper führen defekte Zilien und Fehlfunktionen in involvierten Proteinen zu unterschiedlichsten Krankheitsbildern wie z.B. zystische Nieren und Netzhautdegeneration. Diese Krankheiten sind unter dem Begriff Ziliopathien zusammengefasst. Um den Mechanismus, der Ziliopathien zu Grunde liegt zu verstehen, müssen die Prozesse innerhalb eines Ziliums untersucht werden. Bisher wurde keine Struktur innerhalb eines Ziliums beschrieben, welche Proteine und andere Biomoleküle, die für die Herstellung und Instandhaltung von Zilien notwendig sind, bereitstellt. Allerdings ist der Intraflagellare Transportmechanismus (IFT) bekannt, der mit Hilfe von Motorproteinen und den zwei Proteinkomplexen IFT-A und IFT-B essentielle Proteine von dem Zytoplasma in die Spitze des Ziliums und wieder zurück transportiert. IFT-A besteht aus sechs Komponenten und ist für den retrograden Transport (vom Zilium zurück zur Zelle) wichtig. Vorangegangene Studien beschreiben eine Akkumulation von IFT-B Partikeln in der Spitze des Ziliums aufgrund eines Defektes in dem IFT-A Komplex. Dies führt zu verkürzten und gewölbten Zilien, welche zum Beispiel zu Deformationen im Skelett führen können. Solche Anomalien sind in unterschiedlichen Ziliopathien beschrieben, wie z.B. Sensenbrenner Syndrom. Um den Einfluss von IFT-A auf Ziliopathien zu untersuchen, soll in dieser Studie die Stöchiometrie des Proteinkomplexes in der natürlich vorkommenden Zusammensetzung als auch die Struktur von IFT-A untersucht werden.

Grundlegend für diese Arbeit war die Generierung von stabilen Flp-In Zelllinien, um keinen Einfluss auf die natürlich vorkommende Zusammensetzung von IFT-A durch artifizielle Überexpression zu nehmen. Für die Generierung wurden drei Ankerproteine, sogenannte *Baits* ausgewählt, die in unterschiedlicher Weise mit IFT-A in Zusammenhang stehen: IFT122 ist ein Teil des IFT-A Komplexes, ein assoziiertes Protein TULP3 und LCA5, welches ein labiler und transienter Interaktionspartner von IFT-A ist. Der Einsatz eines internen Proteins als *Bait* (wie z.B. IFT122) führt zu einer deutlichen Veränderung der Komplexzusammensetzung. Im

Vergleich zu LCA5 besitzt TULP3 eine höhere Affinität zum IFT-A Komplex. Dies führt zu einer größeren Menge an aufgereinigtem IFT-A bei der Verwendung von Flp-In (N)-SF-TULP3, weswegen diese monoclonalen Zellen für die folgenden Experimente verwendet wurden.

Um die Stöchiometrie eines Proteinkomplexes zu bestimmen, ist eine absolute Quantifizierung der einzelnen Komponenten des Proteinkomplexes notwendig. Für die absolute Quantifizierung mittels Massenspektrometrie wurden zwei unterschiedliche Methoden verwendet: *Selected Reaction Monitoring (SRM)* und *Parallel Reaction Monitoring (PRM)*. Ein weiterer wichtiger Punkt für eine Absolute Quantifizierung ist die Auswahl von geeigneten und repräsentativen Standardpeptiden. Um einen kostengünstigen Standardmix mit equimolaren Mengen der ausgewählten, proteotypischen Peptide zu generieren wurde die *Equimolarity through Equalizer Peptide (EtEP)* Methode genutzt. Für die Bestimmung der Zusammensetzung von IFT-A wurde PRM auf einem QExactive Plus Massenspektrometer in Kombination mit einem equimolaren Standardmix zur Methode der Wahl. Die Datenanalyse erfolgte mit der freiverfügbaren Software Skyline. Diese Studie ermittelte die Stöchiometrie von natürlich vorkommendem IFT-A, welche während des Auf- und Abbau eines Ziliums variiert. Zusätzlich wurden durch CRISPR/Cas9 ausgewählte IFT-A Proteine (IFT43 und WDR35) in Flp-In (N)-SF-TULP3 Zellen gezielt mutiert, um die Auswirkung von krankheitsassoziierten Mutationen auf die Stöchiometrie des IFT-A zu beschreiben.

Der zweite Teil dieser Arbeit beschäftigt sich mit der Identifizierung von Bindestellen innerhalb des IFT-A mittels chemischen Crosslinking. Die Kombination von chemischem Crosslinking mit Massenspektrometrie legt den Grundstein zur Strukturaufklärung und Generierung eines Strukturmodells von IFT-A. Aufgereinigtes IFT-A wurde mit dem homobifunktionalem Crosslinker Disuccinimidylsuberat (DSS) inkubiert, um primäre Amine in N-Termini als auch in Lysin innerhalb der Proteinsequenzen zu vernetzen. Dafür ist die Amin-Reaktivität der verwendeten N-hydroxysuccinimidyl-Gruppen verantwortlich. Durch die definierte Länge des eingesetzten Crosslinkers (11Å) können Nachbarschaftverhältnisse über vernetzte Peptide identifiziert werden. Für die Identifizierung der vernetzten Peptide mittels Massenspektrometrie ist eine vorangehende Anreicherung dieser Peptide notwendig. *Size exclusion chromatography (SEC)* ist in diesem Zusammenhang eine gern genutzte Methode, die allerdings sehr kostenintensiv, zeit- und arbeitsaufwendig ist. Deswegen ist in dieser Arbeit eine vereinfachte Methode zur Reduktion der Probenkomplexität mittels *3kDa CutOff spin columns* parallel zur SEC verwendet worden und beschrieben. Für die Identifizierung der vernetzten Peptide wurde die Software-Pipeline xQuest/xProphet verwendet. Die anschließende Visualisierung der Daten wurde mittels Xinet durchgeführt. Beide Anreicherungsmethoden erzielten vergleichbare Ergebnisse in der Art und der Positionen der identifizierten Crosslinks. Zusätzlich haben beide Methoden bisher

beschriebene Interaktions-Domänen bestätigt als auch neue, bisher unbeschriebene *Crosslink-Hotspots* identifiziert.

Mit Hilfe dieser Studie konnten sowohl bekannte Interaktionsdomänen innerhalb des IFT-A verifiziert, als auch neue unbekannte Bindestellen identifiziert werden. In Kombination mit der ermittelten Stöchiometrie des IFT-A legt diese Arbeit den Grundstein zur Generierung eines Strukturmodells dieses Proteinkomplexes, welches für die Aufklärung des Einflusses von IFT-A in Ziliopathien unabdingbar ist.

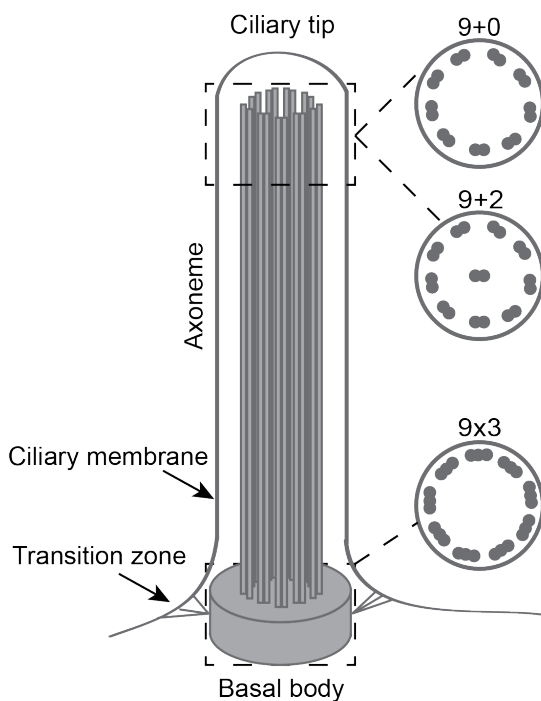
# 1 Introduction

## 1.1 Cilia

Cilia are complex and evolutionary conserved organelles which protrude from the surface of almost every polarized eukaryotic cell [1]. Based on its molecular structure, cilia are classified in motile and immotile cilia. Immotile, or primary cilia, serve as sensing organelles [1] and are specialized as cellular antennae for sensing the environment. They are involved in many different functions, for example in renal development as well as in neurosensory functions like hearing, smell and vision [1, 2]. The increase in surface area of a cell by building up a cilium is ideal to sense its extracellular environment [1, 3, 4].

### 1.1.1 Structure of a cilium

The ciliary structure can be divided into different compartments: basal body, transition zone, axoneme, ciliary membrane and ciliary tip.



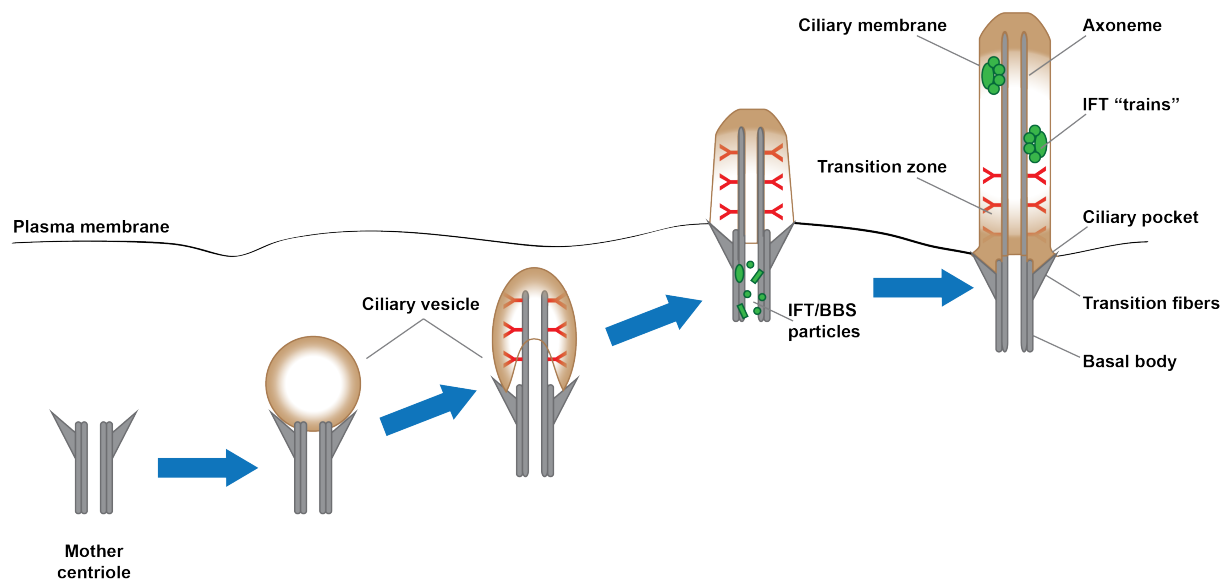
**Figure 1: Structure of a cilium**

A cilium is composed of different compartments: Basal body, transition zone, ciliary membrane, axoneme and ciliary tip. Different types of cilia, '9+0' for primary cilia and '9+2' for motile cilia, are shown within the cross-sections. The axoneme of each cilium is built by nine microtubule doublets with an additional centered microtubule doublet in case of motile cilia ('9+2'). The basal body of a cilium consists of nine microtubule triplets ('9x3') in comparison to the doublets present in the axoneme. Y-shaped transition fibres connect the basal body, the basis of each cilium, to the cilia membrane. Modified after [5].

As illustrated in Figure 1, cilia arise from the basal body which originates from the mother centriole and represents a centriolar barrel [5]. Nearby this initiation point, docking sites for intra flagellar transport (IFT) particles are present. These are involved in the transport of protein cargo along the axoneme of a cilium, described in 1.1.4 [6, 7]. The transition zone is characterized by Y-shaped transition fibres which link the axoneme and the basal body to the ciliary membrane. This specialized region is thought to act as a barrier where only certain proteins may pass to enter the ciliary compartment [8]. The microtubule core of a cilium (axoneme) is composed of nine microtubule doublets, formed by heterodimers of tubulin A and B [9]. Motile cilia consist of nine outer and two central microtubules ('9+2' array of microtubules) in comparison to the '9+0' array of microtubules in primary cilia which are in most cases immotile [10-12]. In contrast to the axoneme, the basal body consists of nine microtubular triplets ('9x3' array of microtubules). The two central microtubules of motile cilia combined with outer and inner dynein arms are necessary for directional ciliary movement [1, 13].

### 1.1.2 Ciliogenesis

The generation of cilia, a process called ciliogenesis, proceeds in different stages [3].



**Figure 2: Ciliogenesis**

Ciliogenesis, the process of generating a cilium, takes place in different steps. After the attachment of a ciliary vesicle to the distal end of a mother centriole, the axoneme is built up by microtubules. The elongation of the axoneme starts at its tip and the ciliary vesicle grows with it forming a ciliary sheath. This structure migrates to the plasma membrane where the ciliary sheath fuses to externalise the cilium. For the completion of a functional cilium, IFT as well as BBS proteins are necessary.

As depicted in Figure 2, ciliogenesis in mammalian cells is initiated by the attachment of a mother centriole to the cell surface where it fuses with a ciliary vesicle. Subsequently, the

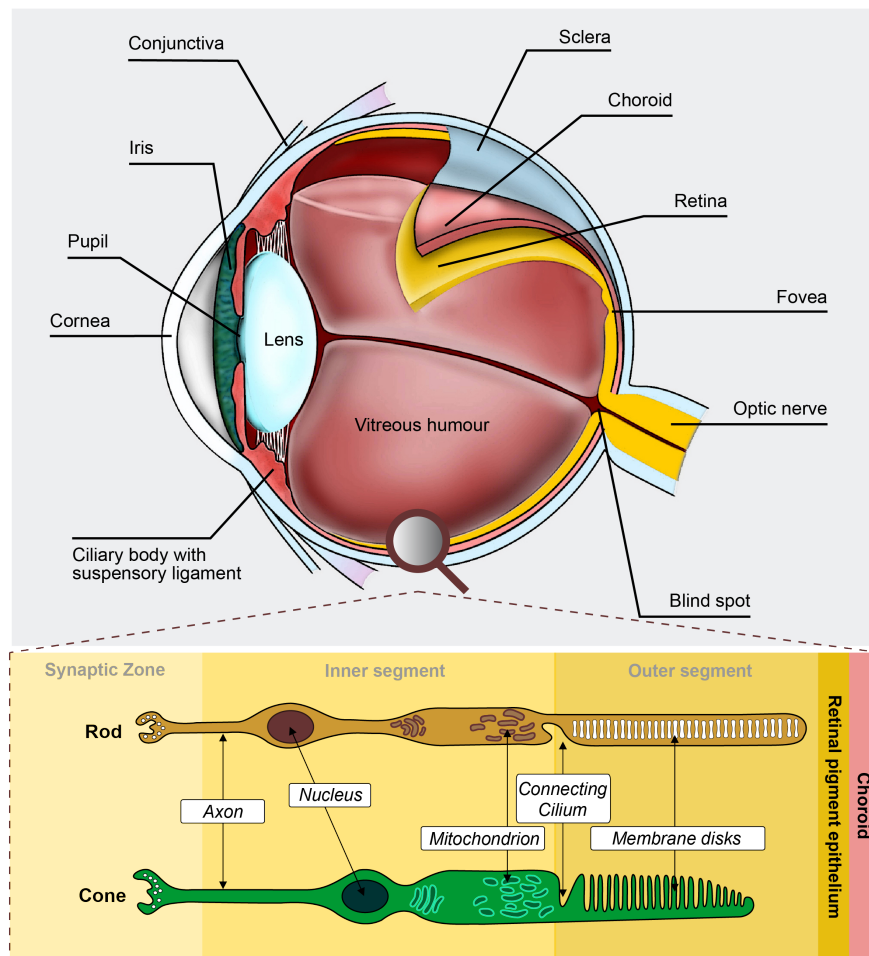
mother centriole differentiates into a basal body which is anchored to the plasma membrane. Through nucleation of microtubules the axoneme is built up and the transition zone is generated. During the elongation of the axoneme the ciliary vesicle fuses with a second vesicle forming the ciliary membrane. While the axoneme consists of doublet microtubules the basal body contains triplet microtubules [3, 13-15].

Ciliogenesis is characterized by a variety of complex processes like the synthesis of macromolecules as well as assembly. Therefore, ciliogenesis must be regulated carefully. In dividing cells, ciliary disassembly has to take place before cell division because the basal body is crucial for the organization of the centrosome for the attachment of the spindle apparatus [14, 16, 17]. As shown by Avasthi and Marshall [3], tubulin deacetylation, ciliary protein methylation as well as ubiquitination are involved in this disassembly of cilia. Centrioles are duplicated and transmitted to the daughter cells where they act as templates for the next generation of cilia. Other regulation mechanisms for ciliogenesis are cell confluence, fluid flow and cell spreading. To control ciliary length, a steady-state assembly is compensated by the permanent removal of microtubular subunits from the tip of cilia. Conversely, this process requires anterograde IFT to provide new axonemal subunits [1].

### **1.1.3 Photoreceptor: a modified sensory cilium**

One specific type of cilia is represented by photoreceptors, present in the retina. The retina is a light-sensitive tissue located at the back of the eye and represents an important part of the central nervous system. It is mainly formed by six types of neurons and one type of glia cells. The outermost neuronal cell layer of the retina is formed by photoreceptors which are light-sensitive, transduce the light stimulus into an electrical signal and transmit the signal to the inner retinal neurons and through the optic nerve to the brain. There are two different types of photoreceptors present in the human retina: rods and cones. Rod photoreceptors are sensitive to dim light, whereas cones are specialized for high acuity daytime vision [18, 19]. Rods and cones have a distinct structure in common. They consist of an inner- and a light-sensitive outer segment [13, 20]. As depicted in Figure 3, the outer segment of photoreceptor cells represents a modified sensory cilium which originates from the basal body. The inner segment is linked to the outer segment of a photoreceptor through the innermost part of the axoneme, which is important as a structural backbone of the membrane disks. The connecting cilium represents the transition zone of this specialized cilia [21] and it plays a central role in photoreceptor function because it connects the metabolic machinery of the cell (inner segment) to the light-sensitive outer segment [22]. The disks of the outer segment, unique to photoreceptors, create a maximally enlarged membrane surface area to harbour a large number of visual pigments. This maximises the visual efficiency. Considering that about 10% of all outer segment disks are

renewed every day, the maintenance of outer segments requires a precise control and an extremely high rate of ciliary transport.



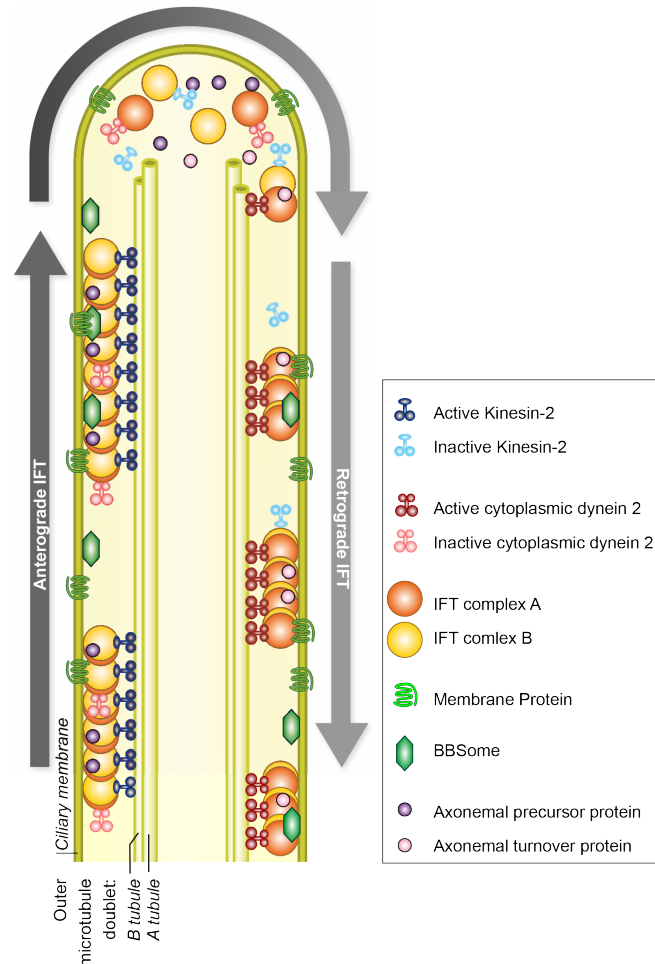
**Figure 3: Structure of the human eye with photoreceptors**

The human eye is a complex optical system. It detects light and converts these physical stimuli into electro-chemical signals which are transmitted through the optical nerve into the visual cortex in the brain. The retina consists of different layers of neurons. The only light-sensitive neurons present are the photoreceptors. There are two types of photoreceptors: rods and cones. While rods provide black and white vision, cones are essential for daytime vision and are sensitive for colour. The connecting cilium, located in the photoreceptor cells, connects the metabolic machinery (inner segment) with the light-sensitive outer segment. The extensive surface area of the membrane disks leads to a maximum of photon capture and consequently in an effective visual transduction.

#### 1.1.4 Intraflagellar transport machinery

Each particle which is essential for the assembly and the maintenance of cilia has to be delivered to the cilium because cilia do not contain any known biosynthesis machinery. Until now, little is known about the mechanisms how ciliary cargo is specifically targeted to the cilium. However, there is some evidence that the biogenesis of cilia is depending on vesicular and intraflagellar transport (IFT). Therefore, the fusion of carrier vesicles with the periciliary

membrane allows the ciliary cargo to be delivered [23-26]. For the bidirectional transport along the axoneme of a cilium, the elaborated and evolutionarily conserved IFT machinery is responsible (see Figure 4).



#### Figure 4: Intraflagellar transport

Intraflagellar transport (IFT) denotes the bidirectional movement along the axoneme of cilia, essential for its assembly and maintenance. Each protein involved in ciliary assembly is transported from the cell to the tip of a cilium. This anterograde IFT is driven by the subcomplex IFT-B in combination with kinesin motor proteins. Responsible for the delivery of protein cargo back from the ciliary tip to the basal body (anterograde) is the protein complex IFT-A along with dynein. Modified after [1].

This IFT is driven by kinesin-2 and dynein motor proteins [27] and the cargo link is established by two distinct multiprotein complexes: IFT-A and IFT-B. IFT-B, consisting of 17 proteins and kinesin are important for the anterograde transport from the cell to the ciliary tip and are involved in the assembly and maintenance of cilia. In contrast, IFT-A which contains six known proteins, together with cytoplasmic dynein are essential for the retrograde transport of protein cargo from the ciliary tip back to the cell body [1, 13, 28]. To enable cargo binding, IFT proteins contain numerous protein-protein interaction motifs [3]. Whereas the molecular composition of



---

the IFT machinery is known, the mechanisms of ciliary assembly as well as the interaction of IFT proteins with correct protein cargo is still unknown, although there is some evidence that there are specific adaptors that link defined cargos to the IFT system [1].

### **1.1.5 Intraflagellar Transport Protein Complex A - IFT-A**

The IFT complex A (IFT-A), which is in the centre of this study, represents one of the two subcomplexes of the IFT machinery. IFT-A comprises six complex components IFT122, IFT140, IFT43, TTC21B, WDR19 and WDR35 and it was shown that three of them, WDR19, IFT122 and IFT140, are forming a core complex [29]. The whole protein complex IFT-A is essential for the retrograde transport of protein cargo from the ciliary tip to the cell body, a process that seems to be crucial for the protein turnover in the cell body but not necessarily of primary importance for ciliary assembly [1]. Due to their importance in axonemal transport, mutations in genes encoding IFT-A proteins may cause ciliopathies as described in more detail in 1.1.6.

Tubby-like protein 3 (TULP3) has been described to interact with IFT-A and thereby promotes the ciliary localization of a subset of G protein-coupled receptors (GPCRs) [1]. Due to a conserved domain, located in the amino terminus, some proteins of the tubby family are able to bind to the core complex of IFT-A [29]. Based on this physical association of TULP3 to IFT-A, TULP3 can be used as bait for affinity purification (AP) of IFT-A and its interaction partners.

### **1.1.6 IFT-A-related ciliopathies**

Cilia are present on almost any polarized cell of the human body. Therefore, defects in cilia and malfunction of ciliary proteins lead to a wide range of human diseases affecting many organ systems with multiple symptoms, including polycystic kidney disease, hydrocephalus and retinal degeneration [1, 2, 6]. According to their biological function, mutation-induced inactivation of IFT-B or kinesins leads to a defective biogenesis of cilia, whereas inactivation of IFT-A or cytoplasmic dynein results in defective ciliary function [13]. Due to its diverse occurrence, any defect in cilia may cause a wide range of diseases such as Bardet-Biedl syndrome (BBS), Joubert syndrome (JBS) and retinitis pigmentosa (RP). These diseases are summarized under the term ciliopathies [2, 30].

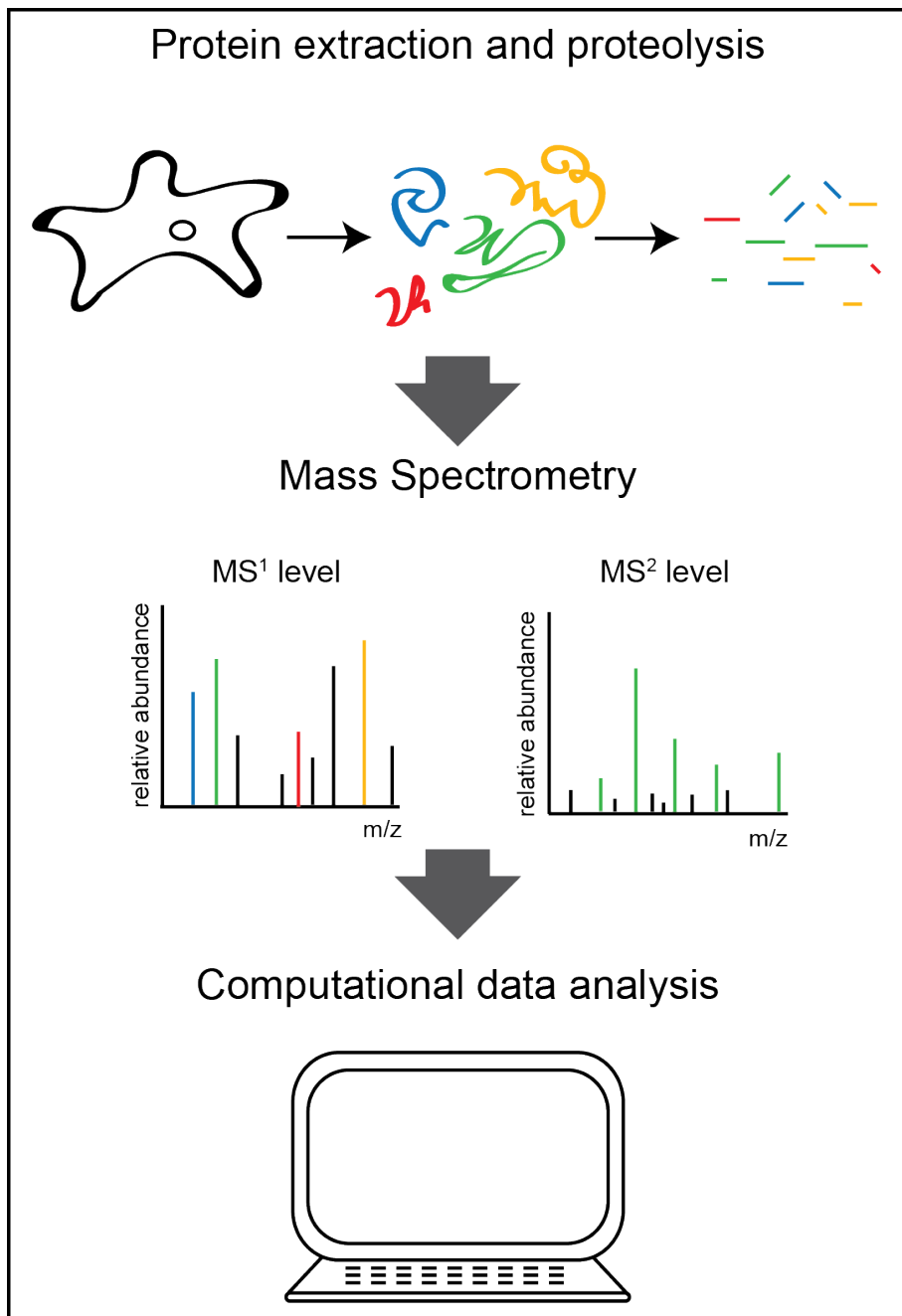
Lack or malfunction of IFT-A proteins results in cilia deformation. Consequently, mutations in genes encoding IFT-A complex components cause different ciliopathies. Due to its involvement in retrograde ciliary transport, malfunction of IFT-A complex components leads to an accumulation of IFT-B particles in the ciliary tip and to the formation of shortened and bulged cilia. Based on this effect and in combination with the requirement of IFT-A proteins for mammalian sonic hedgehog (shh) signalling, mutations in genes encoding IFT-A proteins are a major cause of skeletal ciliopathies [27, 31-38]. Sensenbrenner syndrome for example, is

characterized by skeletal and ectodermal anomalies as well as chronic renal failure and liver fibrosis [39]. Another IFT-A-related ciliopathy is the rare autosomal recessive Mainzer-Saldino syndrome (MSS). This disease is caused by mutations in IFT140 and is characterized by phalangeal cone-shaped epiphyses, chronic renal failure and early-onset retinal dystrophy [40]. To ultimately understand the underlying mechanisms of these diseases it is crucial to determine the exact function of the involved protein complexes. A first step towards this is to determine the precise composition and the stoichiometry of the protein complex IFT-A.

## 1.2 Mass spectrometry

Mass spectrometry (MS) enables the determination of the mass of molecules within a sample. The mass spectrometer analyses ions under vacuum and it consists of three devices: ion source, mass analyser and detector. Ions are generated in and sorted via their mass to charge ratio ( $m/z$ ) before they are detected [41]. The first application of mass spectrometry goes back to the early 20<sup>th</sup> century with steady progress for scientific applications, thenceforth [42, 43]. In 1960's the analysis of small volatile molecules performing mass spectrometry was achievable. Thereby, molecules were separated using gas chromatography. Separated fractions were bombarded with electrons to ionize the small molecules before analysis using mass spectrometry takes place. But this approach was not adaptable to larger non-volatile biomolecules. In the 1980's other ion sources were developed which enabled the ionization of larger biomolecules for example matrix-assisted laser-desorption ionization (MALDI) or electrospray ionization (ESI) [44]. Nowadays, mass spectrometry is a central analytical technique for proteomic strategies and almost replaced other proteomic methods like 2D-gel electrophoresis. The quest to identify and quantify proteins and their interactions in complex samples resulted in many experimental strategies and mass spectrometric methods [45]. Advanced software tools and databases facilitate the analysis and interpretation of complex mass spectra. Although many applications require a steady increase in sensitivity and resolution of used mass spectrometer instruments, this also leads to the disadvantage of a simultaneous increase in background which leads to the identification of non-specific binding partners. There is a wide diversity in mass spectrometric based methods, however the main steps are the same. In often performed bottom-up approaches, proteins are fragmented into smaller peptides using proteases. The resulting peptide mixture is separated by liquid chromatography (LC) before peptide fractions are analysed by MS. The LC system can be directly coupled to the MS using electrospray ionization (ESI) source. Based on an initial full scan or on generated mass lists, precursor ions can be chosen for further fragmentation. The analysis of these fragments enables the detection of the amino acid sequence of the analysed peptide. Using different

software and databases the resulting data are assigned to the corresponding proteins for further evaluation.



**Figure 5: Workflow of a bottom-up mass spectrometry approach**

Samples for mass spectrometry approaches can be prepared from cell lines as well as from various tissues with different methods. Extracted proteins are enzymatically cleaved into smaller peptides using a protease. Peptides are separated by liquid chromatography (LC) which, in case of electrospray ionization, is directly coupled to the mass spectrometer. After peptides entered the mass spectrometer, a precursor scan (MS<sup>1</sup> level) is performed and chosen precursors are further fragmented and sequenced (MS<sup>2</sup> level). For data analysis, different computational tools tailored to the used approach can be used.

### **1.2.1 Structural mass spectrometry of protein complexes**

Malfunction of proteins, involved in critical cellular processes, may lead to disease. To understand the underlying molecular mechanism of diseases, such as ciliopathies, the knowledge about all involved proteins and especially how these are assembled into protein complexes as well as the complex stoichiometry is essential. To determine stoichiometry of involved proteins using mass spectrometry, absolute quantification is the method of choice.

#### **1.2.1.1 Absolute quantification and stoichiometry determination**

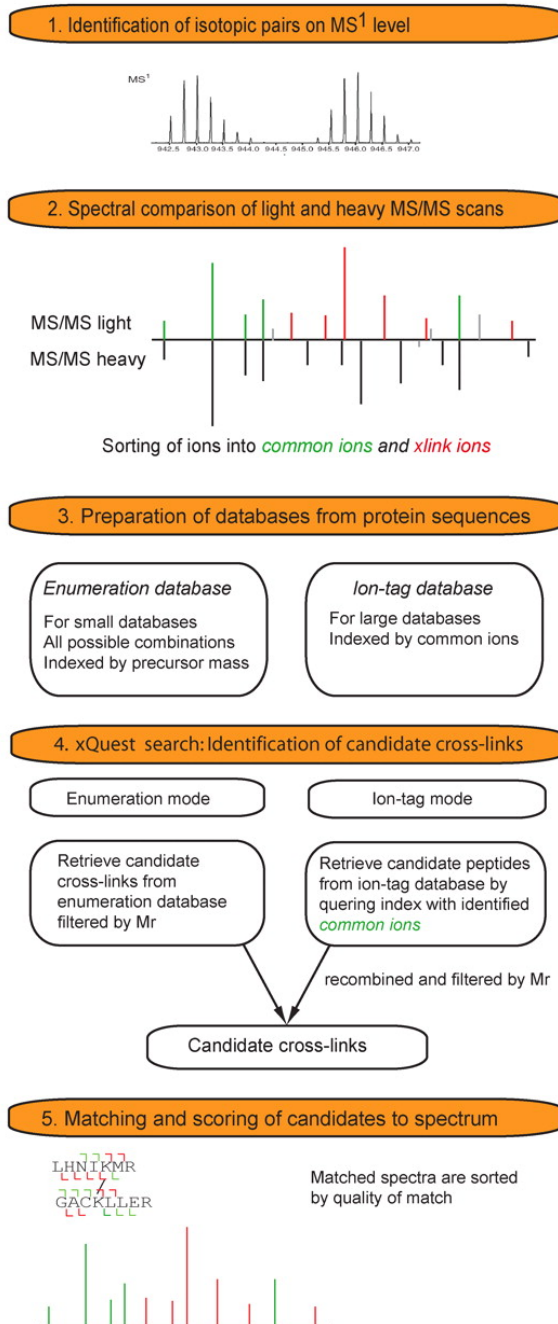
To determine the stoichiometry of a protein complex by mass spectrometry, the absolute quantification of representative peptides is crucial. Different methods are already described based on the addition of a known amount of standard proteins or peptides into the biological sample to be analysed. Thereby, it makes no difference whether the spiked-in representative peptides are synthesized with incorporated stable isotopes (AQUA peptides [46]) or the proteins within the biological sample are isotopically labelled for example using SILAC media. While targeted mass spectrometry is crucial for an absolute quantification of low and high abundant proteins, other quantitative proteomics approaches, for example 2D SDS-PAGE where gel spots were quantified using staining methods, are only useful for high abundant proteins [47]. Absolute quantification of peptides is a difficult but highly precise method to determine the composition and to calculate the stoichiometry of a protein complex of interest. Due to variability in ionization efficiency of selected peptides during electro spray ionization (ESI) and the following ion transfer and detection within the mass spectrometer, absolute quantification represents a challenging method for the analysis of protein complex stoichiometry [48]. The determined stoichiometry can be the basis for the determination of quantitative effects introduced through alterations in cellular state or by disease-associated mutations within a protein complex. The procedure of an absolute quantification approach can be separated into different steps, starting with the selection of representative peptides performing data-dependent shotgun analysis and ending with the quantification of each representative peptide using a defined amount of spiked-in standard peptide (described in more detail in chapter 3.2.5.1) [49].

#### **1.2.1.2 Chemical crosslinking**

Protein complexes play a key role in cellular function. To understand the molecular mechanisms of cellular function, it is important to combine the information of protein-protein interaction studies and the knowledge of the comprehensive composition of protein complexes with structural information. Chemical crosslinking with subsequent LC-MS/MS analysis is a useful method to better characterize the structure of proteins and of protein complexes [50]. It represents a complementary approach to well-established high-resolution techniques like X-ray

crystallography, NMR spectroscopy and electron microscopy [51]. Chemical crosslinking combines the opportunity to investigate protein conformation and information about protein-protein interactions with an additional benefit due to spacial proximity information given by the identification of two chemically cross-linked amino acids [50, 52]. A big advantage of all chemical crosslinking approaches is the potential to detect protein-protein binding sites in their native states. After enzymatic proteolysis cross-linked peptides are identified by mass spectrometry. Commonly used chemical crosslinkers are typically homobifunctional molecules with two lysine-reactive functional groups, like disuccinimide esters. To obtain a definite distance, aliphatic spacers are introduced between the reactive compounds. For LC-MS/MS analysis, peptides are fragmented using for example collision induced dissociation (CID) which leads to complex ion spectra. Therefore, a thorough computational analysis of MS data is essential. To facilitate the detection of cross-linked peptides by the generation of an isotopic signature, the applied crosslinker is used in two isotopic forms [51] as depicted in Figure 6.

## Identification of cross-linked peptides from MS/MS spectra by xQuest



**Figure 6: Identification of cross-linked peptides generating an isotopic signature**

First, isotopically shifted pairs are identified on MS<sup>1</sup> level. MS/MS spectra of isotopic pairs are compared and product ions are divided into common ions (highlighted in green) and cross-linked ions (highlighted in red). Common ions are identical in both spectra while cross-linked ions can be identified through the introduced mass shift. For the analysis, the generation of a focused database, comprising protein sequences of all proteins of interest, is essential. At the end, the software tool xQuest identifies candidate cross-links filtered by mass. This research was originally published in *Molecular & Cellular Proteomics*. Leitner, A., et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Molecular & Cellular Proteomics*. 2010; 9(8): p.1634-49. © the American Society for Biochemistry and Molecular Biology.

---

## 2 Aim of the study

Cilia are found in almost every eukaryotic cell. Therefore, deformation of cilia or non-functional cilia result in a broad range of phenotypes including cystic kidney disease, polydactyly, hydrocephalus, retinal degeneration and many more. Diseases that are caused by malfunction of cilia are summarized under the term ciliopathies.

The intraflagellar transport (IFT) machinery is important for the transport of protein cargo along the ciliary axoneme and can be divided into two protein subcomplexes: IFT-A and IFT-B. While IFT-B is involved in anterograde transport, IFT-A, consisting of six known proteins, is part of the retrograde transport machinery. This retrograde movement, which delivers protein cargo from the tip of a cilium back to its basal body is important for the maintenance and disassembly of cilia as well as for cellular signalling and recycling of components. According to previous knowledge, mutations in genes encoding some IFT-A complex components (WDR19, IFT43, IFT140) lead to syndromic (e.g. Sensenbrenner syndrome), as well as non-syndromic (e.g. nonsyndromic retinal dystrophy) forms of ciliopathies [33, 34, 40, 53]. To date, the function of the IFT-A complex on molecular level is only poorly understood. In contrast, there are many studies published dealing with the architecture of IFT-B and its involvement in ciliopathies [54-57]. To understand the mechanisms underlying IFT-A-related ciliopathies and to improve the understanding of the cellular function of IFT-A, the determination of the stoichiometric composition of IFT-A as well as the investigation of its structural information are the focus of this present study.

### Objective 1: Absolute quantification of purified IFT-A in its native compositions

To determine the native complex stoichiometry of IFT-A and to circumvent an effect of artificial over expression [58], the first aim of this study is the generation of monoclonal cell lines, stably expressing SF-TAP tagged IFT-A proteins, using the Flp-In system. The complete and intact purification of the whole IFT-A protein complex is a crucial step for a successful stoichiometry determination. Therefore, validation by SDS-PAGE with further silver staining and following quantitative mass spectrometry is obligatory. The stoichiometry determination of purified IFT-A in its native composition by absolute quantification combines a targeted mass spectrometry approach with representative standard peptides. To generate an economic equimolar standard mix, the “Equimolarity through Equalizer Peptide” (EtEP) is the method of choice in this study [59].

Based on its role in retrograde transport of protein cargo from the ciliary tip to the basal body, composition of IFT-A may change during the assembly and disassembly of cilia. Therefore, the

study of compositional changes in IFT-A stoichiometry during different ciliary stages is another important aim of this study. To investigate the composition of IFT-A in different cellular stages, hTERT-RPE1 cells, cultivated in different conditions, shall be used for the purification of IFT-A. Subsequently, the assembly of the IFT-A should be compared.

Objective 2: Effect in complex composition of IFT-A generating disease-related mutations of single components of the IFT-A protein complex

The last but not least step to understand the underlying molecular mechanism of IFT-A-related diseases is the analysis of induced changes in the complex stoichiometry.

For the introduction of stoichiometry changes in the protein complex of interest, the genome of Flp-In monoclonal lines, expressing TAP-tagged TULP3, are subjected to gene editing using the CRISPR/Cas9 system. Thereby, cell lines carrying mutations either in the WDR19 or IFT43 gene should be generated. To study the composition of IFT-A in disease-related mutations, the absolute quantification of purified IFT-A in validated and characterized mutant cell lines is another crucial step of this study.

Objective 3: Structural investigation of the IFT-A complex as basis for the prediction of functional relevant domains

Another aim of this study is the refinement of the stoichiometric investigations of IFT-A, achieved in this study, with structural information of this protein complex using chemical crosslinking. Therefore, disuccinimidyl suberate (DSS), a homobifunctional chemical crosslinker, should be used to gain proximity information of protein components of IFT-A. This elaborate approach of chemical crosslinking of TAP-purified IFT-A has to be optimised regarding the amount of cells as well as crosslinker concentration.



## 3 Material and Methods

### 3.1 Material

#### 3.1.1 Equipment

Analytical balance ABJ 120-4M	Kern & Sohn GmbH, Balingen, Germany
Autoclave DX-150	Systemec, Wettengel, Germany
CO <sub>2</sub> -Incubator HeraCell 150i	Heraeus, Hanau, Germany
Compartment dryer T20	Heraeus, Hanau, Germany
Developer Curix 60	Agfa, Mortsel, Belgium
Freezer (-80°C) Forma 900 Series	Thermo Fisher Scientific, Waltham, MA, USA
Freezer Liebherr Comfort	Liebherr, Bulle, Switzerland
Gel documentation device Easy RH	Herolab, Wiesloch, Germany
Ice machine AF200	Scotsman, Vernon Hills, IL, USA
Incubator INB 300	Memmert, Schwabach, Germany
IntelliMixer	NeoLab, Heidelberg, Germany
Laboratory balance S72	Kern & Sohn GmbH, Balingen, Germany
Laboratory hood model 854006.1	Wesemann, Syke, Germany
Laminar flow HeraSafe HS 12	Heraeus, Hanau, Germany
Laminar flow MSC 12	Thermo Fisher Scientific, Waltham, MA, USA
Magnetic stirrer and heater MR Hei-Standard	Heidolph, Schwabach, Germany
Megafuge 16	Heraeus, Hanau, Germany
Microscope PrimoVert	Zeiss, Göttingen, Germany
Microwave	Siemens, München, Germany
Multifuge X3R	Heraeus, Hanau, Germany
Neubauer counting chamber	Marienfeld, Lauda-Königshofen, Germany
pH meter PB-11	Sartorius, Göttingen, Germany
Platform shaker Duomax 1030	Heidolph, Schwabach, Germany
Power Supply Consort	Consort, Turnhout, Belgium
Power Supply PowerPak Basic	Bio-Rad, Hercules, CA, USA
Refrigerated Vapor Trap RVT400-230	Thermo Fisher Scientific, Waltham, MA, USA
Refrigerator	Liebherr, Bulle, Switzerland
Roller Mixer RM5	Assistant, Sondheim, Germany
Rotating incubator Infors HAT	Ecotron, Bruckmühl, Germany
Speedvac Concentrator SPD111V-230	Thermo Fisher Scientific, Waltham, MA, USA
Table top centrifuge 5415D	Eppendorf, Hamburg, Germany
Table top centrifuge Fresco17, refrigerated	Heraeus, Hanau, Germany
Table top centrifuge Pico 21	Heraeus, Hanau, Germany

Thermoblock MBT 250	ETG, Illmenau, Germany
Thermocycler Primus	MWG-Biotech, Ebersberg, Germany
Thermocycler Primus 96 Plus	MWG-Biotech, Ebersberg, Germany
Thermomixer Comfort	Eppendorf, Hamburg, Germany
Thermomixer Univortemp	Universal Labortechnik, Leipzig, Germany
Transluminator Fluo Link (312nm)	Bachofer, Reutlingen, Germany
Ultrapure water purification system Nanopure	Thermo Fisher Scientific, Waltham, MA, USA
Ultrasonic bath Sonorex Digitec	Bandelin, Berlin, Germany
UV/VIS Spectrometer T70	PG Instruments Limited, Lutterworth, UK
Vacuum pump 2522Z-02	Welch, Niles, IL, USA
Vortex Mixer	NeoLab, Heidelberg, Germany
Waterbath WNB 14 with shaker SV1422	Memmert, Schwabach, Germany

### 3.1.2 Consumables and labware

Accu-Jet® pro pipetting aid	Brand, Wertheim, Germany
Amersham Hyperfilm ECL	GE Healthcare, Waukesha, WI, USA
Baffled flask 250/500ml	Neolab, Heidelberg, Germany
Blue rack for 6/12 Tubes	GLW Storing Systems, Würzburg, Germany
Box for pipette tips	Gilson, Middleton, WI, USA
Cell culture dish, 6 well	BD Biosciences, Franklin Lakes, NJ, USA
Cell scraper	Sarstedt, Nümbrecht, Germany
Centrifugation tubes 15/50ml	Greiner bio-one, Kremsmünster, Austria
Comb 10/15 well for 0.75mm gels	Bio-Rad, Hercules, CA, USA
Cryo Tube Rack	Nunc, Rochester, NY, USA
Cryobox 0.5/1.5/2.0ml tubes	Carl-Roth, Karlsruhe, Germany
Culture tube 14ml	BD Biosciences, Franklin Lakes, NJ, USA
Cuvette 1.5ml	Sarstedt, Nümbrecht, Germany
Drigalski spatula	Carl-Roth, Karlsruhe, Germany
Filter system 0.22µm	Corning, NY, USA
Gel loading tips	Carl-Roth, Karlsruhe, Germany
Glas plate with 0.75mm spacer	Bio-Rad, Hercules, CA, USA
Gloves purple nitrile	Kimberly-Clark, Irving, TX, USA
Gloves soft nitrile	Paul Hartmann, Heidenheim, Germany
Graduated measuring glass	Duran Group, Wertheim, Germany
Hybond-P PVDF Transfer membrane	GE Healthcare, Waukesha, WI, USA
Hyperscassette	GE Healthcare, Waukesha, WI, USA
Hyperscreen	GE Healthcare, Waukesha, WI, USA
Icebath	Neolab, Heidelberg, Germany

Inlays for Cryobox for 0.5/1.7/2.0ml tubes	Carl-Roth, Karlsruhe, Germany
Inoculation loop	Carl-Roth, Karlsruhe, Germany
Laboratory bottles, 50-2000ml	Neolab, Heidelberg, Germany
MicroSpin Columns	GE Healthcare, Waukesha, WI, USA
Multiwell plate, 96wells	Greiner bio-one, Kremsmünster, Austria
Parafilm sealing foil	Brand, Wertheim, Germany
Pasteur capillary pipette	VWR International, West Chester, PA, USA
Pasteru Pipette PP	VWR International, West Chester, PA, USA
PCR rack	Carl-Roth, Karlsruhe, Germany
Petridish 90 x 14.2mm	VWR International, West Chester, PA, USA
pH indicator sticks	Carl-Roth, Karlsruhe, Germany
Pipettes 2/10/20/200/1000µl	Gilson, Middleton, WI, USA
Pipette tips 1-10/1-200/101-1000µl	Sarstedt, Nümbrecht, Germany
Pipettes (serological) 2/5/10/25/50ml	BD Biosciences, Franklin Lakes, NJ, USA
Polypropylen insert with bottom spring	Sigma-Aldrich, St. Louis, MO, USA
Reaction tube 0.2	Eppendorf, Hamburg, Germany
Reaction tube (safe-lock) 0.5/1.5/5.0ml	Eppendorf, Hamburg, Germany
Reaction tube 15ml	Sarstedt, Nümbrecht, Germany
Reaction tube 50ml	BD Biosciences, Franklin Lakes, NJ, USA
Scalpell	VWR International, West Chester, PA, USA
Short plate	Bio-Rad, Hercules, CA, USA
Spin Columns, 3MWCO, 30MWCO	GE Healthcare, Waukesha, WI, USA
Stage Tips C-18, 200µl	Thermo Fisher Scientific, Waltham, MA, USA
Sterilfilter Millex 0.22 µm	Merck, Darmstadt, Germany
Tissue Dishes, 10cm, Nunclon Surface	Nunc, Rochester, NY, USA
Tissue Dishes, 14cm, Nunclon Surface	Nunc, Rochester, NY, USA
Whatman chromatography paper	GE Healthcare, Waukesha, WI, USA

### 3.1.3 Chemicals

1,4-Dithiothreitol (DTT)	Merck, Darmstadt, Germany
2-Iodacetamide (IAA)	Merck, Darmstadt, Germany
2-Mercaptoethanol	Sigma-Aldrich, St. Louis, MO, USA
2-Propanol LC grade	Merck, Darmstadt, Germany
2-Propanol p.a.	Merck, Darmstadt, Germany
Acetonitrile LC-MS CHROMASOLV®, ≥99.9%	Sigma-Aldrich, St. Louis, MO, USA
Agar-Agar	Carl-Roth, Karlsruhe, Germany
Agarose	Lonza, Basel, Switzerland
Ammonium Bicarbonate	Sigma-Aldrich, St. Louis, MO, USA

Ammonium Persulfate	Merck, Darmstadt, Germany
Ampicillin Sodium Crystalline	Carl-Roth, Karlsruhe, Germany
Bis-Acrylamid/Acrylamid (37.5:1:30%)	Serva Elektrophoresis, Heidelberg, Germany
Bromphenolblue	Sigma-Aldrich, St. Louis, MO, USA
Chloramphenicol	Carl-Roth, Karlsruhe, Germany
Chloroform p.a.	Merck, Darmstadt, Germany
Dimethyl Sulfoxide	Applichem, Darmstadt, Germany
EDTA Disodium Salt Dihydrate	Applichem, Darmstadt, Germany
Ethanol p.a.	Merck, Darmstadt, Germany
Ethidiumbromide	Applichem, Darmstadt, Germany
Ficoll 400	Sigma-Aldrich, St. Louis, MO, USA
Formaldehyde Solution (37%)	Sigma-Aldrich, St. Louis, MO, USA
Glacial Acetic Acid p.a.	Merck, Darmstadt, Germany
Glycerol	Carl-Roth, Karlsruhe, Germany
Glycin	Carl-Roth, Karlsruhe, Germany
HEPES	Sigma-Aldrich, St. Louis, MO, USA
Hydrochloric Acid p.a.	Merck, Darmstadt, Germany
Kanamycin Sulfate	Carl-Roth, Karlsruhe, Germany
Magnesium Chloride	Sigma-Aldrich, St. Louis, MO, USA
Methanol LC-MS grade	Merck, Darmstadt, Germany
Methanol LC-MS grade	VWR International, West Chester, PA, USA
Methanol p.a.	Merck, Darmstadt, Germany
Nonidet P40	Roche, Penzberg, Germany
OrangeG	Sigma-Aldrich, St. Louis, MO, USA
Polyethylenimine, Linear (MW 25,000)	Polysciences, Warrington, PA, USA
Ponceau S	Sigma-Aldrich, St. Louis, MO, USA
Potassium Chloride p.a.	Carl-Roth, Karlsruhe, Germany
RapiGest SF Surfactant	Waters, Milford, MA, USA
Silver Nitrate	Merck, Darmstadt, Germany
Sodium Chloride p.a.	Merck, Darmstadt, Germany
Sodium Dodecylsulfate, Pellet	Sigma-Aldrich, St. Louis, MO, USA
Sodium Dodecylsulfate, 20%	Applichem, Darmstadt, Germany
Sodium Hydroxide pellets p.a.	Carl-Roth, Karlsruhe, Germany
Sodium Thiosulfate Pentahydrate, supra pure	Merck, Darmstadt, Germany
TEMED p.a.	Merck, Darmstadt, Germany
Trifluoroacetic Acid, for protein seq.	Merck, Darmstadt, Germany
Tris(hydroxymethyl) Aminomethane (Tris ultrapure)	Sigma-Aldrich, St. Louis, MO, USA

Trypsin from porcine pancreas, proteomics grade	Sigma-Aldrich, St. Louis, MO, USA
Tryptone/ Peptone from Casein	Carl-Roth, Karlsruhe, Germany
Tween® 20	Sigma-Aldrich, St. Louis, MO, USA
Water, HPLC grade	Merck, Darmstadt, Germany
Water, HPLC grade	VWR International, West Chester, PA, USA
Water bath stabilizer, AKASOLV Aqua Care	Carl-Roth, Karlsruhe, Germany
Yeast Extract	Carl-Roth, Karlsruhe, Germany

### 3.1.4 Special reagents

<sup>12</sup> C <sub>6</sub> , <sup>14</sup> N <sub>2</sub> - Lysine	Silantes, München, Germany
4.4.5.5.-D <sub>4</sub> -L-Lysine	Silantes, München, Germany
<sup>13</sup> C <sub>6</sub> , <sup>15</sup> N <sub>2</sub> -L-Lysine	Silantes, München, Germany
<sup>12</sup> C <sub>6</sub> , <sup>14</sup> N <sub>4</sub> - Arginine	Silantes, München, Germany
<sup>13</sup> C <sub>6</sub> - Arginine	Silantes, München, Germany
<sup>13</sup> C <sub>6</sub> , <sup>15</sup> N <sub>4</sub> -L-Arginine	Silantes, München, Germany
Adenosin 5'-Diphosphate (ADP)	Sigma-Aldrich, St. Louis, MO, USA
AGFA Developer G153	Röntgen Bender, Baden-Baden, Germany
AGFA Fixer G354	Röntgen Bender, Baden-Baden, Germany
Anti-FLAG-M2-agarose	Sigma-Aldrich, St. Louis, MO, USA
Blotting Grade Blocker, nonfat dry	Bio-Rad, Hercules, CA, USA
Bovine Serum Albumin (BSA)	PAA, Pasching, Austria
Coomassie Brilliant Blue	Merck, Darmstadt, Germany
D-Desthiobiotin	IBA, Göttingen, Germany
Dialysed Fetal Bovine Serum (FBS)	Sigma-Aldrich, St. Louis, MO, USA
Didesoxyadenosine 5'-triphosphate (ATP)	Sigma-Aldrich, St. Louis, MO, USA
Disuccinimidyl Suberate (DSS) H12/ D12	Creative Molecules Inc
dNTP Mix, 40µM	New England Biolabs, Ipswich, MA, USA
Dulbecco's Modified Eagle Medium	Sigma-Aldrich, St. Louis, MO, USA
Dulbecco's PBS	Sigma-Aldrich, St. Louis, MO, USA
ECL plus Western Blotting Substrate	Thermo Fisher Scientific, Waltham, MA, USA
ECL Western Blotting Substrate	Thermo Fisher Scientific, Waltham, MA, USA
Effectene transfection reagent	Qiagen, Hilden, Germany
Fetal Bovine Serum (FBS)	Sigma-Aldrich, St. Louis, MO, USA
Flag peptide	Sigma-Aldrich, St. Louis, MO, USA
GeneRuler™1kb Plus DNA ladder	Thermo Fisher Scientific, Waltham, MA, USA
L-Glutamine, 200mM	Life Technologies, Carlsbad, CA, USA
Hygromycin B (50mg/ml)	Invitrogen, Carlsbad, CA, USA
PageRuler Plus, 250kDa prestained	Fermentas, Burlington, Canada

---

PageRuler, 170kDa prestained	Fermentas, Burlington, Canada
Penicillin/ Streptomycin	Life Technologies, Carlsbad, CA, USA
Phosphatase Inhibitor Cocktail 2	Sigma-Aldrich, St. Louis, MO, USA
Phosphatase Inhibitor Cocktail 3	Sigma-Aldrich, St. Louis, MO, USA
Proline	Silantes, München, Germany
Protease Inhibitor Cocktail Complete	Roche, Penzberg, Germany
Protein Assay Dye Reagent	Bio-Rad, Hercules, CA, USA
Protein G PLUS-Agarose	Santa Cruz, Santa Cruz, CA, USA
Protein A PLUS-Agarose	Santa Cruz, Santa Cruz, CA, USA
Puromycin Dihydrochloride GIBCO™ (sterile) (10mg/ml)	Thermo Fisher Scientific, Waltham, MA, USA
SILAC DMEM	Thermo Fisher Scientific, Waltham, MA, USA
Strep Tactin Superflow (50% suspension)	IBA, Göttingen, Germany
Strep-TAG Elution Buffer	IBA, Göttingen, Germany
Trypsin EDTA	Life Technologies, Carlsbad, CA, USA

### 3.1.5 Buffers, solutions and media

In this study, deionized water is described as dH<sub>2</sub>O and ddH<sub>2</sub>O refers to ultra-pure water.

#### 3.1.5.1 *E.coli* culture

LB-Medium  
1% (w/v) Tryptone/ Peptone from Casein  
0.5% (w/v) Yeast Extract  
1% (w/v) NaCl  
adjust to pH 7.0 using NaOH  
in dH<sub>2</sub>O

LB-Agar  
1% (w/v) Tryptone/ Peptone from Casein  
0.5% (w/v) Yeast Extract  
1% (w/v) NaCl  
1% (w/v) Agar-Agar  
in dH<sub>2</sub>O

#### 3.1.5.2 Mammalian cell culture

Cryo Medium  
90% FBS  
10% DMSO

Growth Medium  
Dulbecco's Modified Eagle Medium  
10% FBS  
0.5% Penicillin/ Streptomycin

PEI Transfection Reagent  
1mg/ml Polyethylenimine (PEI)  
in dH<sub>2</sub>O

SILAC Medium Light  
SILAC DMEM  
10% FBS (dialysed)  
2% L-Glutamine  
0.5% Penicillin/ Streptomycin  
2mM Proline  
0.55mM <sup>12</sup>C<sub>6</sub>, <sup>14</sup>N<sub>2</sub> Lysine  
0.4mM <sup>12</sup>C<sub>6</sub>, <sup>14</sup>N<sub>4</sub> Arginine

---

SILAC Medium Medium	SILAC DMEM 10% FBS (dialysed) 2% L-Glutamine 0.5% Penicillin/ Streptomycin 2mM Proline 0.55mM 4,4,5,5- D <sub>4</sub> -L-Lysine 0.4mM <sup>13</sup> C <sub>6</sub> Arginine
SILAC Medium Heavy	SILAC DMEM 10% FBS (dialysed) 2% L-Glutamine 0.5% Penicillin/ Streptomycin 2mM Proline 0.55mM <sup>13</sup> C <sub>6</sub> , <sup>15</sup> N <sub>2</sub> -L-Lysine 0.4mM <sup>13</sup> C <sub>6</sub> , <sup>15</sup> N <sub>4</sub> -L-Arginine
Flp-In Growth Medium	Dulbecco's Modified Eagle Medium 10% FBS 1% Penicillin/ Streptomycin 2mM L-Glutamine 100µg/ml Zeocin
Flp-In Transfection Medium	Dulbecco's Modified Eagle Medium 10% FBS 2mM L-Glutamine
Flp-In Post-Transfection Medium	Dulbecco's Modified Eagle Medium 10% FBS 2mM L-Glutamine 1% Penicillin/ Streptomycin



---

Flp-In Selection Medium	Dulbecco's Modified Eagle Medium 10% FBS 2mM L-Glutamine 1% Penicillin/ Streptomycin 100µg/ml Hygromycin
-------------------------	--

CRISP Selection Medium	Dulbecco's Modified Eagle Medium 10% FBS 0.05% Penicillin/ Streptomycin 8µg/ml Puromycin
------------------------	---

### 3.1.5.3 Agarose Gels

Orange G (6x)	250mg/ml Ficoll 400 0.5% (w/v) SDS 50mM EDTA 1 spatula tip Orange G in ddH <sub>2</sub> O
---------------	---

TAE-buffer (50x)	2M Tris 50mM EDTA 1M Acetic Acid in dH <sub>2</sub> O
------------------	--

### 3.1.5.4 SDS-PAGE, Coomassie staining and Western blot analysis

Blocking Solution	5% Blotting Grade Blocker in TBST (1x)
-------------------	---

Coomassie Staining Solution	0.4% (w/v) Coomassie Brilliant Blue in dH <sub>2</sub> O
-----------------------------	---

Fixer Solution	50% MeOH 12% Acetic Acid in dH <sub>2</sub> O
----------------	---

---

Laemmli buffer (5x)	250mM Tris-HCl pH 6.8 5% SDS 50% Glycerol 500mM 2-Mercaptoethanol 0.05% (w/v) Bromphenol Blue in ddH <sub>2</sub> O
Ponceau-Solution (10x)	0.1% (w/v) Ponceau S 50% Acetic Acid in ddH <sub>2</sub> O
Running buffer (10x)	2M Glycine 250mM Tris 1% (w/v) SDS in dH <sub>2</sub> O
Separation Gel	25% Tris-HCl pH 8.8 12% Acrylamide 0.1% SDS 0.2% TEMED 0.05% Ammonium Persulfate (APS) in ddH <sub>2</sub> O
Stacking Gel	14% Tris-HCl pH 6.8 4.2% Acrylamide 0.1% SDS 0.4% TEMED 0.1% Ammonium Persulfate (APS) in ddH <sub>2</sub> O
TBST (10x)	300mM Tris 1.5M NaCl 1% Tween® 20 in dH <sub>2</sub> O adjust pH to 7.4 using HCl

---

Western buffer (10x)                      1.92M Glycine  
250mM Tris  
in dH<sub>2</sub>O

Western buffer (1x)                      10% Western buffer  
20% MeOH  
In dH<sub>2</sub>O

### 3.1.5.5 Silverstaining

AgNO<sub>3</sub>-Solution                      11.8mM Silver Nitrate (AgNO<sub>3</sub>)  
0.075% Formaldehyde (37%)  
in ddH<sub>2</sub>O

Coomassie Fixer                      50% MeOH  
12% Acetic Acid  
in ddH<sub>2</sub>O

Developer                      1M Sodium Carbonate (Na<sub>2</sub>CO<sub>3</sub>)  
5% Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>-Solution  
0.1% Formaldehyde (37%)  
in ddH<sub>2</sub>O

Fixer                      50% MeOH  
12% Acetic Acid  
0.05% Formaldehyde (37%)  
in ddH<sub>2</sub>O

Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>-Solution                      1.3mM Sodium Thiosulfate (Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>)  
in ddH<sub>2</sub>O

Preserver                      20% EtOH  
2% Glycerol  
in ddH<sub>2</sub>O

### 3.1.5.6 Affinity purification

FLAG-Peptide (25x)                      5mg/ml FLAG-Peptide  
in TBS (1x)

---

Lysis Buffer	0.55% Nonidet P40 2% Protease Inhibitor Cocktail Complete (PIC) 1% Phosphatase Inhibitor Cocktail 2 1% Phosphatase Inhibitor Cocktail 3 in TBS (1x)
Strep-Elution Buffer	10% Strep-TAG Elution Buffer in H <sub>2</sub> O (HPLC grade)
TBS (10x)	300mM Tris 1.5M NaCl in dH <sub>2</sub> O adjust pH to 7.4 using HCl
Washing Buffer	0.12% Nonidet P40 1% Phosphatase Inhibitor Cocktail 2 1% Phosphatase Inhibitor Cocktail 2 in TBS (1x)
<b>3.1.5.7 Chemical crosslinking</b>	
Crosslinking Buffer	50mM HEPES 150mM NaCl 5mM MgCl <sub>2</sub> in ddH <sub>2</sub> O adjust pH to 7.5 using NaOH
Disuccinimidyl Suberate	12.5mM Disuccinimidyl Suberate (H12/D12) in DMSO
FLAG-peptide (25x)	5mg/ml FLAG-peptide in Crosslinking Buffer

CL-Lysis Buffer	0.58% Nonidet P40 2% Protease Inhibitor Cocktail Complete 1% Phosphatase Inhibitor Cocktail 2 1% Phosphatase Inhibitor Cocktail 3 in Crosslinking Buffer
CL-Strep-Elution Buffer (10x)	5.35mg/ml D-Desthiobiotin in Crosslinking Buffer
CL-Washing Buffer	0.12% Nonidet P40 1% Phosphatase Inhibitor Cocktail 2 1% Phosphatase Inhibitor Cocktail 3 in Crosslinking Buffer

### 3.1.6 Kits

Bio-Rad Protein Assay Kit	Bio-Rad, Hercules, CA, USA
EndoFree Plasmid Maxi Kit	Qiagen, Hildem Germany
Enhanced Chemiluminescence Kit, ECLplus	GE Healthcare, Waukesha, WI, USA,
Gel Extraction Kit	Fermentas, Burlington, Canada
GeneJet™ Plasmid Miniprep Kit	Fermentas, Burlington, Canada
peqGold Tissue DNA Mini Kit	Peqlab Biotechnologie GmbH, Erlangen, Germany
PureYield Plasmid Midiprep Kit	Promega, Fitchburg, WI, USA

### 3.1.7 Enzymes

BP Clonase II Enzyme MIX, with proteinase K	Invitrogen, Carlsbad, CA, USA
FastDigest BbsI/Bpil	Thermo Fisher Scientific, Waltham, MA, USA
HindIII	New England Biolabs, Ipswich, MA, USA
LR Clonase II Enzyme MIX, with proteinase K	Invitrogen, Carlsbad, CA, USA
Phusion High-Fidelity DNA Polymerase supplied with HF Reaction buffer (5x) and MgCl <sub>2</sub> (50mM)	Thermo Fisher Scientific, Waltham, MA, USA
PlasmidSafe exonuclease supplied with PlasmidSafe buffer (10x)	Epicentre, Madison, WI, USA
T4 DNA Ligase supplied with Reaction buffer (10x)	Roche, Penzberg, Germany
T4 Polynucleotide Kinasesupplied with Reaction buffer (10x)	Roche, Penzberg, Germany
Taq DNA Polymerase supplied with Reaction buffer (10x)	Fermentas, Burlington, Canada
XhoI	New England Biolabs, Ipswich, MA, USA

### 3.1.8 *E.coli* strains

Library Efficient® DH5α	Invitrogen, Carlsbad, CA, USA
-------------------------	-------------------------------

### 3.1.9 Oligonucleotides

**Table 1: Oligonucleotides for gene editing using CRISPR/Cas9 system**

Oligonucleotides were purchased from Eurofins Genomics.

Oligo name	Sequence 5'-3'
IFT43 sgRNA1_for	caccgctgggaccatctgttactg
IFT43 sgRNA1_rev	gaccctggtagacaagtgaccaaa
IFT43 sgRNA2_for	caccgcctccttccgtctgcagt
IFT43 sgRNA2_rev	cggaggaagaggcagacgtcacaaa
IFT43 HDR1	gctgggactgggaccacctgttcaccgaggtgtcctcagaggtcctcactgagtgggaccactgtagact gagaaagaggaccctgcggggcaggccaggcacacctgag
WDR19 sgRNA1_for	caccgaaataactaactcgtaatat
WDR19 sgRNA1_rev	ctttatgattgagcactatacaaaa
WDR19 sgRNA2_for	caccgaaaagagatatcttcaggc
WDR19 sgRNA2_rev	ccttttctatagaagtccgcaaa
WDR19 HDR1	gaaggagaaaagagatatcttcaggctggaaaattcttctgtgtgtggccaataactaagaattctcacgc gttagtatttgccaagaaaatatacactgactccgcaggaataattg
WDR19 HDR2	ctgaaggagaaaagagatatcttcaggctggaaaattcttctgtgtgtggccaataacttacgccttagcatt tgccaagaaaatatacactgactccgcaggaataattgtagg

**Table 2: Sequencing and PCR Primers**

Primer name	Sequence 5'-3'	Application
attB1	gggacaagttgtacaaaaaagcaggct	Gateway Cloning
attB2	ggggaccactttgtacaagaaagctgggt	Gateway Cloning
BGH polyA signal (P-13-074)	ctgtgccttctagttgccag	FlpIn System
CMV_for	cgcaaattgggcggtaggcgtg	CRISPR/Cas9 System
Hygromycin_rev (P-13-067)	gcaaagtgccgataaacataac	FlpIn System
IFT43_genomfor	ggcattcctgcaggctcag	CRISPR/Cas9 System
IFT43_genomrev	gaggagatggcacagaataagc	CRISPR/Cas9 System
IFT43_HDR1_for	gaccactgtagactgagaaag	CRISPR/Cas9 System

IFT43_HDRfail_for	gaccactgcagacggag	CRISPR/Cas9 System
lacZ-Zeocin_rev (P-13-079)	gggaacaaacggcggattga	FloIn System
lacZ-Zeocin_rev (P-13-075)	gtttcccagtcacgacgtt	FloIn System
pcDNA3_for	ctctggctaactagagaac	CRISPR/Cas9 System
SV <sub>40</sub> _for (P-13-066)	aattagtcagcaaccaggtgtg	FloIn System
SV <sub>40</sub> _for (P-13-070)	tccgccccatggctgactaa	FloIn System
WDR19_genomfor	gaatgtcatctccctactgtctg	CRISPR/Cas9 System
WDR19_genomrev	gccttaggccaaggggt	CRISPR/Cas9 System
WDR19_HDR1_for	gccaataactagaattctcacgcg	CRISPR/Cas9 System
WDR19_HDR2_for	gccaatactacgcttagcatttg	CRISPR/Cas9 System
WDR19_HDRfail_for	gccaatattcacgagtagtatttg	CRISPR/Cas9 System

### 3.1.10 Plasmids

**Table 3: Vectors**

Vector name	Vector type	Tag	Resistance
Gateway® pDONR™221	Donor Vector	-	Kanamycin
Gateway® pENTR223	Entry Vector	-	Spectinomycin
(C)SF-TAP pDEST	Destination Vector	(C)-SF-TAP	Ampicillin, Chlormaphenicol, Neomycin
(N)SF-TAP pDEST	Destination Vector	(N)-SF-TAP	Ampicillin, Chlormaphenicol, Neomycin
pcDNA5/FRT/TO	Destination Vector	-	Hygromycin
pSpCas9n(BB)-2A-Puro	Destination Vector	-	Ampicillin, Puromycin

### 3.1.11 Constructs

**Table 4: Constructs**

Construct name	cDNA (human)	Vector
(C)-SF-TAP pDEST-IFT122	IFT122	(C)-SF-TAP pDEST
(C)-SF-TAP pDEST-LCA5	LCA5	(C)-SF-TAP pDEST
(C)-SF-TAP pDEST-TULP3	TULP3	(C)-SF-TAP pDEST

### 3.1.12 Antibodies

**Table 5: Primary Antibodies**

Antibody	Species	Dilution	Vendor
Anti-FLAG® M2-Peroxidase (HRP)	mouse (monoclonal)	1:1000	Sigma-Aldrich, St. Louis, MO, USA
Anti-βGal	Rabbit (polyclonal)	1:100000	Abcam, Cambridge, UK

**Table 6: Secondary Antibodies**

Antibody	Species	Dilution	Vendor
Anti-rabbit IgG	goat (polyclonal)	1:15000	Jackson ImmunoResearch, West Grove, PA, USA

### 3.1.13 Liquid chromatography and mass spectrometry

Acclaim PepMap RSLC C18 (75µm x 25cm)	Thermo Fisher Scientific, Waltham, MA, USA
Ettan FPLC	GE Healthcare, Waukesha, WI, USA
LTQ Orbitrap Velos	Thermo Fisher Scientific, Waltham, MA, USA
LTQ Orbitrap XL	Thermo Fisher Scientific, Waltham, MA, USA
Orbitrap Fusion	Thermo Fisher Scientific, Waltham, MA, USA
QExactive Plus	Thermo Fisher Scientific, Waltham, MA, USA
Qtrap 5500	AB Sciex, Framingham, MA, USA
Superdex Peptide PC 3.2/30" column	GE Healthcare, Waukesha, WI, USA
Ultimate 3000 Nano-RSLC	Thermo Fisher Scientific, Waltham, MA, USA
µ-Precolumn Acclaim PepMap100 C18 (300µm i.d. x 5mm)	Thermo Fisher Scientific, Waltham, MA, USA

### 3.1.14 Software and databases

#### 3.1.14.1 Software

**Table 7: Software**

Adobe Illustrator CS5.1	Adobe Systems Inc., San Jose, CA, USA
Adobe Photoshop CS5	Adobe Systems Inc., San Jose, CA, USA
Chromeleon 2.1.4	Thermo Fisher Scientific, Waltham, MA, USA
EndNote X6	Thomson Reuters, New York, NY, USA
MaxQuant 1.5.3.30	MPI of Biochemistry, Martinsried, Germany
MS Office 2010 (Word, Excel, PowerPoint)	Microsoft, Redmond, WA, USA
OriginPro 9.0	OriginLab Corporation, Northampton, MA, USA
Perseus 1.5.2.6	MPI of Biochemistry, Martinsried, Germany
Proteome Discoverer Daemon 2.0	Thermo Fisher Scientific, Waltham, MA, USA
RawMeat	Vast Scientific,



---

Scaffold 4.0	Proteome Software Inc., Portland, OR, USA
Skyline 3.5	MacCoss Lab Software,
Vector NTI Suite 9.0	Invitrogen, Carlsbad, CA, USA
XCalibur 2.07	Thermo Fisher Scientific, Waltham, MA, USA
Xinet	Rappsilber Laboratory, Edinburgh, UK
xQuest/xProphet	ETH Zürich, Switzerland

### 3.1.14.2 Databases

**Table 8: Databases**

Ensembl Genome Browser	<a href="http://www.ensembl.org">http://www.ensembl.org</a>
NCBI	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>
NCBI Blast	<a href="http://www.ncbi.nlm.nih.gov/Blast.cgi">http://www.ncbi.nlm.nih.gov/Blast.cgi</a>
NCBI Nucleotide	<a href="http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide">http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide</a>
NCBI Protein	<a href="http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein">http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein</a>
NCBI PubMed	<a href="http://www.ncbi.nlm.nih.gov/sites/entrez">http://www.ncbi.nlm.nih.gov/sites/entrez</a>
Swiss-Prot	<a href="http://www.expasy.ch/sprot/">http://www.expasy.ch/sprot/</a>
UniProt	<a href="http://uniprot.org/">http://uniprot.org/</a>

## **3.2 Methods**

### **3.2.1 Molecular biology**

#### **3.2.1.1 Enzymatic DNA treatment**

The basis of molecular biology is the use of different enzymes to generate DNA constructs for specific needs.

#### **DNA restriction digestion**

An important tool in molecular biology is the use of restriction endonucleases. These enzymes recognise a specific basepair sequence and thereby create a targeted double-strand cut at or next to its specific recognition site. Depending on the applied enzyme, this cut results in blunt or sticky ends. For DNA restriction digestion, buffers and reaction conditions were chosen as recommended by NEB (<http://www.neb.com>). Digestion of 1µg of double-stranded DNA was performed using 20U of restriction endonuclease.

To create corresponding ends, 5µg of the plasmid containing target cDNA as well as 5µg of the destination vector were incubated with the same restriction enzyme for 2h at 37°C. For double digestion, buffer conditions with best reactivity for both enzymes were chosen. If reaction conditions of selected endonucleases were not compatible, DNA restriction was performed sequentially.

#### **Annealing and phosphorylating of oligonucleotides**

Short double-stranded DNA (dsDNA) fragments can be generated through the annealing of complementary synthetic oligonucleotides. The annealing of single-stranded DNA (ssDNA) occurs through the formation of hydrogen bonds between complementary nucleobases A/T and G/C. This reaction leads to basepairs forming dsDNA. To enable subsequent ligation of blunt ended dsDNA into restricted plasmid vector, phosphorylation of the hydroxy group (-OH) at the 5' end is necessary.

To phosphorylate and anneal single-stranded oligonucleotides, 1µl of each complementary fragment (100µM) was incubated with 1µl T4 polynucleotide kinase (T4 PNK) in 1xT4 PNK buffer in a total reaction volume of 10µl. Incubation was performed in a thermocycler. ssDNA was phosphorylated for 30min at 37°C, followed by a denaturation for 5min at 95°C and subsequently annealing of complementary oligonucleotides by cooling down slowly to 25°C with a decrease of 5°C/min as described previously [60].

## Ligation

The ligation of a restriction enzyme digested DNA fragment with a vector backbone can be introduced by the formation of phosphodiester bonds between the 3' hydroxy group and the 5' phosphate group of both molecules. Ligation was enzymatically catalysed by T4 DNA Ligase and performed as described by Ran and colleagues [60].

### 3.2.1.2 *E.coli* culture and plating

*E.coli* were cultivated in LB-medium. Depending on the antibiotic-resistance, encoded on the introduced plasmid, the LB-medium was supplemented with Ampicillin (100µg/ml), Kanamycin (50µg/ml), Chloramphenicol (30µg/ml) or Spectomycin (100µg/ml) for selection. *E.coli* were cultured in different volumes of LB-medium according to following experimental needs.

### 3.2.1.3 Transformation and cryoconservation of *E.coli*

For the uptake of pure DNA into *E.coli*, plasmid DNA constructs are transformed into the bacterium using a heat shock protocol. Thereby, competent bacteria are heated up to 42°C to permeabilise the cell wall which allows DNA to enter the cell.

Transformation of DNA constructs into competent cells was performed using 50µl of library efficient DH5α and either 1µl of vector DNA (100ng/µl) or 5µl of ligation product. Incubation was performed for 1h on ice. Subsequently, cells were heat shocked for 45sec at 42°C before sample was cooled down on ice for 2min. After 250µl S.O.C. Medium was added, cells were incubated for 1h at 37°C under constant agitation for initial bacteria growth. For selection, 25-100µl were spread on a pre-warmed LB-Agar plate containing the appropriate antibiotic. Plates were incubated overnight at 37°C. The next day, formed colonies were transferred to 5ml LB-medium supplemented with the appropriate antibiotic and cultured for at least 5h. For cryoconservation of bacteria at -80°C, 500µl of bacteria suspension were mixed with 500µl sterile 80% glycerol. To cultivate bacteria strains again, a small amount of the frozen cryoconservation stock was transferred to LB-medium using an inoculation loop. To initiate growth, cells were cultured overnight at 37°C under constant agitation. For the purification of plasmid DNA, remaining bacterial culture was used.

### 3.2.1.4 DNA isolation from *E.coli*

To purify plasmid DNA from bacterial cultures, different kits were used. Depending on the culture volume, the GeneJet™ Plasmid Miniprep (for 5ml culture), PureYield Plasmid Midiprep (for 100ml culture) or EndoFree Plasmid Maxi Kit (for 100ml culture) was used according to the

manufacturer's instructions. To determine the DNA concentration, the absorbance at 260nm and 280nm was measured using a UV/VIS photometer and calculated the following way:

$$c_{\text{DNA}} [\mu\text{g/ml}] = \text{OD}_{260} \cdot V \cdot F$$

$c_{\text{DNA}}$  concentration of DNA solution

$\text{OD}_{260}$  optical density at 260nm

$V$  dilution factor

$F$  multiplication factor

50 for dsDNA

40 for RNA

33 for ssDNA

20 for oligonucleotides

Additionally, calculation of the ratio of  $\text{OD}_{260}/\text{OD}_{280}$  can be used to determine the purity of the plasmid DNA. Highly pure DNA preparations have ratios of 1.8 or higher.

### 3.2.1.5 Polymerase chain reaction - PCR

Polymerase chain reaction (PCR) is a fast and effective way for the in vitro amplification of DNA fragments. This reaction is divided into three steps (Denaturation, Annealing and Elongation) which are repeated for several cycles (30-35). At the very beginning, an initial denaturation of the dsDNA into ssDNA is performed for 2min at 96°C. Subsequently, every amplification cycle starts with a denaturation for 30sec at 98°C, followed by the binding of oligonucleotide primers to their complementary sequence, each on one strand, at a primer specific annealing temperature. The last step includes the DNA synthesis at 72°C. This elongation step is catalysed by a thermostable DNA polymerase (e.g. Phusion High-Fidelity DNA Polymerase or DreamTaq DNA Polymerase). After amplification cycles are completed, a final extension step is performed for 10min at 72°C. To terminate PCR, the temperature is decreased to 15°C. Based on different template length as well as primer melting temperatures, the following PCR program (see Table 9) is adjusted for every individual PCR. The elongation time is adjusted according to template length, whereas the annealing temperature is varied according to the melting temperature ( $T_m$ ) of used primers.

**Table 9: PCR program**

Initial Denaturation	96°C	2min	} up to 35 cycles
Denaturation	98°C	30sec	
Annealing	60°C (55°C-71°C)	40sec	
Elongation	72°C	1min (1kb/min)	
Final extension	72°C	10min	
Cool down	15°C	∞	

For the validation of Flp-In monoclonal lines using PCR, 60ng of genomic DNA was used and mixed with 1µl dNTP mix, 1.5µl MgCl<sub>2</sub> (50mM), 2.5µl of both primers (10µM) and 0.5µl Phusion High-Fidelity DNA Polymerase (2U/µl) in a final volume of 50µl in 1x Phusion HF reaction buffer.

To validate mutations generated using the CRISPR/Cas9 system, PCR was performed with 1µl of genomic DNA mixed with 0.2µl dNTPs (each 10mM), 0.3µl DreamTaq DNA Polymerase and 1µl of each primer (10pM) in 1x GreenTaq Buffer with a total reaction volume of 20µl. Genomic DNA was extracted applying the peqGOLD Tissue DNA Mini Kit according to the manufacturer's instructions.

### 3.2.1.6 Cloning of plasmid expression vectors

Cloning is a commonly used method to generate defined DNA constructs at the level of individual bacterial cells and to produce large amounts of plasmid DNA through cultivation of individual bacterial clones.

#### Gateway cloning

A high efficient and universal cloning method is the so called Gateway cloning, commercialized by Invitrogen. This approach is based on site specific recombination at defined sequences derived from the attachment (att) sites of bacteriophage lambda [61]. A big advantage of Gateway cloning in comparison to classical cloning is the opportunity to enable cloning at sites where no suitable restriction enzymes are available. Additionally, this method allows a simplified transfer of cDNA inserts between a variety of destination vectors for different purposes. After recombination of an attB-flanked PCR product with an attP-donor vector (BP-reaction), an attL-including entry clone as well as an attR-containing by-product is formed. This reaction is catalysed by the BP Clonase<sup>TM</sup> enzyme mix containing bacteriophage λ integrase and E.coli integration host factor. The reaction product is transformed into competent bacteria and selection is performed taking advantage of the kanamycin resistance of the created entry

clones. The recombination of the attL sites of the entry vector with the attR site of the destination vector (LR reaction) is mediated by LR Clonase™ enzyme mix (bacteriophage  $\lambda$  integrase, excisionase and *E.coli* integration host factor) and results in an attB site carrying expression clone and attP1 site containing by-product. The reaction is transformed into competent bacteria and clones selected with ampicillin.

The gene of interest (GOI) was first generated performing PCR (see 3.2.1.5). To link attB sites to selected PCR fragments, another PCR was performed using attB primers (listed in Table 2). Thereafter, the BP-reaction was performed, incubating 3 $\mu$ l of attB-comprising PCR product with 1 $\mu$ l of pDONOR vector (90ng/ $\mu$ l) and 1 $\mu$ l BP-clonase II mix for 4h at 25°C. 0.5 $\mu$ l proteinase K was added to terminate the reaction by incubating the mixture for 10min at 37°C. After transformation and kanamycin selection, plasmid DNA was isolated. For the LR-reaction, 1 $\mu$ l of DNA (100ng/ $\mu$ l) was incubated with 1 $\mu$ l pDEST vector (100ng/ $\mu$ l), supplemented with 2 $\mu$ l Miniprep elution buffer and 1 $\mu$ l LR-clonase II mix for 4h at 25°C. The LR-reaction was terminated adding 0.5 $\mu$ l proteinase K for 10min at 37°C. Thereafter, the complete LR-reaction product was transformed into *E.coli* followed by selection on ampicillin-containing plates.

### **Cloning sgRNA into a vector backbone**

The CRISPR/Cas9 gene editing system (described in 3.2.2.6) relies on a target-specific single-guided RNA (sgRNA) that together with the Cas9 protein forms a ribonucleoprotein complex capable to introduce a double-strand break into DNA at a defined sequence. The plasmid-borne CRISPR/Cas9 system used in this study requires an insertion of the template for the sgRNA into the parental plasmid vector. A fast and easy cloning procedure which combines DNA restriction and ligation was used for this purpose [60].

The annealing product (see 3.2.1.1) was diluted 1:200 in ddH<sub>2</sub>O and incubated with 100ng of the vector, 0.5 $\mu$ l T4 DNA Ligase, 1 $\mu$ l DTT (10mM), 1 $\mu$ l ATP (10mM) and 1 $\mu$ l FastDigest BbsI/BpiI in 20 $\mu$ l 1x DNA Ligase reaction buffer for 1h in a thermocycler with the following cycle: 5min at 37°C followed by 5min at 21°C. To digest any residual linearized DNA, 1.1 $\mu$ l of ligation reaction was treated with 1 $\mu$ l PlasmidSafe exonuclease supplemented with 1.5 $\mu$ l ATP (10mM) in a total volume of 15 $\mu$ l of 1x PlasmidSafe buffer.

### **3.2.2 Mammalian cell culture**

#### **3.2.2.1 Cryopreservation and thawing of cells**

To conserve mammalian cell stocks, cells of a confluent grown 10cm culture dish were washed with dPBS, detached with Trypsin/EDTA and dispersed in growth medium. After centrifugation for 3min at 500xg, the cell pellet was resuspended in 5ml cryo medium (see 3.1.5.2) and distributed into five cryopreservation tubes. A slow freezing protocol was performed to avoid intracellular freezing. First, the aliquots were incubated for 10min at 4°C followed by 2h at -20°C with further overnight incubation at -80°C. For long-term storage, cell stocks were transferred to liquid nitrogen.

Frozen cell stocks were rapidly defrosted in a water bath at 37°C. To prevent DMSO-mediated cytotoxicity, cells were quickly resuspended in prewarmed growth medium and seeded in culture dishes. The following day, when cells have already adhered to the plate, growth medium was exchanged to remove remaining cryo medium and cell debris.

#### **3.2.2.2 Maintenance and growth of cells**

HEK293T cells and hTERT-RPE1 cells were cultured in a humidified atmosphere containing 5% CO<sub>2</sub>. They were grown at 37°C in Dulbecco's Eagle Medium supplemented with 0.5% Penicillin/Streptomycin and 10% FBS until they reached, after 3-4 days, 100% cell confluence. Afterwards, confluent cells were washed with 5ml prewarmed dPBS, detached by incubation with 2ml Trypsin/EDTA for 1-3min at RT and resuspended in prewarmed growth medium. To maintain cells in culture, they were regularly split on new culture dishes in a ratio of 1:10 (or 1:5 for SILAC-labelled cells).

Newly generated Flp-In monoclonal lines as well as manipulated Flp-In monoclonal lines were cultivated in Flp-In Selection Medium (see 3.1.5.2) in a humidified atmosphere containing 5% CO<sub>2</sub>. After 3-4 days, cells were washed with 5ml prewarmed dPBS and detached by incubation for 1-3min with 2ml Trypsin/EDTA solution. Afterwards, cells were resuspended in prewarmed Flp-In Selection Medium and plated on new dishes in a ratio of 1:5 (or 1:3 for SILAC-labelled cells).

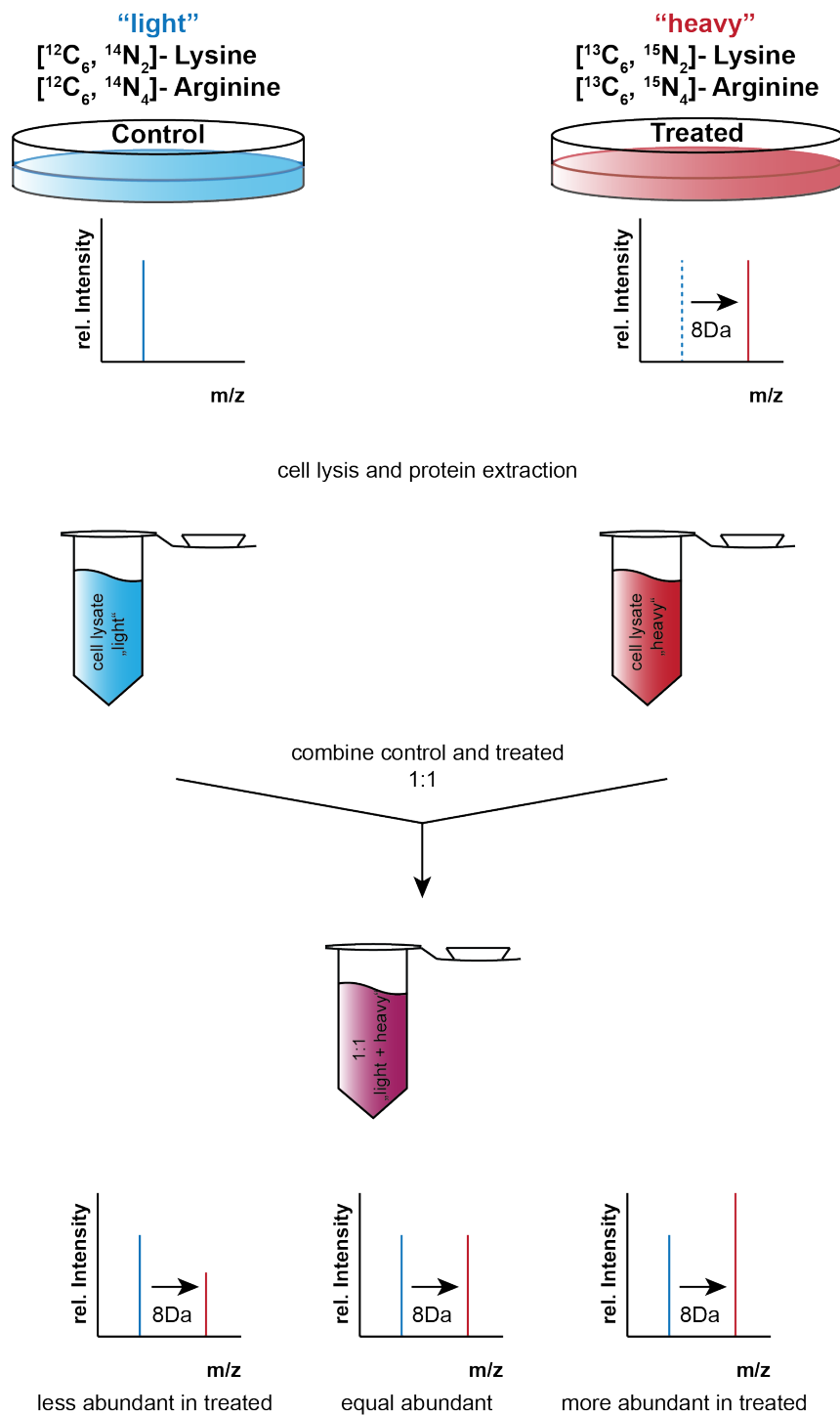
#### **3.2.2.3 Stable isotope labelling of amino acids in cell culture - SILAC**

Stable isotope labelling of amino acids in cell culture (SILAC) is a common approach for metabolic labelling of mammalian cell lines. This approach is well-established for quantitative mass spectrometry and is based on the incorporation of stable isotopically labelled amino acids into the proteome of the cell. As shown in Figure 7, this method can be used to pool samples of different labels to compare the proteome of different cell conditions. For a SILAC approach,

growing cells are crucial and at least five cell cycles are required for a complete incorporation. A complete incorporation of isotopically labelled and essential (semi-essential) amino acids results in a defined mass shift as depicted in Figure 7. Typically, lysine and arginine are used as labelled amino acids resulting in at least one incorporated isotopically labelled amino acid at the end of each tryptic peptide [41].

For SILAC labelling, cells were cultured in SILAC DMEM supplemented with heavy isotope-substituted lysine and arginine (see 3.1.5.2). Introduction of 'heavy' lysine results in a mass shift of 8Da compared to the light amino acid, whereas the chosen 'heavy' arginine leads to a specific mass shift of 10Da. To avoid arginine to proline conversion, 2mM proline was added to the medium [62, 63].





**Figure 7: Stable isotope labelling of amino acids in cell culture (SILAC)**

For metabolic labelling of amino acids in cell culture, mammalian cells are cultured in SILAC medium containing isotopically labelled lysine and arginine. To ensure a complete incorporation of both labelled amino acids, at least five cell cycles are necessary. Combining cells with two different labels enables the quantitative comparison of two conditions based on the introduced and defined mass shift which can be distinguished using mass spectrometry. E.g. the mass shift of 8Da is based on the incorporation of one ‘heavy’-labelled lysine in the analysed tryptic peptide.

#### 3.2.2.4 Transfection of mammalian cell lines

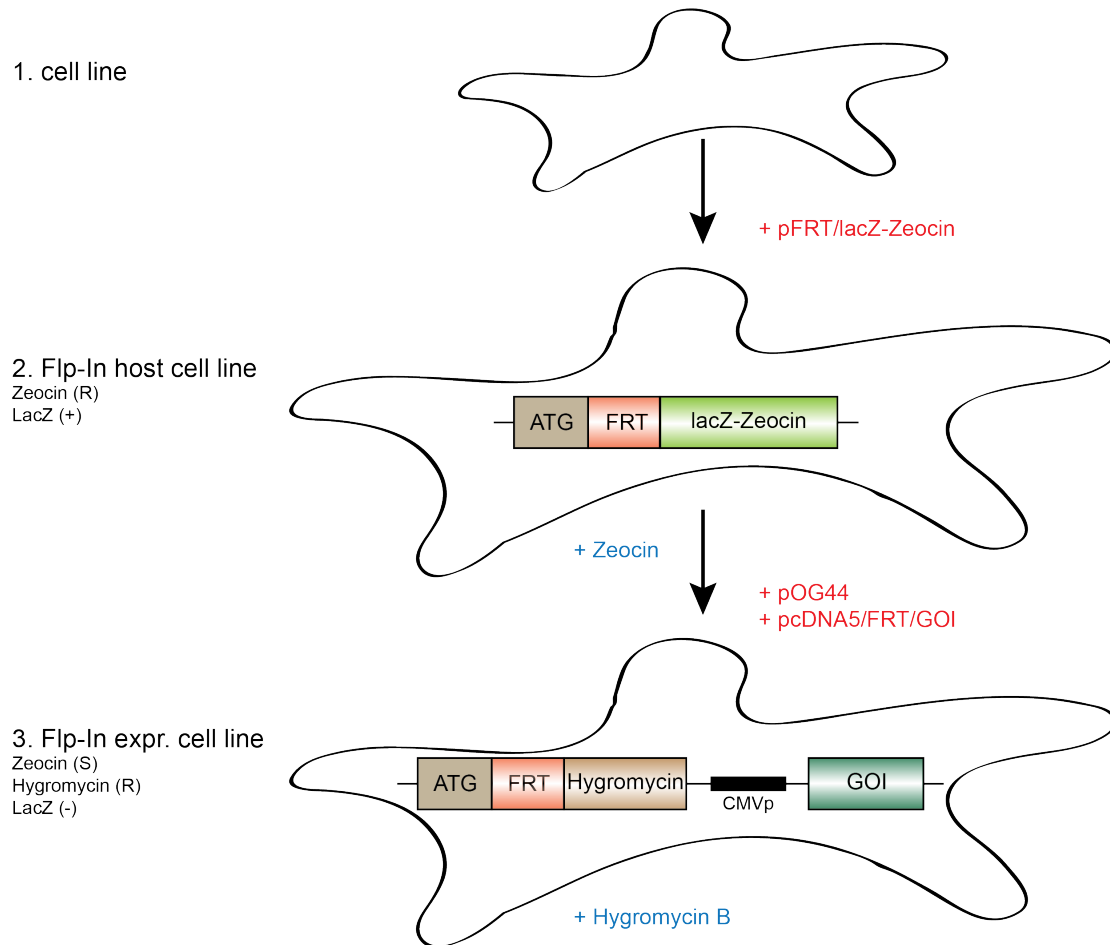
Transfection of mammalian cell lines is a powerful analytical method to study protein function, generating genetically modified cells by the introduction of foreign nucleic acids into the mammalian cells either in a transient or a stable way. In transiently transfected cells, a transgene is only expressed for a limited time period, whereas in stably transfected cells the transgene is integrated into the genome [64].

To enable transfection of mammalian cell lines, 50-80% confluency of cells was achieved by seeding the cells one day before transfection. For the amount of cells of one 14cm culture dish, 8µg DNA was diluted in 1ml PEI transfection reagent (1mg/ml) (see 3.1.5.2). Complex formation of DNA and PEI was initiated by mixing and incubating the solution for 10min at RT. Afterwards, DNA-PEI mixture was added dropwise to the cells. After transient transfection, cells were cultured for another two days before they were harvested.

#### 3.2.2.5 Generation and maintenance of stable expression cell lines (Flp-In)

Using the Flp-In™ system of Invitrogen for the generation of stably transfected cell lines allows the integration and expression of a gene of interest (GOI) at a specific genomic location. After introducing a Flp Recombination Target (FRT) site into the genome of the chosen mammalian cell line, the GOI, inserted in an expression vector, is integrated via Flp recombinase-mediated DNA recombination [65].

For the generation of isogenic stable cell lines using the Flp-In™ system, three different vectors were used. First, the pFRT/lacZeo target site vector was used to generate a Flp-In™ host cell line. In the FRT/lacZeo construct, the ATG initiation codon is followed by the FRT site which serves as binding and cleavage site for the recombinase (see Figure 8). To verify a successful integration of the pFRT/lacZeo plasmid, mammalian cells are selected for zeocin resistance and expression of lacZ can be assayed by western blot followed by immunostaining. To integrate the gene of interest into the parental Flp-In™ host cell line, the chosen GOI has to be cloned into a pcDNA5/FRT expression vector. This vector additionally contains a hygromycin resistance gene with an implemented FRT site. After cotransfection of this expression vector and pOG44, which expresses the Flp recombinase, a homologous recombination between the FRT sites takes place. This leads to a Flp-In™ expressions cell line which can be selected for hygromycin resistance, zeocin sensitivity or lack of β-galactosidase activity. An inserted human CMV promoter is responsible to control and drive the expression of the GOI.



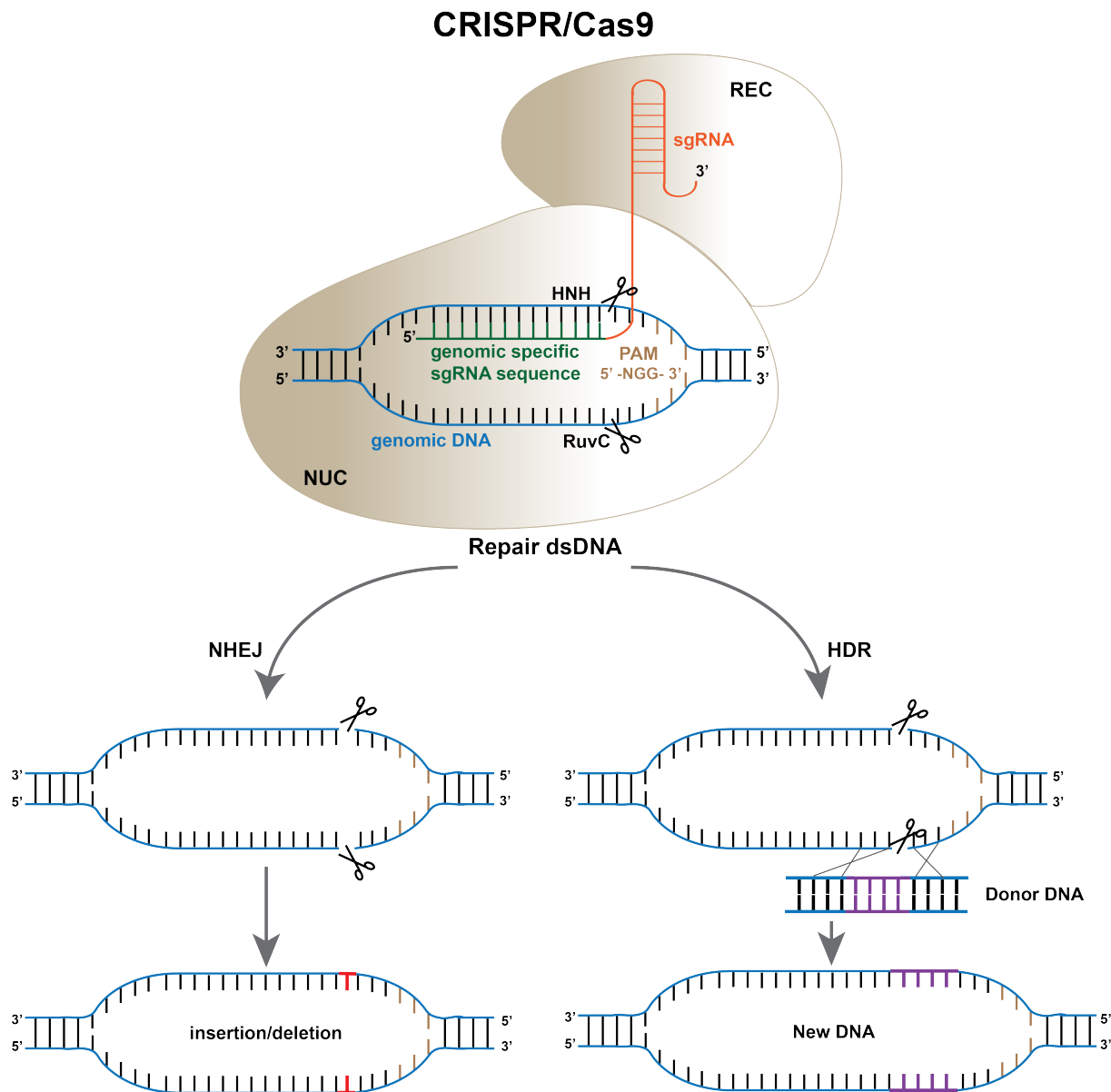
**Figure 8: Generation of stable cell lines using the Flp-In system**

First, mammalian cell lines were transfected with a FRT/lacZ-Zeocin target site vector to generate the Flp-In host cell line which is characterized by zeocin resistance. To introduce the gene of interest into the genome of the generated Flp-In host cell line, co-transfection with pOG44 (recombinase) and pc/DNA5/FRT/GOI was performed. Generated Flp-In expression cell lines are characterized by zeocin sensitivity, hygromycin B resistance as well as by the lack of  $\beta$ -galactosidase expression.

Flp-In 293 host cells were transfected in 10cm culture dishes using 0.8 $\mu$ g of either pcDNA5/FRT/(N)-SF-IFT122, pcDNA5/FRT/(N)-SF-TULP3 or pcDNA5/FRT/(N)-SF-LCA5 and 7.2 $\mu$ g of pOG44 (FLP recombinase) by means of lipofection (lipofectamin2000) according to manufacturing recommendation. Control transfections were done with 0.8 $\mu$ g of expression vector but without pOG44. Medium was changed after 24h post transfection. After two days, cells were split in a ratio of 1:4 on 10cm dishes using growth medium supplemented with 100 $\mu$ g/ml hygromycin for selecting only cells carrying the plasmid which contains the hygromycin resistance gene. Two weeks later, single clones were picked. Each clone was transferred to one well of a 24-well plate containing 1ml growth medium supplemented with 100 $\mu$ g/ml hygromycin. Cells were expanded stepwise to 6-wells and 10cm culture dishes. Stable expression level of the transgene was tested by western blot analysis (see 3.2.3.4).

### 3.2.2.6 Gene editing using the CRISPR/Cas9 system

A new technology for gene editing is the innovative CRISPR/Cas9 system. This technique uses a targeted nuclease which facilitates genome alterations at a defined site.



**Figure 9: Gene editing using the CRISPR/Cas9 system**

Specific sgRNA, containing a 20-nucleotide guide sequence, targets the nuclease Cas9 (NUC) to a specific DNA sequence. After the guide sequence has paired with the DNA target, Cas9 generates a double-stranded break 3bp upstream of the protospacer adjacent motif (PAM). After a nuclease-induced double-stranded break occurs, two types of DNA damage repair mechanisms can follow: The error-prone non-homologous end joining (NHEJ) or the high-fidelity homology-directed repair (HDR) pathway. In the NHEJ pathway, DNA repair is processed and re-joined by endogenous DNA repair machinery. This method can lead to random mutations which can result for example in gene knockout. To achieve a precise editing of the genomic DNA, a repair template (plasmid or ssODN) can be supplied to induce HDR.

---

In this system, Cas9 nuclease is guided by a short RNA molecule (sgRNA) to generate a precise double-stranded break at a target DNA locus. As shown in Figure 9, DNA damage repair can occur through two different pathways: First, error-prone non-homologous end joining (NHEJ) which leads to random mutations (e.g. knock-out, non-mediated decay) or second, high-fidelity homology-directed repair (HDR) [60].

Flp-In cells expressing (N)-SF-TAP-tagged TULP3 (see 3.2.2.5) were transfected with an bicistronic vector expressing a specifically designed sgRNA sequence as well as Cas9 alone or together with an additional HDR construct. As a control, an additional transfection was done with the 'empty' vector (Cas9 gene only). For transfection, 500ng DNA was mixed with 66µl PEI and added dropwise to the cells. After 48h, selection was initiated adding DMEM supplemented with 8µg/ml Puromycin and cells were cultivated for additional 48h. Subsequently, cells were grown for some days in DMEM and transferred to 10cm culture dishes as well as seeded at very-low density (1:400) into a 14cm culture dish for single-clone selection. For single-clone selection, cells were washed and trypsinated using 20ml trypsin-solution (1:20 in PBS). After some minutes, cells started to detach and could be soaked and transferred to 24-well plates. Some weeks later, single clones were characterized by PCR.

### **3.2.3 Protein chemistry**

#### **3.2.3.1 Protein extraction**

To extract proteins from a 14cm dish of cultured cells, growth medium was removed and cells were washed with dPBS. Plates were placed on ice and 1ml lysis buffer (see 3.1.5.6) was added. Cells were detached using a cell scraper and the suspension was transferred to reaction tubes followed by incubation for 20-30min at 4°C under constant agitation. To separate cell debris from clear protein extract, lysates were centrifuged at 10,000xg for 10min at 4°C. The supernatant, containing the clear protein extracts, was transferred to fresh tubes. Protein concentration was determined using a Bradford assay (for more detail see below in 3.2.3.2) and lysates stored at -80°C for long-term storage or directly used for further experiments.

#### **3.2.3.2 Quantification of protein concentration**

The total amount of protein was determined using a Bradford assay. This colorimetric assay is based on an absorbance shift of the dye Coomassie brilliant blue G250. The absorbance maximum of the pure dye is at 465nm. After complex formation with proteins, this maximum is shifted to 595nm [44]. In this work, the commercial Bio-Rad protein assay kit was used. The Dye Reagent Concentrate was diluted 1:5 with ddH<sub>2</sub>O and 1ml of the diluted reagent was added to an appropriate amount of protein sample. For determination, a standard curve with known concentrations of BSA was generated. Therefore, standards of 2-8µg BSA were prepared,

diluted in the corresponding cell lysis buffer and the absorption was subsequently measured at 595nm using the diluted dye as reference.

### 3.2.3.3 SDS-PAGE

SDS-PAGE is a widely used technique to separate proteins according to their molecular weight. SDS represents an anionic detergent which binds with a ratio of 1.4g SDS per gram protein. As a result, the intrinsic charge of the protein is effectively masked by the negative charge of the detergent. Due to this nearly constant charge-to-mass ratio of the SDS-protein complexes these complexes can be separated according to their molecular masses by electrophoreses in a SDS-containing polyacrylamide gel. In the SDS-PAGE approach, small proteins are moving faster and farther through the gel than larger molecules.

Gels were casted using the Mini-Protein 3 system chambers with 0.75mm spacers. Solutions for the stacking as well as the separation gel were prepared as given in 3.1.5.4. The separation gel solution was filled between the glass plates first and covered immediately with isopropanol. Polymerisation of acrylamide/bisacrylamide is initiated by ammonium persulfate (APS) and catalysed by N,N,N',N'-tetramethylethylenediamine (TEMED). After 30min, isopropanol was removed and stacking gel solution was casted on top. For sample separation, a comb with the desired well number was inserted before polymerization. Gels were placed into gel chambers, filled with running buffer and combs were removed. Defined protein amount (5-10 $\mu$ g) was mixed with 5x Laemmli buffer (see 3.1.5.4) and denatured for 2min at 96°C and cooled down. The sample as well as 7 $\mu$ l of a size standard (PageRuler 170kDa/ 250kDa) was loaded on the stacking gel. Electrophoresis was initiated by applying 80V until all proteins entered the separation gel. After that, voltage was increased to 120V and separation was performed until the bromophenol blue dye reached the bottom of the separation gel. For subsequent staining, proteins were fixed using EtOH/HAc/ddH<sub>2</sub>O mixtures.

### Coomassie Staining of SDS-gels

As far as sufficient amounts of protein (200-400ng per lane) are used, Coomassie staining is the method of choice because it is simple and cheap. Nevertheless, the time-consuming destaining of Coomassie stained gels is a big disadvantage [44].

After fixation of the proteins by incubating the gel three times for 10min in Fixer-Solution, staining was performed for less than 10min in Coomassie solution (see 3.1.5.4). To lower down background staining, the gel was washed for three times with Fixer-Solution for 10min each.

### Silver staining of SDS-gels

Silver staining enables the detection of low amounts of protein (1-10ng per lane) [44]. Thereby,  $\text{Ag}^+$ -ions are forming complexes with Glu, Asp and Cys-residues which, after incubation with alkaline formaldehyde, are reduced to elementary Ag [44].

For silver staining of proteins, SDS-PAGE gels were fixed two times using Fixer-Solution for 30min each. Subsequently, gels were transferred to 50% EtOH and incubated two times for 20min followed by some minutes in  $\text{Na}_2\text{S}_2\text{O}_3$ -Solution. After briefly washing the gel two times in ddH<sub>2</sub>O, the gels were incubated for 20min in  $\text{AgNO}_3$ -Solution. Gels were washed again in ddH<sub>2</sub>O before proceeding to the developing step (2-5min incubation in Developer-Solution). To stop the reaction, Coomassie-Fixer was added and the gel fixed for 10min under constant agitation. For long-term storage and preservation, gels were finally incubated two times for 10min in Preserver-Solution. Recipes of all buffers and solutions are listed in 3.1.5.5.

#### 3.2.3.4 Western blot analysis and immunostaining

To immobilize proteins after SDS-PAGE separation, they can be transferred to nitrocellulose or polyvinylidene fluoride (PVDF) membranes. Subsequently, antibodies can be used to bind appropriate antigens immobilized on the membrane and enable the detection through enzymatic activities coupled to the primary or the secondary antibody [66].

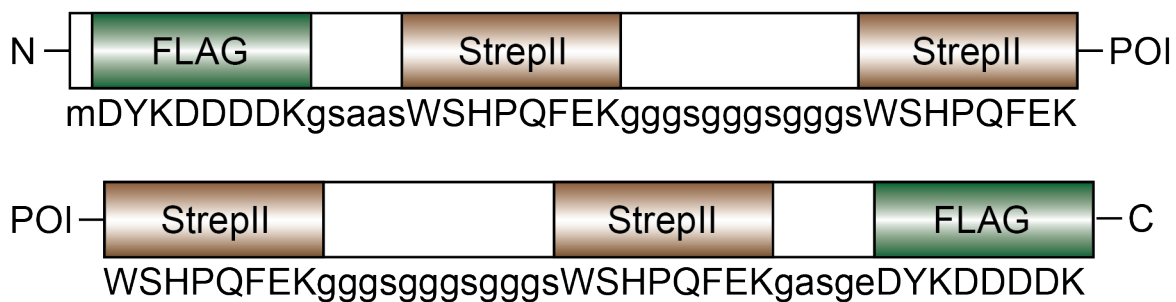
After SDS-PAGE, proteins were immobilized onto PVDF membranes using tank blotting. First, the black side of the holder cassette was covered with one soaked (in western buffer, see 3.1.5.4) sponge followed by two soaked whatman papers. Subsequently, the SDS-polyacrylamide gel was placed on top of the two whatman papers. The PVDF membrane was activated for at least 30sec in methanol before it was placed on the gel. Two additional soaked whatman papers, followed by one soaked sponge, were put on top. To remove air bubbles, a blot roller was used in between every step. After closing and placing the cassette into the blotting chamber, the chamber was filled with western buffer. Western blotting was performed at 90V for 90min at 4°C. To visualize immobilization of proteins on the PVDF membrane, proteins can be stained using Ponceau-Solution (see 3.1.5.4) and destained again by incubation for 10min in TBST. After blocking the membrane for at least 30min using Blocking solution (3.1.5.4), membrane was incubated either over night with a protein specific primary antibody or for 1.5h with a HRP-coupled primary antibody recognizing the FLAG-epitope. Afterwards, the membrane was washed three times with TBST. Once a protein specific primary antibody was used, the blot was further incubated for 1h with a HRP-coupled secondary antibody which recognises the specific  $F_c$  portion of the primary antibody. Before the detection was carried out

using ECL or ECLplus system according to the manufacturer's instructions, the membrane was washed again for three times with TBST. The ECL and ECLplus systems are based on chemiluminescence which is generated by the reaction of horseradish peroxidase (HRP) with luminol, present in ECL reagent, and was detected on light-sensitive films (hyperfilm, GE Healthcare).

### 3.2.4 Analysis of protein-protein interactions

#### 3.2.4.1 Tandem Affinity Purification (TAP)

To study protein interactions, Strep FLAG tandem affinity purification (SF-TAP) is a common and well described method [67]. Overexpressed proteins are fused to a SF-TAP tag with a molecular weight of 4.6kDa. As shown in Figure 10, this SF-TAP tag consists of a Strep-tag II as well as a FLAG tag. While the N-terminal SF-TAP tag starts with FLAG tag followed by two StrepII tags, the C-terminal (C)-SF-TAP tag starts with the tandem StrepII tag followed by the FLAG tag fused to the C-terminus of the POI [68].

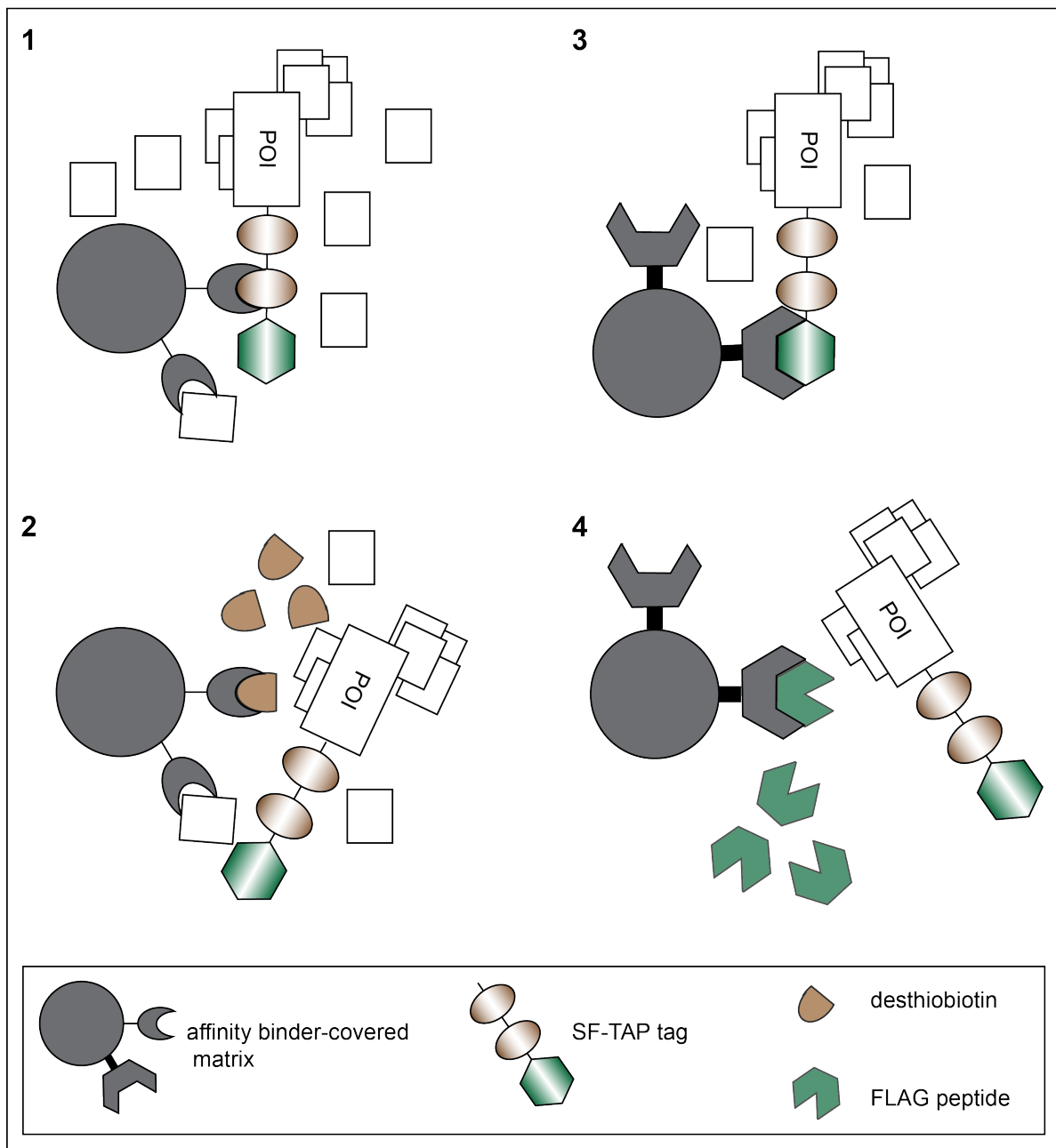


**Figure 10: Protein sequence of (N)/(C)-SF-TAP tag**

For tandem affinity purification (TAP), so called Strep-FLAG tags are fused to the protein of interest (POI) by means of fusion the open reading frame of the POI with the nucleotide sequence encoding the SF-TAP tag. Through these tags the molecular weight of the fusion protein increases by 4.6kDa. The structure of the N-terminal SF-TAP tag starts with FLAG tag followed by two StrepII tags, whereas the (C)-SF-TAP tag starts with the tandem StrepII tag followed by the FLAG tag fused to the C-terminus of the POI.

For affinity purification of proteins using the Strep-tag II, Strep Tactin sepharose beads can be used and the bound proteins can be eluted with D-Desthiobiotin. In contrast, FLAG-tagged proteins bind to anti-FLAG-M2 agarose and elution can be done using FLAG peptide (see Figure 11).





**Figure 11: Tandem Affinity Purification (SF-TAP)**

To purify protein complexes of interest and to study protein-protein interactions, tandem affinity purification (TAP) can be performed. Thereby, overexpressed proteins that are fused to a Strep-FLAG TAP tag either on the N-terminus or the C-terminus of the protein of interest (POI) and their interaction partners can be purified. In the TAP approach two purification steps are performed sequentially. First, Strep-Tactin sepharose beads are used to purify the POI (1). After washing and eluting bound proteins using D-Desthiobiotin (2) a second purification step is performed using FLAG-M2-agarose beads (3), recognizing the FLAG tag of the POI. To elute bound proteins, a FLAG peptide with a higher affinity to the FLAG tag is added to the reaction (4).

---

Cell lysates were incubated with a corresponding volume of packed Strep-Tactin sepharose beads (25 $\mu$ l pure beads per 14cm culture dish) for 1.5h at 4°C on an end-over-end shaker. Before use, Strep-Tactin sepharose beads were washed once using 500 $\mu$ l Lysis Buffer followed by two washing steps using 500 $\mu$ l Washing Buffer, each. Subsequently, beads were centrifuged for 30sec with 5,000xg at 4°C and two thirds of the supernatant were discarded before beads were transferred to 1.5ml reaction tubes comprising the cell lysates. After the incubation step, beads and bound proteins were transferred to microspin columns and washed three times using 500 $\mu$ l Washing Buffer (see 3.1.5.6) by pushing through all remaining liquid. To elute bound proteins, spin columns were closed with a plug. 100 $\mu$ l Strep-Elution Buffer per cells of one 14cm culture dish were added to the beads (still remaining in the microspin columns) and incubation was done for 10min at 4°C. During this elution step, beads were manually mixed every two minutes. For elution of bound proteins, the plug of the spin columns was removed and columns were transferred to fresh 2.0ml reaction tubes before centrifugation was performed for 1min at 1,000xg at 4°C. Afterwards, a corresponding volume (12.5 $\mu$ l pure beads per 14cm culture dish) of washed anti-FLAG-M2-agarose beads (washing procedure similar to Strep-Tactin sepharose beads described above) was transferred to Strep-Tactin eluates and incubated for 1.5h at 4°C on a shaker. Then, beads were washed for three times using 500 $\mu$ l Washing Buffer and bound proteins were eluted by incubating the FLAG-M2-agarose beads with 200 $\mu$ l Flag peptide for 10min at 4°C followed with constant agitation followed by centrifugation at 1,000xg for 1min at 4°C.

#### **3.2.4.2 One-step FLAG affinity purification**

This one-step affinity purification protocol using anti-FLAG-M2-agarose beads (which is also a part of the SF-TAP approach described in 3.2.4.1) was published previously by our group [69]. This approach is based on the pulldown of proteins tagged with a FLAG-tag using anti-FLAG-M2 agarose followed by elution of the bound proteins using FLAG peptide.

After protein extraction from cell cultures of two confluent 14cm culture dishes per sample, the cleared cell lysate was incubated with washed (see 3.2.4.1) anti-FLAG-M2-agarose beads (12.5 $\mu$ l pure beads per 14cm dish) for 2h at 4°C on a shaker. Beads were then centrifuged for 1min at 5,000xg, transferred to micro spin columns and washed three times with 500 $\mu$ l Washing Buffer each. To elute bound proteins, spin columns were closed and beads were incubated using 200 $\mu$ l FLAG peptide per sample for 10min at 4°C. After centrifugation (1,000xg, 1min, 4°C), eluates were used for protein precipitation (described in 3.2.4.3) followed by an in-solution digestion (see 3.2.4.4) and targeted mass spectrometry (described in chapter 3.2.6.2).

### 3.2.4.3 Methanol-chloroform precipitation

To precipitate purified proteins, a methanol-chloroform precipitation was performed according to a protocol developed by Wessel and Flugge in 1984 [70]. This method allows an efficient protein precipitation even in presence of detergents and salts. For this approach, the ratio of methanol (MeOH), chloroform (CHCl<sub>3</sub>) and water (H<sub>2</sub>O) is crucial.

In this work 200µl volumes of protein containing solution were used for protein precipitation. Thereby, 800µl MeOH were added to 200µl aqueous sample and mixed by vortexing. After centrifugation at 9,000xg for 30sec at RT, 200µl CHCl<sub>3</sub> were added and the solution was vortexed and centrifuged again. Subsequently, the solution was mixed with 600µl HPLC H<sub>2</sub>O, vortexed and centrifuged at 16,000xg for 1min to collect precipitated proteins in the interphase. The upper phase was discarded. After the addition of 600µl MeOH and vortexing, proteins were pelleted via centrifugation for 2min at 16,000xg at RT. Protein pellet was dried for further sample preparation.

### 3.2.4.4 In-solution tryptic proteolysis

To perform mass spectrometric-based proteomics, small peptides instead of large proteins are necessary. Therefore, enzymatic digestion of precipitated proteins was carried out as described previously [69]. A commonly used enzyme for proteolysis is trypsin which is a pancreatic serine protease that hydrolyses peptide bonds at the carboxyl group of arginine and lysine except for the presence of a proline residue linked to this very carboxyl group. Prior to the tryptic digestion, samples are treated with dithiotreitol (DTT) to cleave disulfide bridges of cysteine-containing peptides and resulting cysteine residues are alkylated by 2-iodoacetamide (IAA).

Precipitated proteins (10-20µg) were dissolved in 30µl 50mM ammonium bicarbonate (ABC) and 4µl RapiGest, a surfactant that improves solubility of the precipitated proteins, by vortexing harshly. Subsequently, 1µl DTT (100mM) was added and the solution was mixed and incubated for 10min under constant agitation at 60°C. After the sample was cooled down, 1µl IAA (300mM) was added and incubation was carried out for 30min at RT in complete darkness. For enzymatic digestion of proteins, 2µl trypsin solution (1µg/µl in 1mM HCl) was added and proteins were digested overnight at 37°C. The day after, enzymatic cleavage was stopped by adding 1.7µl TFA (100%). The whole solution was transferred to polypropylene inserts and incubated for 10min at RT. To obtain a clear peptide solution, centrifugation was performed for 15min at 16,000xg and the supernatant was used for further experiments.

### 3.2.4.5 Desalting via stop-and-go extraction tips

To prepare peptide mixtures for mass spectrometry analysis, it is necessary to remove interfering agents like detergents and salts. For this step, stop-and-go extraction tips (StageTips) can be used according to the protocol reported by Rappsilber and colleagues [71]. These tips are filled with C18 matrix with a binding capacity of 10µg of peptides. While peptides adsorb to the C18 material, salts and detergents pass through. For the elution of bound peptides, organic solvents can be used.

StageTips were equilibrated with 20µl 80/5 solution (80% ACN, 5% TFA) followed by a washing step with 20µl 0/5 solution (5% TFA). After loading the acidified sample onto the C18 material, the tip was washed again with 20µl 0/5 solution. Peptides were eluted using two sequential elution steps. First, with 20µl 50/5 solution (50% ACN, 5% TFA) followed by 20µl 80/5 solution. To remove organic solvents and to concentrate purified peptides, the sample was evaporated to 2-5µl using a vacuum centrifuge. Prior to LC-MS/MS analysis, sample was filled up to 15µl using 0.5% TFA.

### 3.2.4.6 In-gel digestion of silver stained proteins

In-gel tryptic proteolysis was performed to prefractionate samples before LC-MS/MS analysis as described before [69].

After silver staining of SDS-PAGE-separated proteins (3.1.5.4), stained protein bands were cut out and sliced into 2-4 small pieces. Pieces were transferred to a 96 well plate and wells filled with HPLC H<sub>2</sub>O. To destain gel pieces, 100µl of a 1:1 ratio of potassium hexacyanoferrat II (30mM) and sodium thiosulfate (100mM) were added to each well and incubation was performed for 5min at RT on a shaker. The destaining solution was discarded and each well was washed three times for a few seconds with 100µl HPLC H<sub>2</sub>O to remove remaining destaining solution. For enzymatic digestion of proteins, gel pieces were incubated with 100µl DTT (5mM) per well for 30min at 60°C under agitation before solution was discarded. Then 100µl IAA (25mM) were added. After incubation for 45min at RT in darkness, IAA solution was removed and gel slices were washed for 5min using 100µl ACN (40%) per well. This step was repeated once and subsequently 100µl ACN (100%) were added for 2min. After gel slices were dried under laminar flow, 40µl trypsin solution (0.01µg/µl in 50mM ABC) were added to each well and tryptic cleavage was performed overnight at 37°C. The next day, enzymatic cleavage was quenched by adding 10µl TFA solution (2.5%) for 15min. The supernatant was transferred to reaction tubes. Gel slices were further incubated for 15min with 80µl 50/0.5 solution (50% ACN, 0.5% TFA) and the supernatant was added to the supernatant of the previous step. After adding 80µl 100/0.5 solution (100ACN, 0.5% TFA) to the gel pieces for additional 15min, the

---

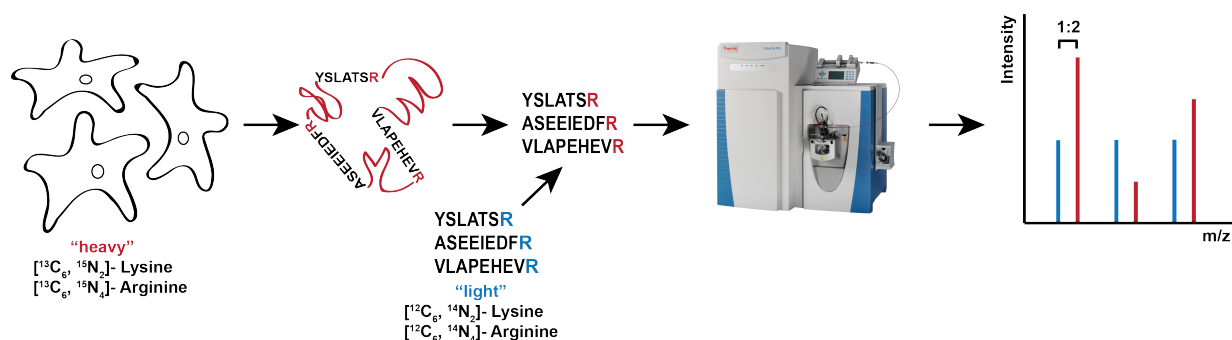
supernatant was added to the pooled sample. To remove organic solvent, the volume of the samples were reduced to 2-5 $\mu$ l using a vacuum centrifuge. For LC-MS/MS analysis, samples were resuspended in 0.5% TFA and filled up to a total volume of 15 $\mu$ l. Afterwards, half of the sample was injected per MS run.

### **3.2.5 Stoichiometric and structural investigations of IFT-A**

#### **3.2.5.1 Absolute quantification of purified IFT-A**

In disease state, protein complexes may change their composition as well as the ratio of the components within the protein complex. To understand the underlying mechanism of diseases, alterations in protein complex stoichiometry must be taken into consideration. To properly address these alterations, relative quantification of proteins and protein complexes is insufficient. In contrast, absolute quantification enables the determination of the absolute amount of a protein of interest in a complex sample instead of comparing different protein amounts in different samples or under different conditions. Performing absolute quantification is more demanding than relative quantification in both the sample preparation and in the data analysis. To achieve an accurate quantification, which is necessary for applications like stoichiometry determination or absolute quantification a targeted mass spectrometry approach is highly beneficial (for more detail see 3.2.6.2). Therefore, the combination of an internal standard mixture with targeted mass spectrometry is the chosen method for the absolute quantification of the ciliopathy-related protein complex IFT-A. No matter what type of targeted mass spectrometry technique is used, it always starts with the selection of a target protein set based on the objective. Tryptic digestion of the chosen target proteins leads to multiple peptides. At least two representative peptides per protein are chosen to generate reliable data for the absolute quantification [72]. A careful selection of representative peptides is essential: The peptides must be proteotypic, which means that they are specific for only one protein and ideally between 8-15 amino acid residues large. Additionally, it is useful to select peptides which are most likely being observed in further shotgun experiments analysed on the same mass spectrometer instrument. Peptides containing residues which are prone to modification like oxidation, deamidation or phosphorylation should also be avoided. To choose target proteins as well as precursor ions, a data-dependent shotgun analysis which is commonly performed in mass spectrometry based proteomic studies, is used [45]. Selected representative peptides are then synthesized and combined to create a standard mix of known peptide concentration. After spiking in a known amount of the standard mix into an isotopically coded biological sample containing the purified protein complex of interest, targeted mass spectrometry is performed to quantify each of the chosen peptides as shown in Figure 12. Calculating the ratio of analysed peak-areas of isotopically coded and non-coded peptide leads to the absolute amount of each

peptide in the biological sample. Hence, the stoichiometry of the protein complex of interest can be determined by normalising the ratios to the least abundant protein, which is set to 1.



**Figure 12: Workflow for absolute quantification**

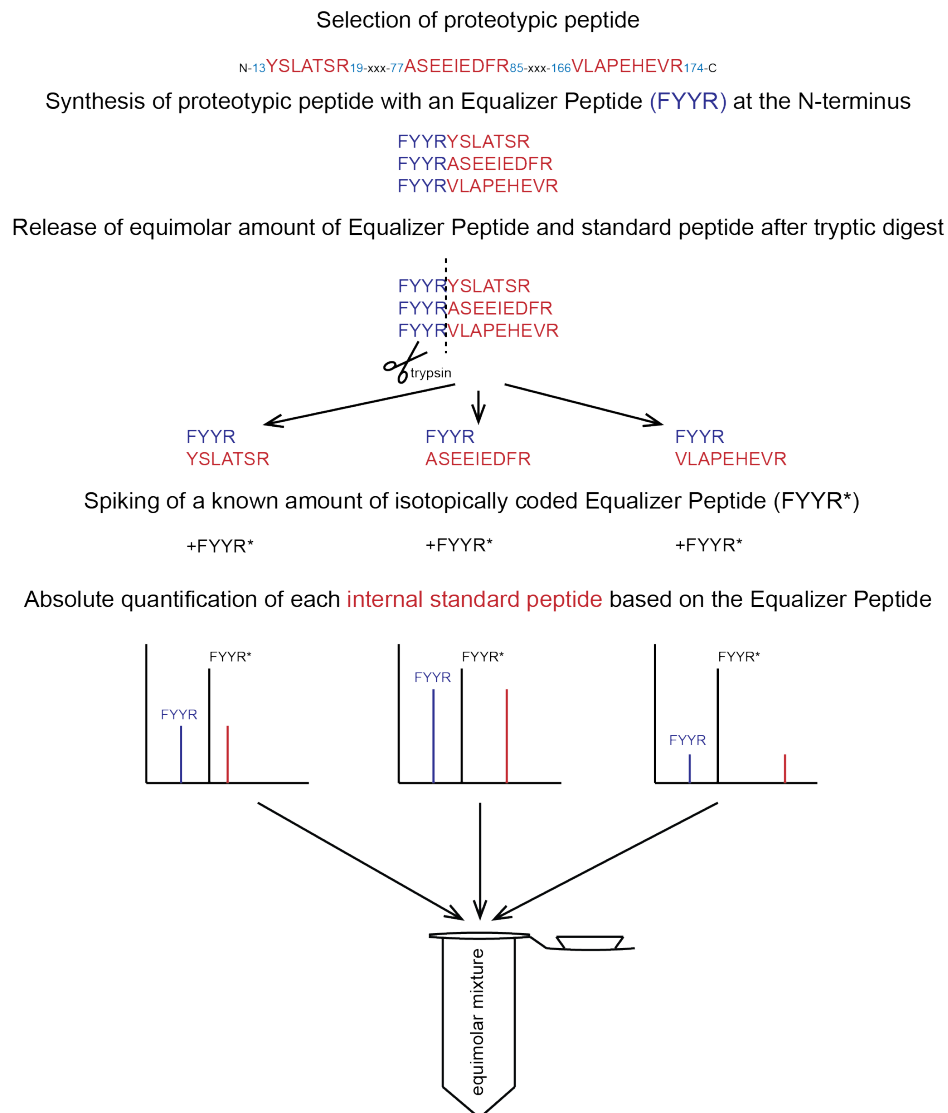
To calculate the absolute amount of a protein complex, mammalian cell lines were cultivated in SILAC media to introduce an isotopic label. After affinity purification of the protein complex of interest and further tryptic digestion, a known amount of non-coded synthetic peptides was spiked into the biological sample. After targeted mass spectrometry analysis, the ratio of 'light' to 'heavy' peptide was calculated and then used to determine the absolute amounts of the selected proteins and the stoichiometry of the protein complex of interest.

### Generating the standard peptide mix

For the absolute quantification of IFT-A, an internal standard mixture containing at least three proteotypic peptides for each of the six complex components was generated. These peptides need to be specific for the given protein (proteotypic), they should not contain amino acids which are prone to post-translational modification and they need to be identifiable by mass spectrometry. To see which peptides are feasible to mass spectrometry analysis, shotgun experiments of affinity purified IFT-A were performed as described in 3.2.6.1 (for sample preparation see 3.2.4.2). The obtained peptides were evaluated with respect to their specificity as well as to their amino acid composition.

#### *Equimolar peptide mix*

To generate an internal standard mix for the use of absolute quantification, it is necessary that the absolute amount of each peptide within the mix is known. Since the determination of peptide amount using amino acids analysis is quite expensive, a standard mix was generated applying the economic "Equimolarity through Equalizer Peptide" protocol which was published by Holzmann et al. [59].



**Figure 13: Generating equimolar standard mixture performing EtEP**

To generate an equimolar mixture of standard peptides, the economic “Equimolarity through Equalizer Peptide” (EtEP) was used. Selected proteotypic peptides were synthesized containing an “Equalizer Peptide” (EP) sequence (FYYR) at the N-terminus. Additionally, the sole EP was also ordered in an isotopically coded manner (FYYR\*) and its amount was analysed via amino acid analysis. Due to the introduced and artificial tryptic cleavage site, enzymatic digestion using trypsin leads to a release of equimolar amounts of the EP and of the corresponding standard peptide. A known amount of the isotopically coded EP was spiked into each peptide solution. Afterwards, the ratio of labelled to unlabelled EP was used to quantify the absolute amount of each representative peptide. Finally, a standard mixture comprising equimolar amounts of each peptide was generated.

As shown in Figure 13, selected peptides were ordered containing an “Equalizer Peptide Sequence” (FYYR-) at the N-terminus of each peptide. Additionally, the sole “Equalizer Peptide” (EP) was ordered in an isotopically coded manner with a known absolute amount (determined via amino acid analysis). An artificial tryptic cleavage site was introduced between the

---

“Equalizer Peptide Sequence” (FYR-) and the proteotypic peptide sequence. After tryptic proteolysis, this tryptic cleavage site leads to a release of an equimolar amount of the EP and the corresponding proteotypic peptide. Then, a known amount of the isotopically coded EP was spiked into each individual peptide solution. After targeted mass spectrometry was performed, the amount of each proteotypic peptide was determined by the ratio of isotopically labelled to unlabelled EP. This ratio was used to quantify the absolute amount of each proteotypic peptide in the corresponding stock solution. Finally, a standard mixture comprising equimolar amounts of each proteotypic peptide was generated.

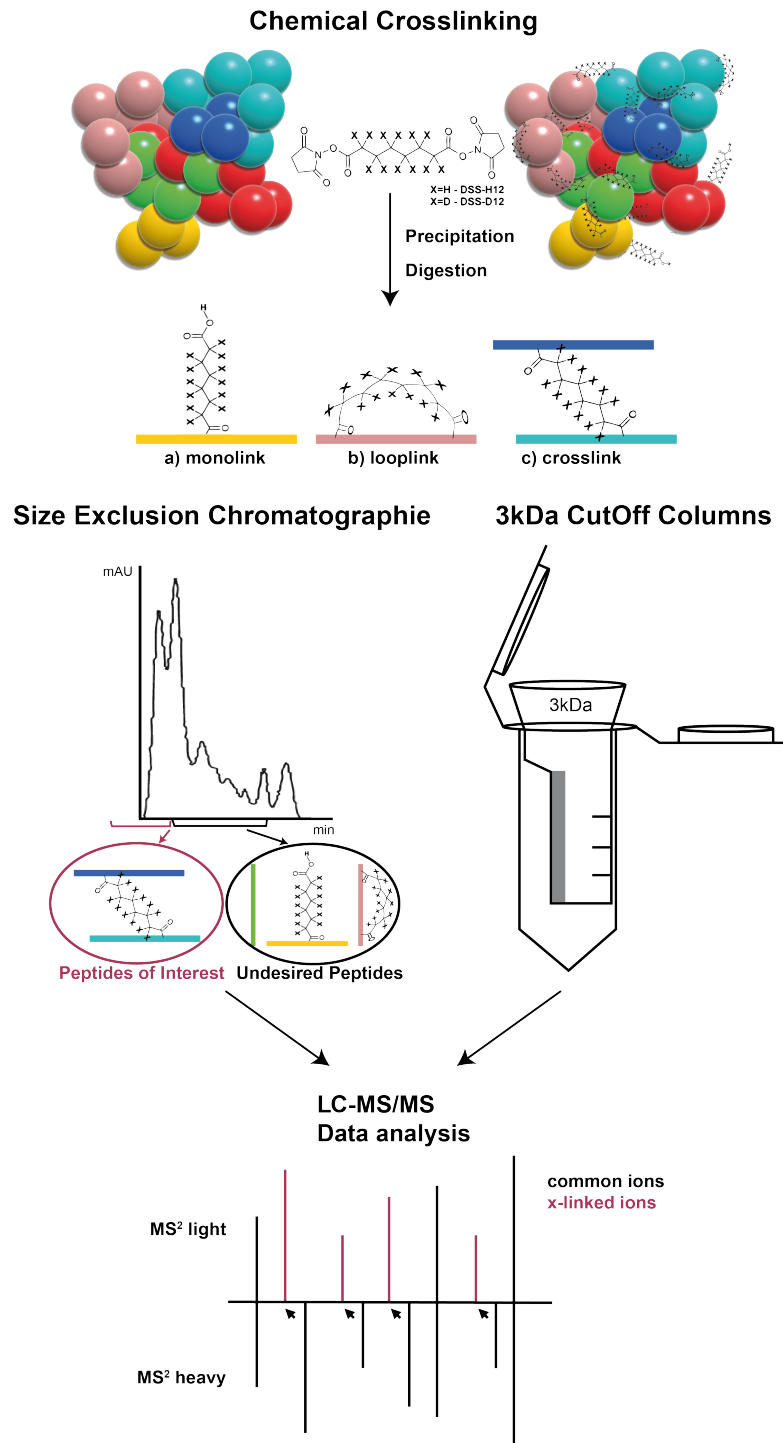
#### *Adjusted peptide mix*

To avoid differences in the dynamic range of quantification upon targeted mass spectrometry, another standard mix containing adjusted amounts of standard peptides was generated. The amount of each peptide was adjusted to the amount of biological peptide present in the purified sample. Therefore, a known amount of synthetic peptides was spiked into a purified IFT-A sample (see 3.2.4.2). After targeted mass spectrometry followed by data analysis was performed, the ratio of isotopically coded to non-coded peptides was determined to adjust the amount of each peptide present in the adjusted standard mix.

#### **3.2.5.2 Chemical crosslinking of purified protein complex**

Chemical crosslinking of purified proteins is an effective method to study protein-protein interactions and to get structural information of a protein complex of interest. After tandem affinity purification (TAP) of IFT-A, the purified protein complex is chemically cross-linked using the bifunctional crosslinker disuccinimidyl suberate (DSS). This chemical crosslinker contains two lysine-reactive N-hydroxysuccinimide (NHS)-esters. To facilitate the following computational analysis (see 1.2.1.2) a 1:1 ratio of isotopically-coded and non-coded crosslinker is used. After tryptic digestion of cross-linked proteins, the complex sample is prefractionated either using size exclusion chromatography (SEC) as described in 3.2.5.3 or performing 3kDa Cut-Off spin column filtration (see 3.2.5.4). Only early eluting fractions of SEC or the retarded fraction of spin columns, containing large cross-linked peptides, are used for LC-MS/MS analysis. This workflow is shown in Figure 14.





**Figure 14: Workflow of chemical crosslinking of purified IFT-A**

For chemical crosslinking, the purified protein or protein complex of interest was incubated with a defined amount of the chemical crosslinker disuccinimidyl suberate (DSS). After quenching the reaction, cross-linked proteins were precipitated using a general MeOH/CHCl<sub>3</sub> protocol. Subsequently enzymatic cleavage was performed using trypsin. To enrich cross-linked peptides, two different methods were used. Either peptides were fractionated performing size exclusion chromatography (SEC) on a FPLC or sample complexity was reduced using conventional 3kDa Cut-Off spin column. Fractions containing large, cross-linked molecules were used for further LC-MS/MS analysis.

---

Cells from six 14cm culture dishes were harvested and pooled for one analysis. Cells were washed with 5ml dPBS and proteins were extracted using 1ml amine-free Lysis Buffer (see 3.1.5.7 ) per culture dish. In situ lysis was done for 30min at 4°C on an end-over-end shaker followed by centrifugation at 10,000xg for 10min to remove cell-debris. The supernatant, containing the cleared cell lysate, was first incubated with washed anti-FLAG-M2-agarose beads (12.5µl pure beads per 14cm culture dish) for 1.5h at 4°C under constant agitation. Then, beads were centrifuged for 1min at 5,000xg and transferred to micro spin columns. After washing the beads three times with Washing Buffer, the beads were incubated with 200µl amine-free FLAG-peptide for 10min at 4°C to elute bound proteins. The eluate was then incubated with washed Strep-Tactin sepharose beads (25µl pure beads per culture dish) for 1.5h at 4°C under constant agitation. Afterwards, centrifugation and washing was performed as described before and beads were incubated with 4-fold bead volume of amine-free Strep-Elution Buffer for 10min at 4°C and centrifuged for 1min at 1,000xg. To reduce sample volume, solutions were transferred to 30kDa Cut-Off spin columns and centrifuged for 13min at 13,000xg at 4°C. The supernatant was collected and protein concentration was determined performing Bradford Assay (see 3.2.3.2).

Purified protein complexes were chemically cross-linked using a defined amount of disuccinimidyl suberate (DSS) (either 5-, 25-, 40-, 50- or 90-fold excess of protein to chemical crosslinker) under constant agitation for 1h at RT. The reaction was quenched through the addition of Tris/HCl pH 7.5 to a final concentration of 100mM. Cross-linked proteins were precipitated using MeOH/CHCl<sub>3</sub> protocol (see 3.2.4.2) and in-solution digestion using trypsin was performed as described in 3.2.4.4. To prefractionate cross-linked peptides, either size exclusion chromatography (SEC) as described in see 3.2.5.3 or 3kDa CutOff spin column filtration was performed (see 3.2.5.4) to reduce sample complexity. After LC-MS/MS analysis, the free-available software pipeline xQuest/xProphet was used for data processing. To visualize identified crosslinks afterwards, the free available software tool Xinet was applied. Thereby, a csv-file extracted from the data file generated by xQuest/xProphet and containing protein ID as well as the crosslinking position within the linked peptides is uploaded. Additionally, a fasta sequence containing each of the proteins of interest has to be uploaded. To annotate known domains within the proteins, described domain localisations within a protein can be uploaded using an additional csv file.

### **3.2.5.3 Size exclusion chromatography of chemically cross-linked peptides**

Size exclusion chromatography (SEC) is a well described method [51] for the prefractionation of a peptide mix after chemical crosslinking. Thereby, molecules are separated due to their

molecular size. While small peptides are retained in the pores of the column material, larger molecules can pass and elute in the early fractions.

Cross-linked peptides were resuspended in 12 $\mu$ l SEC Buffer (30%ACN, 0.1%TFA) and separated using a Superdex Peptide PC 3.2/30" column at a flow-rate of 50 $\mu$ l/min on an Ettan FPLC. Fractions of 100 $\mu$ l were collected and UV traces were recorded at 214nm, 254nm and 280nm. Only early eluting fractions, containing large cross-linked peptides, were used for further LC-MS/MS analysis and volumes were reduced using a vacuum centrifuge. For LC-MS/MS analysis samples were filled up to 7.5 $\mu$ l with 0.5% TFA and whole sample concentration was injected.

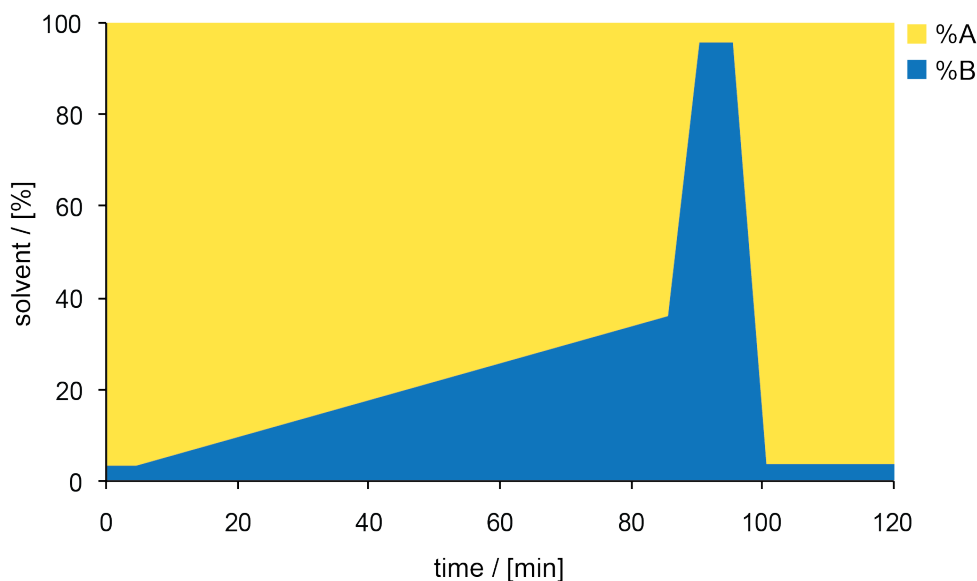
#### **3.2.5.4 Prefractionation using 3kDa CutOff spin column filtration**

To simplify the prefractionation of chemically cross-linked peptides, 3kDa Cut-Off spin columns were used to reduce the sample complexity. The cross-linked peptide mix was filled up to 100 $\mu$ l using ABC solution (50mM) and transferred to spin columns. Centrifugation was performed twice at 4°C: first at 10,000xg for 40min followed by 10min at 13,000xg. The volume of retained sample was reduced using a vacuum centrifuge. For LC-MS/MS analysis, 0.5% TFA was added to a volume of 7.5 $\mu$ l to the sample and whole sample was injected as one shot.

#### **3.2.6 LC-MS/MS analysis**

LC-MS/MS analysis was performed on an Ultimate3000 nano-HPLC coupled by a nanospray ion source to either an OrbitrapVelos for interaction study, to a QExactive Plus for the absolute quantification approach or to an OrbitrapFusion mass spectrometer for the analysis of cross-linked peptides.

Peptide mixtures were injected and loaded onto a nano trap column ( $\mu$ -Precolumn 300 $\mu$ m i.d. x 5mm, packed with Acclaim PepMap100 C18, 5 $\mu$ m, 100Å; Dionex) at a flow rate of 30 $\mu$ l/min in 98% buffer C (0.1% trifluoroacetic acid in HPLC-grade water) and 2% buffer B (80% acetonitrile (ACN), 0.08% formic acid (FA) in HPLC-grade water). After 3min, peptides were eluted and separated on an analytical column (75 $\mu$ m x 25cm, packed with Acclaim PepMap RSLC, 2 $\mu$ m, 100Å; Dionex) at a flow rate of 300 nl/min by a linear gradient from 2% up to 30% of buffer B in buffer A (2% acetonitrile, 0.1% formic acid) for either 82min for interaction study as well as for the absolute quantification, 147min for the single fraction after crosslinking using 3kDa Cut-Off spin columns or 35min for each fraction collected after crosslinking performing SEC. Afterwards, a short steep gradient from 30% to 95% buffer B was performed to elute remaining peptides. Eluting peptides were analysed in a LTQ OrbitrapVelos, QExactive Plus or OrbitrapFusion mass spectrometer. A LC gradient profile is depicted in Figure 15.



**Figure 15: LC gradient profile**

This figure shows the solvent composition of the nano pump over time. The concentration of solvent A is shown in yellow, while the concentration of buffer B is shown in blue. A LC gradient of 120min run time in total is depicted here. This gradient starts with 2% of solvent B for 5min. Then the concentration of B slowly increases up to 35% during a period of 80min followed by a strong increase of buffer B in 5min starting from 35% up to 95%. After 5min, concentration of %B rapidly decreases to 2% and stays at 2% solvent B for the last 20min.

### 3.2.6.1 Data-dependent analysis

For data-dependent analysis, full scan MS spectra measured on an OrbitrapVelos were acquired in a mass range from  $m/z$  300-1500 with a resolution of 30,000. The ten most intense precursor ions were selected, if they exceeded an intensity threshold of 200 and if they were at least doubly charged. Selected precursor ions were subsequently fragmented using collision-induced dissociation (CID) with the CID collision energy set to a value of 35. These parent ions were excluded from further isolation for 20sec. Resulting fragment ions were detected in the linear ion trap. Lock mass option was activated and set to a background signal with a  $m/z$  value of 445.12003 [73].

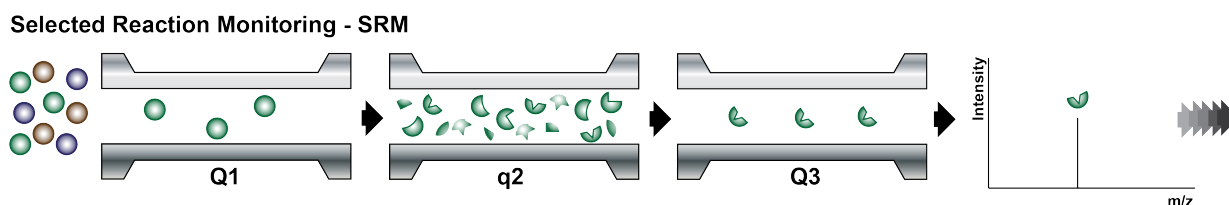
For the data-dependent analysis of cross-linked peptides measured on an Orbitrap Fusion, a pre-scan was performed in a mass range from  $m/z$  500-2000 with a resolution of 120,000. The ten most intense precursor ions were selected if they were at least triply charged and sequentially fragmented using collision-induced dissociation (CID) with the CID collision energy set to 35%. Resulting fragment ions were detected in the linear ion trap and selected precursor ions were excluded from further isolation for 20sec. Target value for MS scan was set to  $5 \times 10^5$

and for MS/MS fragmentation a target value of  $1 \times 10^4$  was chosen. Maximum injection times for MS scan as well as MS/MS fragmentation was 100msec. Lock mass option was activated and set to a background signal with  $m/z = 593.15761$ .

### 3.2.6.2 Targeted mass spectrometry

#### Selected Reaction Monitoring – SRM

Selected Reaction Monitoring (SRM) or also called Multiple Reaction Monitoring (MRM) is a reliable approach for the quantification of low abundant analytes within a complex sample mixture and is generally performed on a triple-quadrupole mass spectrometer. In this method, a predefined precursor ion is filtered. After fragmentation of the filtered precursors, a specific fragment ion is collected and analysed. As shown in Figure 16, quadrupole 1 (Q1) and quadrupole 3 (Q3) serve as mass filters, while the second quadrupole (q2) acts as collision cell where the selected precursor ions are fragmented. This set of predefined precursor and fragment ions is called a transition. Several of those transitions are monitored over time. The addition of retention time windows to the SRM assay achieves a high-selective MS-based proteomic approach. In general the 2-4 most intense fragment ions are chosen as transitions for each peptide [74].

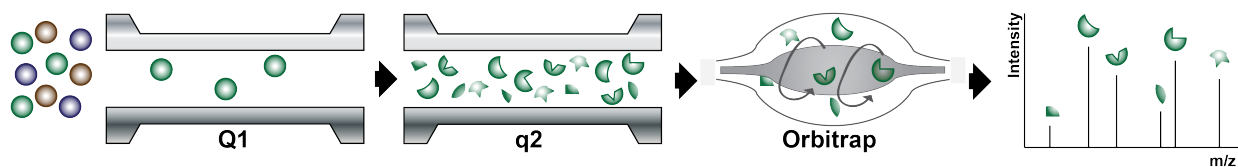


**Figure 16: Selected Reaction Monitoring – SRM**

Selected Reaction monitoring, short SRM, is a targeted mass spectrometry approach performed on a triple quadrupole instrument. Whereas quadrupole 1 (Q1) and quadrupole 3 (Q3) serve as mass filters, quadrupole 2 (q2) represents a collision cell. After selecting a predefined precursor ion in Q1 and fragmentation of it in q2, predefined fragment ions are filtered and analysed in Q3. The combination of predefined precursor and fragment ion is called a transition. In SRM approaches, transition can only be measured sequentially.

#### Parallel Reaction Monitoring – PRM

Another approach of targeted mass spectrometry is Parallel Reaction Monitoring (PRM) which is a common approach performed on quadrupole and Orbitrap containing hybrid mass spectrometers. Like in SRM, quadrupole 1 (Q1) acts as mass filter where predefined precursor ions are selected. After fragmentation in quadrupole 2 (q2), the collision cell, all corresponding fragment ions can be measured in an Orbitrap mass analyser in parallel (see Figure 17).

**Parallel Reaction Monitoring - PRM****Figure 17: Parallel Reaction Monitoring – PRM**

Parallel Reaction Monitoring (PRM) is a targeted mass spectrometry method performed on hybrid mass spectrometers containing a quadrupole as mass filter and an Orbitrap as mass analyser. After filtering a defined precursor ion in quadrupole 1 (Q1) and fragmenting it in the collision cell (quadrupole q2) all resulting fragment ions are analysed in parallel using an Orbitrap mass analyser.

### 3.2.7 Data processing

#### 3.2.7.1 Data-dependent analysis

Raw data were processed using Proteome Discoverer Daemon (Thermo Scientific; version 2.0) and Scaffold (Proteome Software Inc.; version 4.0) for qualitative analysis. All MS/MS spectra were analysed using Proteome Discoverer Daemon. Proteome Discoverer Daemon parameters were set to search the SwissProt\_2014\_04 database pre-selected for homo sapiens (2014\_04, 20,199 entries) using Trypsin/P as digestion enzyme. Searches were performed with a fragment ion mass tolerance of 0.5Da, permitted precursor charge states of two, three or four and an ion tolerance of 10ppm. Carbamidomethylation of cysteine was specified in Mascot as fixed modification, whereas deamidation of asparagine and glutamine as well as oxidation of methionine were specified as variable modifications. To evaluate MS/MS based peptide and protein identification, the software Scaffold was used. Peptide identifications were accepted with a probability > 95.0% by the Peptide Prophet algorithm [75] with a Scaffold delta-mass correction. Protein identifications were accepted with at least two identified peptides per protein and a probability > 99.0% by the Protein Prophet algorithm [76]. Proteins were only considered as interaction partners of the used bait, if they were not detected in the control experiments.

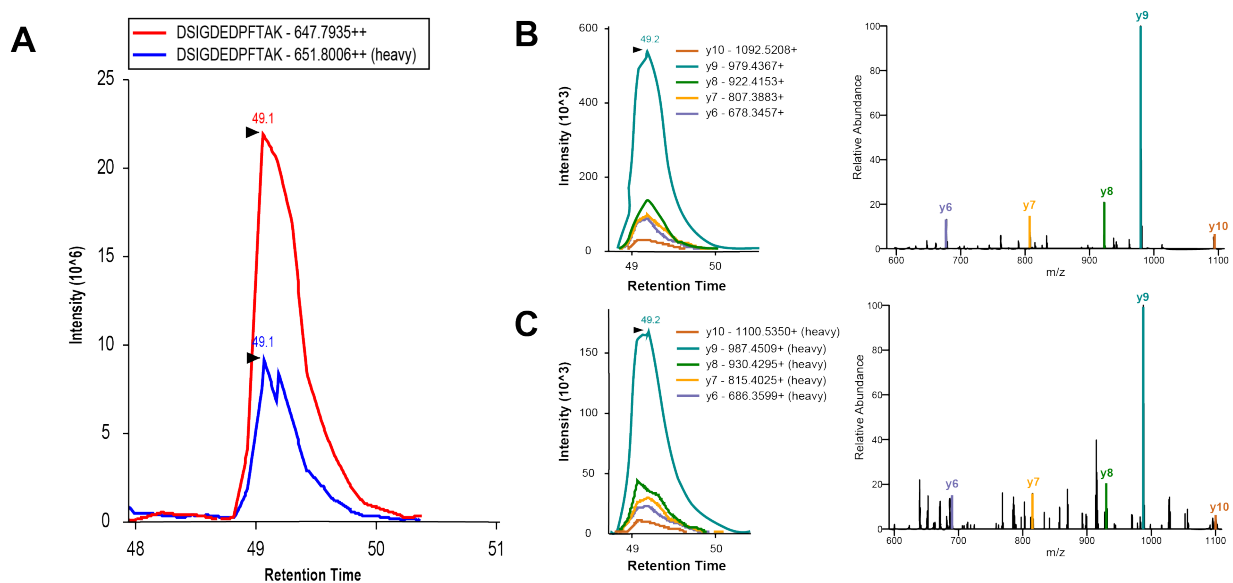
For label-free quantification (LFQ), MS raw data were analysed using MaxQuant software (1.5.3.30). Trypsin/P was set as cleaving enzyme. Carbamidomethylation of cysteine was selected as fixed modification, protein acetylation and oxidation of methionine were set as variable modifications. A maximum of two missed cleavages per peptide were allowed. False discovery rates (FDR) for proteins and peptides were both set to 1%. The first search tolerance for precursor ions was set to 20ppm following the main search option which was enabled with 4.5ppm. Fragment ion match tolerance was set to 20ppm. The human sub-set of the human proteome reference set provided by SwissProt (Release 2014\_04, 544,996 entries) was used for peptide and protein identification. For automated detection of contaminants like keratins, the MaxQuant contaminant database search was enabled. A minimum of two unique and razor peptides with a minimum length of seven amino acids had to be detected for protein quantification. The option "match between runs" was enabled with a match time window of 0.7min and an alignment time window of 20 min.

For the analysis of chemically cross-linked peptides, the previously described open-source xQuest/xProphet software pipeline was used. Thereby, cross-linked peptides were identified by xQuest search and validated afterwards by xProphet [51, 52, 77].

### 3.2.7.2 Targeted mass spectrometry

For data analysis of targeted mass spectrometry, obtained raw data were processed using the open source software tool Skyline. This application was developed to generate methods for targeted proteomics and to evaluate resulting data [78].

For data analysis using Skyline, it is essential to adjust parameter settings before raw files are uploaded. Peptide settings were set as follows: Trypsin [KR | P] was chosen as enzyme and no missed cleavage was allowed. Only peptides with a sequence length between 8 and 25 amino acids were considered. Introduced 'heavy' isotopic label of peptides was defined as well [(13C(6)15N(2)/ (C-term K) and (13C(6)15N(4)/ C-term R)]. For transition settings, allowed charges of precursor ions were set to 2 and only resulting y-ions were chosen for the quantification. After raw data processing, Skyline depicts identified peptides as shown in Figure 18. After evaluating automated integration of the software, some refinements were performed manually. For the quantification of each peptide, the ratio of labelled to non-labelled peptide was calculated by Skyline.



**Figure 18: Data analysis using Skyline**

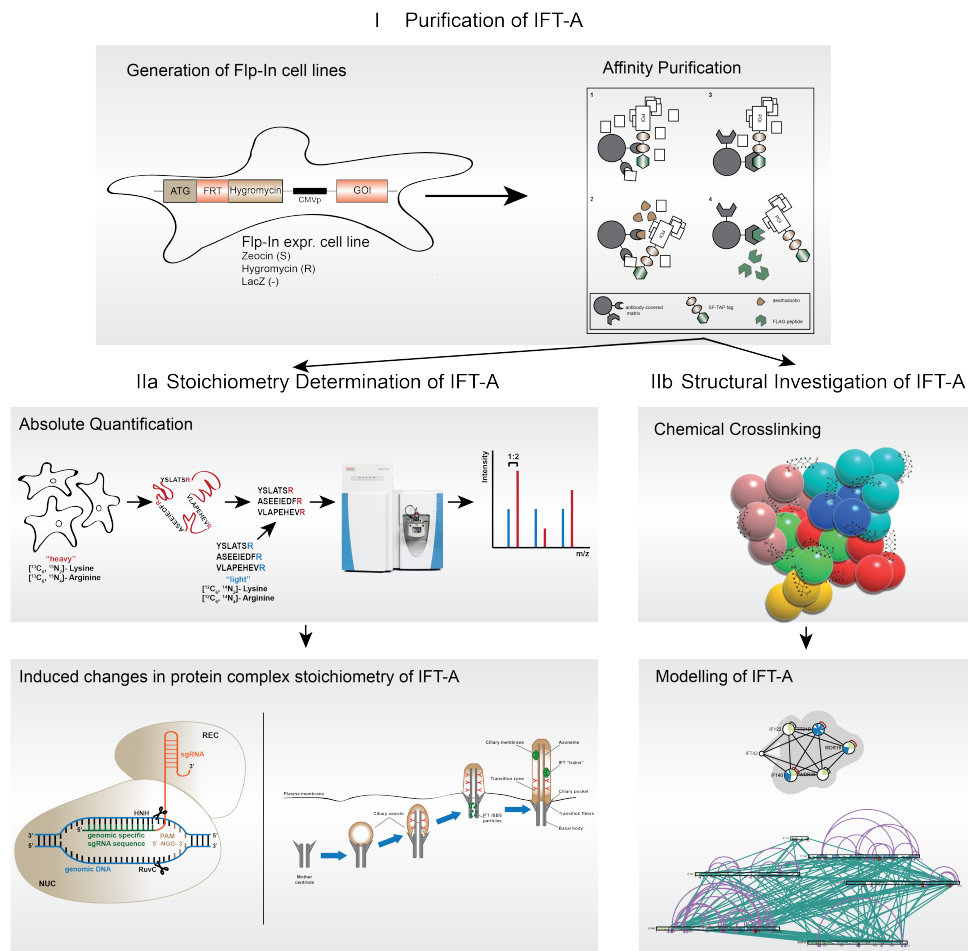
Skyline can be used to visualise raw data, generated by targeted mass spectrometry. Before rawfiles are analysed using Skyline, peptide as well as transition parameters must be set. Uploading rawfiles containing data of isotopically coded (biological peptides) and non-coded peptides (synthetic peptides) facilitates the illustration of corresponding peaks. On peptide level the ratio of labelled to unlabelled peptide is shown (A). To get more information about analysed transitions, peak areas as well as mass spectra of product ions of SILAC-labelled peptide (C) and of synthetic peptide (B) are depicted.



## 4 Results

### 4.1 Experimental workflow

Stoichiometric and structural investigations of the protein complex IFT-A are crucial for the understanding of the underlying mechanisms of IFT-A-related ciliopathies.



**Figure 19: Experimental workflow**

In order to obtain stoichiometric as well as structural information of IFT-A, the basis of this study was the generation of stable cell lines using the Flp-In system. Determination of the stoichiometry of proteins forming IFT-A was enabled by absolute quantification of purified IFT-A. Therefore, targeted mass spectrometry was combined with an internal standard mix. Furthermore, potential alterations and perturbations in stoichiometry of the IFT-A were analysed by two different strategies: To analyse alterations in IFT-A stoichiometry during assembly and disassembly of cilia, IFT-A was purified during different ciliary cell stages and used for absolute quantification. The second experiment to study stoichiometric changes of the protein complex was the introduction of mutations into IFT-A components by gene editing using the CRISPR/Cas9 system. For the structural investigation of purified IFT-A, chemical crosslinking was performed using the homobifunctional chemical crosslinker disuccinimidyl suberate (DSS). Resulting proximity information in combination with determined complex stoichiometry represent the basis for the modelling of IFT-A and to understand its role in ciliopathies.

For this study, purification of IFT-A in its naturally occurring composition is crucial. As hypothesized in 2013 by Gibson et al., overexpression, especially transient overexpression of one single protein, may cause an alteration in the corresponding complex assembly [58]. Therefore, stable cell lines with a single copy of the transgene integrated into the genome were generated using the Flp-In system. In this work, three Flp-In cell lines that stably express either (N)-SF-TAP-tagged IFT122, (N)-SF-TAP-tagged TULP3 or (N)-SF-TAP-tagged LCA5 were generated to enable the purification of IFT-A protein complex by the use of interacting as well as associated proteins as baits. To gain information about the complex composition of IFT-A, absolute quantification of all six described complex components (IFT122, IFT140, IFT43, TTC21B, WDR19 and WDR35) was performed. For the investigation of potential alterations within the stoichiometry of the IFT-A, the protein complex was analysed during different stages of ciliary assembly and disassembly. On top, CRISPR/Cas9 system was used to generate mutations based on Flp-In (N)-SF-TULP3 cells to study potential perturbations on complex composition introduced by mutations within IFT-A components.

The second part of this study was the structural investigation of the IFT-A. Therefore, purified IFT-A was chemically cross-linked using the homobifunctional crosslinker disuccinimidyl suberate (DSS). This crosslinker enables links between lysine residues as well as accessible N-termini and yields proximity information within the complex of interest [51]. The combination of both, stoichiometric and structural information, is the basis for the computational modelling of IFT-A and to improve basic knowledge about the function of IFT-A and its role in ciliopathies.

## **4.2 Generating Flp-In monoclonal cell lines**

For the purification of the protein complex IFT-A, cell lines stably expressing either (N)-SF-IFT122, (N)-SF-TULP3 or (N)-SF-LCA5 were produced using the Flp-In system. The three chosen baits are known to be either a part of the IFT-A (IFT122), associated with IFT-A without being an integral part (TULP3) or like LCA5, a labile and rather transiently bound interaction partner of the protein complex. Flp-In monoclonal lines were successfully generated for all three baits using HEK293 cells as described in 3.2.2.5.

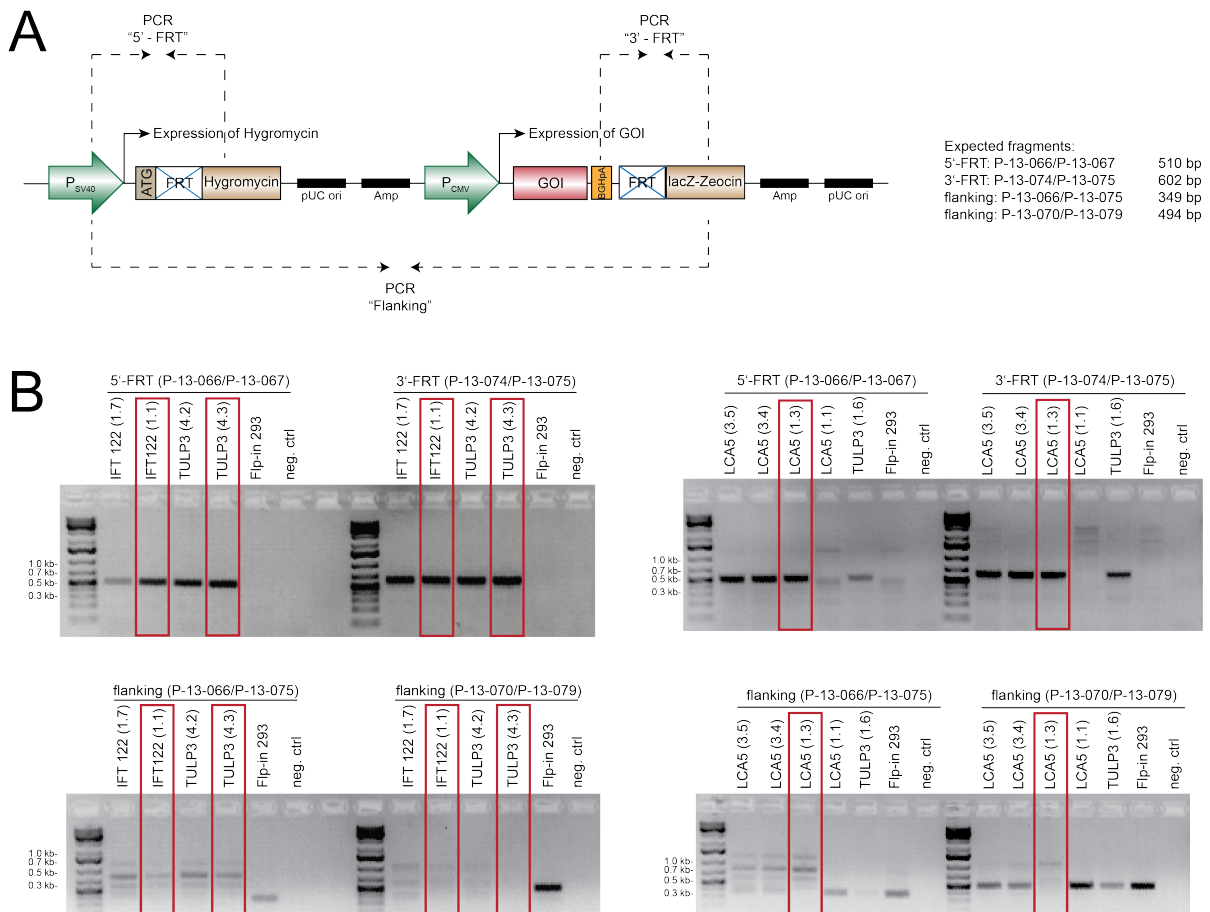
### **4.2.1 Behaviour of generated Flp-In monoclonal cell lines in cell culture**

In comparison to non-transfected HEK293 cells, generated Flp-In monoclonal lines were relatively slow growing, whereas the morphology of the cells stayed the same. Also the reaction to trypsination showed no difference to normal cells independent of the bait.

## 4.2.2 Validation of integration

To ensure a proper integration and expression of the gene of interest (GOI) into the Flp-In 293 host cell line (see 3.2.2.5), several validation experiments were performed: 1) PCR-amplification of specific regions characterizing the expression vector and its integration site and 2) analysis of the protein expression level through SDS-PAGE followed by western blotting.

### 4.2.2.1 PCR-based validation of the proper integration of the transgene



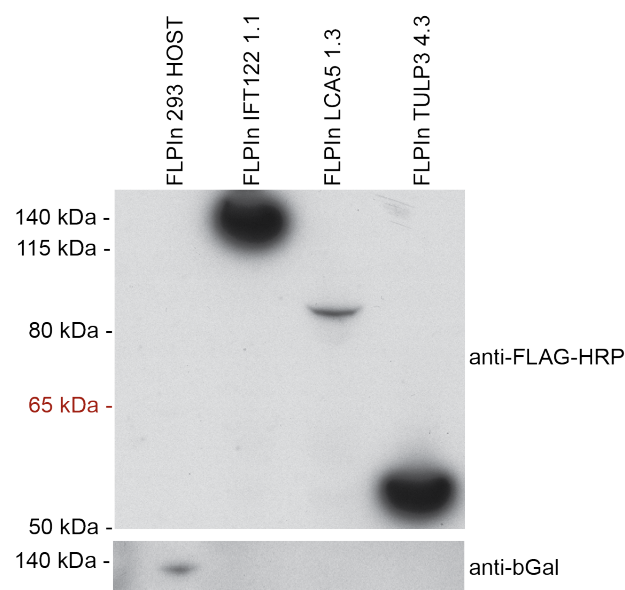
**Figure 20: Validation of generated Flp-In monoclonal lines via PCR**

PCR was performed to validate a proper integration of the GOI into the Flp-In host cell line. A: Different primer sets were used to amplify either the 5'-FRT and 3'-FRT-region (P-13-066/P-13-067; P-13-074/P-13-075) or the "flanking" region (P-13-066/P-13-075; P-13-070/P-13-079). After integration of the GOI construct into the Flp-In host line, the sequence length between the SV<sub>40</sub> promoter and lacZ-Zeocin gene is too big for a successful PCR. Therefore, no PCR fragment of 349/494bp is expected for monoclonal cell lines that underwent integration of the GOI construct. After the integration of the GOI construct into the Flp-In host cells, recombinant Flp-In lines contain an additional hygromycin resistance gene sequence as well as a BGHpolyA site which allows PCR at 5'-FRT and 3'-FRT with PCR products of 510bp and 602bp, respectively. B: Positive Flp-In monoclonal lines, highlighted in red, showed a proper integration of the GOI construct into the Flp-In host cell line. Whereas the 5'-FRT and 3'-FRT sites were amplified, the "flanking" regions were not. Those three clones were used for further experiments.

Polymerase chain reaction (PCR) was performed to validate a successful and proper integration of the GOI construct into the genome of the Flp-In host cell line. As shown in Figure 20A, few PCRs with different primer sets allowed the identification of recombinant clones. In the parental Flp-In host cell line, the “flanking” region between the SV40 promoter and the lacZ-Zeocin gene is short and can be amplified. Yielding PCR fragments with a length of either 349bp or 494bp (based on the chosen primer combination) are shown in Figure 20B (lanes of Flp-In 293). Based on the elongation of the flanking region by an integration of the GOI construct at the Flp-In site, this “flanking” region cannot be amplified anymore. In contrast, recombinant Flp-In clones possessing an additional hygromycin resistance gene sequence as well as a BGHPolyA site enable an amplification of the 5'- and the 3'- FRT sites with fragment sizes of 510bp or 602bp, respectively (see Figure 20A). As highlighted in Figure 20B, one clone per each GOI construct was chosen for further experiments (red boxes).

#### 4.2.2.2 Validation using western blotting in combination with immunostaining

For the validation of a successful expression of the Strep-FLAG-tagged bait protein, protein extracts (20µg protein) of the generated cell lines were separated performing SDS-PAGE followed by western blotting with further immunostaining (see 3.2.3.3 and 3.2.3.4).



**Figure 21: Validation of bait protein expression of Flp-In monoclonal cell lines**

Western blot analysis was performed to validate a successful integration of the gene of interest (GOI) into the genome of a mammalian cell line, using the Flp-In system of Invitrogen. For SDS-PAGE, 20µg extracted proteins of generated Flp-In monoclonal lines containing SF-TAP-tagged IFT122, TULP3 or LCA5 as bait as well as of Flp-In 293 host cells were used. After transferring separated proteins on a PVDF membrane via western blot, the FLAG tag of integrated GOI's was detected using anti-FLAG-HRP antibodies. After integration of the GOI into Flp-In 293 host cell line, protein expression of  $\beta$ -galactosidase was not detectable.

---

An anti-FLAG-HRP antibody was used for the detection of the expression of the chosen bait proteins. As shown in Figure 21, all tested monoclonal cell lines expressed a bait protein of the appropriate size ((N)-SF-IFT122: 147kDa; (N)-SF-LCA5: 55kDa; (N)-SF-TULP3: 86kDa). Successful integration of the GOI construct into the Flp-In host cell line eliminates expression of the  $\beta$ -galactosidase gene. Therefore, a second immunostaining was performed using an anti- $\beta$ -galactosidase antibody. Whereas a  $\beta$ -galactosidase expression was detectable in Flp-In host cell lysates, no  $\beta$ -galactosidase expression was detected in the recombinant Flp-In lines.

### **4.3 Establishing targeted mass spectrometry for absolute quantification of IFT-A**

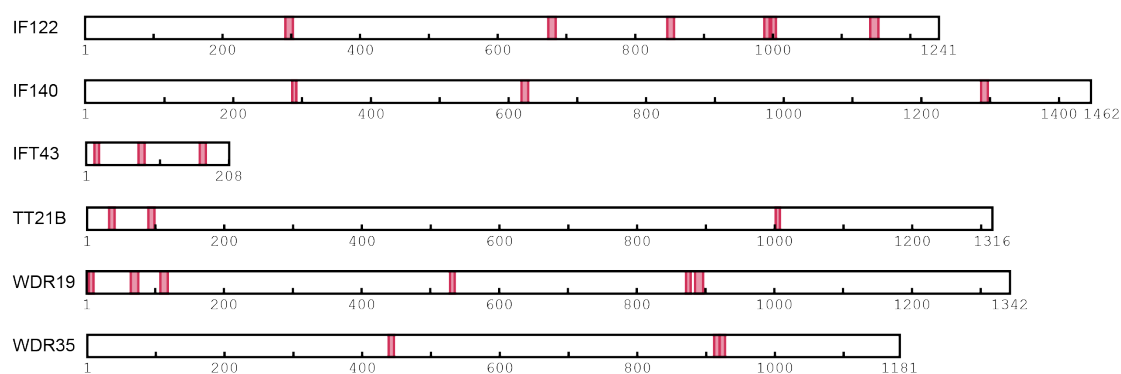
#### **4.3.1 Selection of representative peptides of IFT-A**

Absolute quantification of proteins combines the application of representative and proteotypic standard peptides of defined quantity with targeted mass spectrometry. As outlined in 3.2.5.1, the selection of representative peptides for this elaborative technology is a crucial step. To eliminate ambiguous calls, only tryptic peptides which are exclusively present in the target protein were identified by Basic Local Alignment SEARCH Tools (BLAST). Length of the amino acid sequence as well as amino acid composition of chosen peptides were also filtered to avoid amino acids which are prone to post-translational modifications. Figure 22 lists at least three representative standard peptides for each complex component which were chosen for absolute quantification of purified IFT-A. In Figure 22A, the localisation of the chosen peptides within the corresponding protein is given by the position of the amino acid residue (highlighted in blue), whereas the distribution of the representative peptides on the protein sequence is depicted by red bars in Figure 22B. To determine the absolute amount of each of the peptides performing the economic “Equimolarity through Equalizer Peptide” (EtEP), all of the peptides were synthesized with an artificial tryptic peptide at the N-terminus.

A

IFT122	N-292GEYILLGGSDK302-xxx-674YLELISSIEER684-xxx-847WDEAFALGEK856-xxx -988DTPSGISK995-xxx-998ILFTLAK1004-xxx-1142DSIGDEDPFTAK1153-C
IFT140	N-302FWDIER307-xxx-635SHLFVDEGLK644-xxx-1303AHGALTEAYK1312-C
IFT43	N-13YSLATSR19-xxx-77ASEEIEDFR85-xxx-166VLAPEHEVR174-C
TTC21B	N-33YGSDPVFR40-xxx-90EAILESDAR98-xxx-1002LEDVPR1007-C
WDR19	N-4IFSLLLEK10-xxx-65DGDVLAVIAEK75-xxx-108VGSFLAVGTVK118-xxx -529LVFIDEK535-xxx-872AASVYIR878-xxx-885VGDLLPHVSSPK896-C
WDR35	N-439THVIAASK446-xxx-912EIGSLLAR919-xxx-920YASHLLEK927-C

B



**Figure 22: Proteotypic peptides of IFT-A**

To determine the complex stoichiometry using absolute quantification, the combination of a targeted mass spectrometry approach with an internal standard of proteotypic peptides is essential. A: At least three proteotypic peptides were selected as representatives for each IFT-A complex component. The peptide sequence of the chosen representative peptides is highlighted in red while the position of the peptide within the protein sequence is depicted in blue. B: The distribution of the representative peptides (red bars) on the protein sequence is given.

#### 4.3.2 Generation of an equimolar standard mix performing EtEP

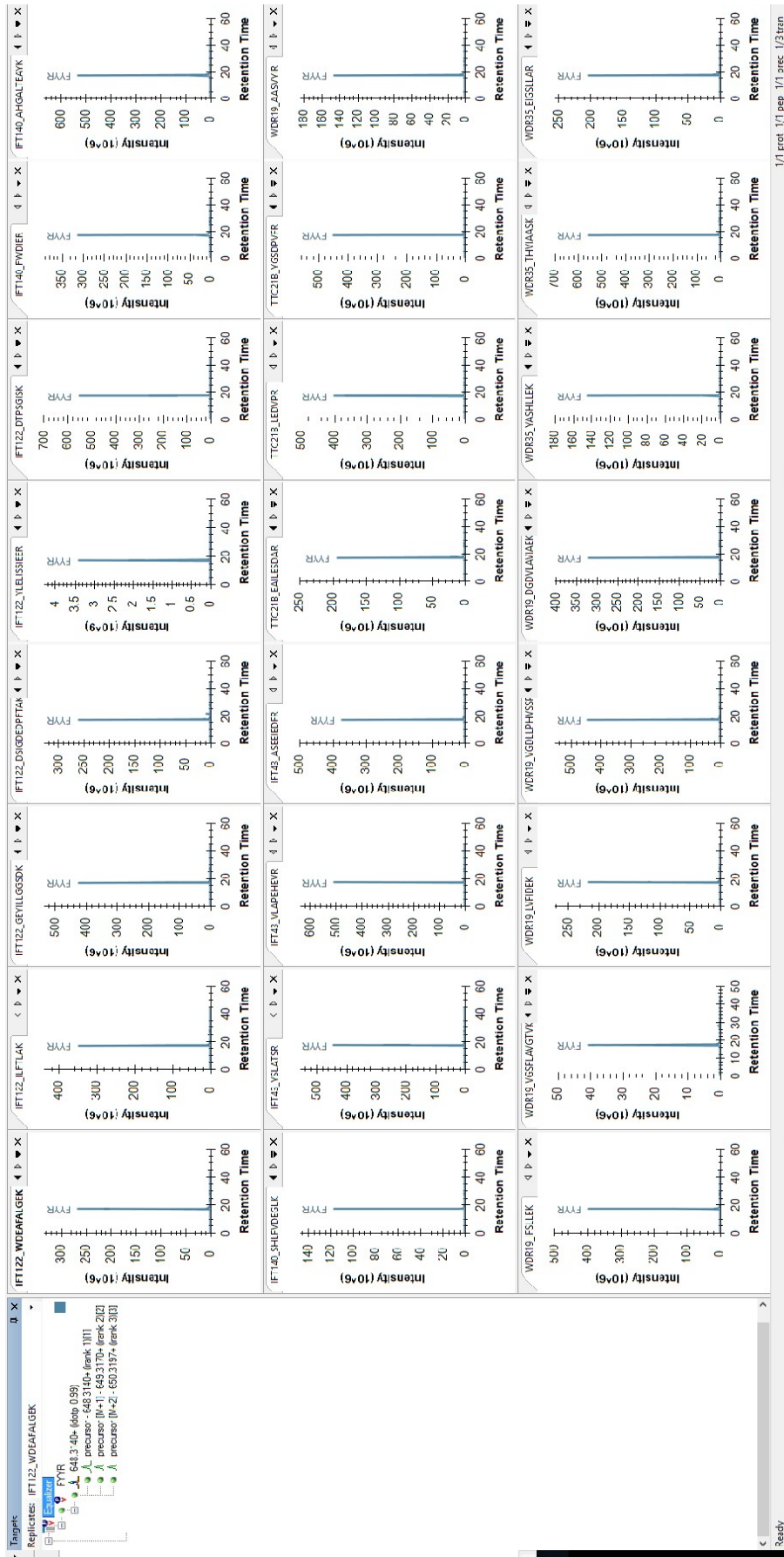
A standard mix with a known amount of each standard peptide is the basis for the absolute quantification of a protein complex of interest. However, the determination of the absolute amount of each peptide by amino acid analysis is very expensive. Therefore, quantification of the absolute amount was done by a cost-efficient method called “Equimolarity through Equalizer Peptide” (EtEP), previously described by Holzmann et al. in 2009 [59] (see 3.2.5.19). In EtEP, the absolute amount of only one peptide, the equalizer peptide (EP), is determined by amino acid analysis. Furthermore, this EP needed to be synthesized in an isotopically labelled form (FYR\*). All representative peptides were synthesized with an attached equalizer peptide

sequence (FYYR-) at the N-terminus. This artificial peptide introduces an additional tryptic cleavage site which is essential for this EtEP method.

#### **4.3.2.1 Validating tryptic proteolysis**

Proteolysis of each EP-fused representative peptide using trypsin results in an equimolar quantity of released EP and synthetic standard peptide. An efficient tryptic digestion was validated performing LC-MS/MS analysis on a QExactive Plus instrument, as shown in Figure 23. Therefore, generated rawfiles were analysed by the software Skyline to validate the tryptic digestion by evaluating the release of the artificial equalizer peptide FYYR. In contrast, an unsuccessful tryptic cleavage of the synthesized peptides would result in no release of the artificial tryptic peptide FYYR.





**Figure 23: Release of Equalizer Peptide after efficient tryptic proteolysis**

As a first control and to confirm an effective proteolysis which is crucial for the generation of an equimolar standard mix performing the “Equimolarity through Equalizer Peptide” (EtEP) method, the release of the equalizer peptide (FYR) was analysed using Skyline. As depicted, proteolysis of all tested representative synthetic peptides was successful. Y-axis was scaled differently, according to the peak intensities of released EP.

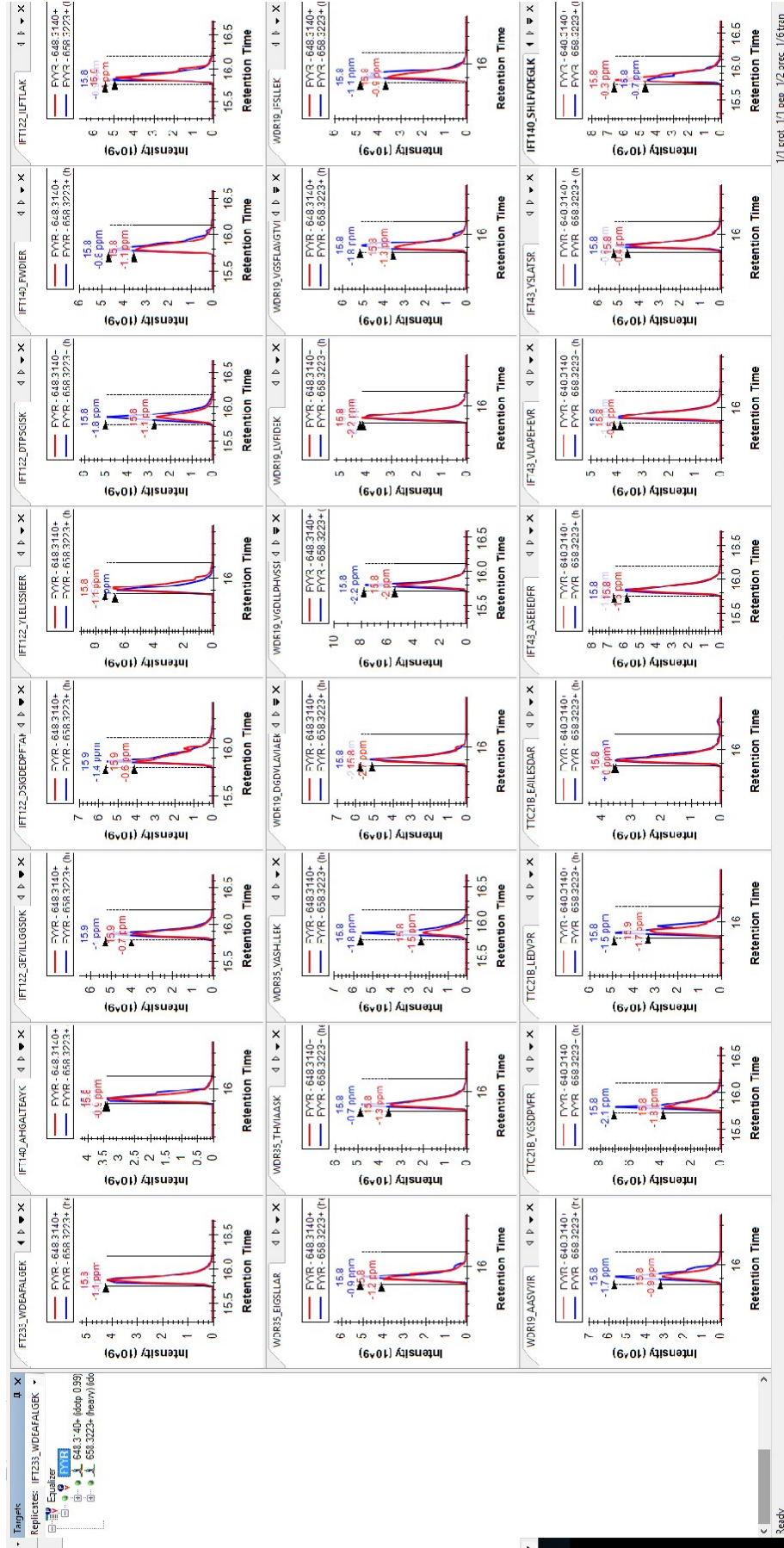


---

#### 4.3.2.2 Quantification of standard peptides using isotopically labelled EP

After a successful proteolysis, the amount of each standard peptide was determined by direct in-sample comparison of the released equalizer peptide (EP) with a known amount (8.7 $\mu$ g) of the isotopically labelled EP spiked into the trypsin-treated peptide samples. After targeted mass spectrometry was performed, rawfiles were analyzed by Skyline again to determine the ratio of unlabelled EP (FY $YR$ ), released by proteolysis, to the spiked in and isotopically labelled EP (FY $YR^*$ ). Figure 24 shows the elution profiles of the spiked-in labelled EP (depicted in blue) and of the released EP of each standard peptide (highlighted in red). The analysis showed that the amount of released equalizer peptide varied considerably. Differences in synthesis yield and solubility (depending on the amino acid composition of the synthetic peptide) may be responsible for this variation. To eliminate error-prone quantifications, based on differences in abundance of the EP's and to consider the dynamic range of the mass spectrometer, the amount of spiked labelled equalizer peptide was adjusted to the individual amounts of released EP. Results are shown in Figure 25.





**Figure 25: Quantification of synthetic peptides with adjusted amounts of spiked EP**

Based on the dynamic range of used mass spectrometers, large differences between the released EP and the spiked-in labelled EP may cause error-prone quantifications. Therefore, the amount of isotopically labelled EP, spiked into each proteolysis-based representative peptide mixture, was adjusted to the detected abundance of released equalizer peptide shown in Figure 24.

The ratio of unlabelled versus isotopically labelled equalizer peptide, as calculated by Skyline, was used for the determination of the concentration of each representative peptide. Therefore, the following formula was used:

$$\frac{A_{\text{Peptide}}}{A_{\text{EP}}} = \frac{c_{\text{Peptide}}}{c_{\text{EP}}} \cdot d_{\text{EP}}$$

$$c_{\text{Peptide}} = \frac{c_{\text{EP}}}{d_{\text{EP}}} \cdot \frac{A_{\text{Peptide}}}{A_{\text{EP}}}$$

with

A : peak area

c : concentration

$c_{\text{EP}} = 661,396 \text{ pmol}/\mu\text{l}$

$d_{\text{EP}} = \text{dilutionfactor}$

The known concentration of the isotopically labelled EP of 661.4 pmol/ $\mu\text{L}$  together with the dilution factor (see Table 10) was used to calculate the concentration of each standard peptide. Table 10 also lists the volume of each peptide sample to generate the equimolar standard mix containing 550pmol of each representative peptide.

**Table 10: Calculation for an equimolar standard mixture**

Protein	Peptide	Mean $A_{\text{Peptide}}/A_{\text{EP}}$	%CV	Dilution factor	$c_{\text{Peptide}} / [\text{pmol}/\mu\text{L}]$	$V_{\text{Peptide}} / [\mu\text{l}]$
IFT122	WDEAFALGEK	0.9130	12.2	2.5	241.5	2.3
	ILFTLAK	0.9415	2.8	2.5	249.1	2.2
	GEYILLGGSDK	0.7826	3.3	2	258.8	2.1
	DSIGDEDPFYAK	0.9049	6.0	1	598.5	0.9
	YLELISSIEER	1.7207	19.2	12.5	91.0	6.0
	DTPSGISK	0.5255	14.9	1.6	217.2	2.5
IFT140	FWDIER	0.7564	8.0	2	250.2	2.2
	AHGALTEAYK	0.8787	22.3	2	290.6	1.9
	SHLFVDEGLK	1.1666	7.4	10	77.2	7.1
IFT43	YSLATSR	0.8799	11.0	2.5	232.8	2.4
	VLAPEHEVR	0.8532	4.5	2	282.2	1.9
	ASEEIEDFR	0.9707	15.0	3	214.0	3
TTC21B	EAILESDAR	0.8781	11.8	3	193.6	2.8
	LEDVPR	0.7365	8.2	1.5	324.7	1.7
	YGSDPVFR	0.7916	4.9	2	261.8	2.1
WDR19	AASVYIR	0.6822	11.0	5	90.2	6.1
	IFSLLEK	0.7371	7.9	2	243.7	2.3
	VGSFLAVGTVK	0.7949	15.4	20	26.3	20.9
	LVFIDEK	0.8887	16.6	4	147.0	3.7
	VGDLLPHVSSPK	0.7794	5.1	2	257.8	2.1
	DGDVLAVIAEK	0.8810	13.7	2	291.3	1.9
WDR35	YASHLLEK	0.4942	7.7	4	81.7	6.7
	THVIAASK	0.8046	2.4	1.5	354.8	1.6
	EIGSLLAR	0.8871	4.3	4	146.7	3.7

%CV: percent coefficient of variation

### 4.3.3 Adjusted standard mix for absolute quantification

For a most-reliable quantification of peptides, the dynamic range of the used mass spectrometer must be considered. This parameter gives the range in which accurate measurements of masses can be made [79], depending on the linearity of ion signal to analyte concentration in this range. To reliably quantify labelled and its corresponding non-labelled peptide, similar amounts of spiked isotopically labelled standard peptide and biological peptide within one sample is important. Therefore, previous data of absolute quantification of IFT-A using an equimolar standard mix (for more details see 4.4.2.2) was used to determine the amount of each peptide within purified IFT-A. Based on this knowledge, a mixture consisting of adjusted amounts of each synthetic peptide was created. Table 11 shows the composition of this adapted standard mix. In most cases, representative peptides for one protein are present in the same amount and therefore adjusted the same way. In contrast, for some proteins (IFT122 and WDR19), representative peptides showed different abundances. In this case, each representative peptide was adjusted separately.

**Table 11: Calculation for a standard mixture with adjusted amount of synthetic peptides**

Protein	Peptide	Mean $A_{\text{Pep}}/A_{\text{EP}}$	%CV	Dilution factor	$C_{\text{Peptide}} / [\text{pmol}/\mu\text{l}]$	$V_{\text{Peptide}} / [\mu\text{l}]$	$n_{\text{Peptide}} / [\text{pmol}]$
IFT122	WDEAFALGEK	0.9130	12.2	2.5	241.5	17.6	4250
	ILFTLAK	0.9415	2.8	2.5	249.1	9.0	2250
	GEYILLGGSDK	0.7826	3.3	2	258.8	3.2	825
	DSIGDEDPFTAK	0.9049	6.0	1	598.5	1.4	825
	YLELISSIEER	1.7207	19.2	12.5	91.0	46.7	4250
IFT140	DTPSGISK	0.5255	14.9	1.6	217.2	2.3	500
	FWDIER	0.7564	8.0	2	250.2	7.0	1750
	AHGALTEAYK	0.8787	22.3	2	290.6	6.0	1750
IFT43	SHLFVDEGLK	1.1666	7.4	10	77.2	22.7	1750
	YSLATSR	0.8799	11.0	2.5	232.8	1.1	250
	VLAPEHEVR	0.8532	4.5	2	282.2	0.9	250
TTC21B	ASEEIEDFR	0.9707	15.0	3	214.0	1.2	250
	EAILEDAR	0.8781	11.8	3	193.6	2.2	425
	LEDVPR	0.7365	8.2	1.5	324.7	1.3	425
WDR19	YGSDPVFR	0.7916	4.9	2	261.8	1.6	425
	AASVYIR	0.6822	11.0	5	90.2	5.5	500
	IFSLEK	0.7371	7.9	2	243.7	19.5	4750
	VGSFLAVGTVK	0.7949	15.4	20	26.3	95.1	2500
	LVFIDEK	0.8887	16.6	4	147.0	7.3	1075
	VGDLLPHVSSPK	0.7794	5.1	2	257.8	8.7	2250
WDR35	DGDVLAVIAEK	0.8810	13.7	2	291.3	3.7	1075
	YASHLLEK	0.4942	7.7	4	81.7	10.1	825
	THVIAASK	0.8046	2.4	1.5	354.8	2.3	825
	EIGSLLAR	0.8871	4.3	4	146.7	5.6	825

%CV: percent coefficient of variation

Although it is a good idea to eliminate error-prone concentration differences, the coefficient of variance obtained in experiments with the adjusted standard mix were much higher than for absolute quantification using an equimolar standard mix (data not shown). This worse performance is most likely due to the additional steps in sample handling. Because of that, the

---

Flp-In (N)-SF-TULP3 monoclonal line in combination with the equimolar standard mix was used for further investigations.

#### **4.3.4 Parameter settings for targeted mass spectrometry**

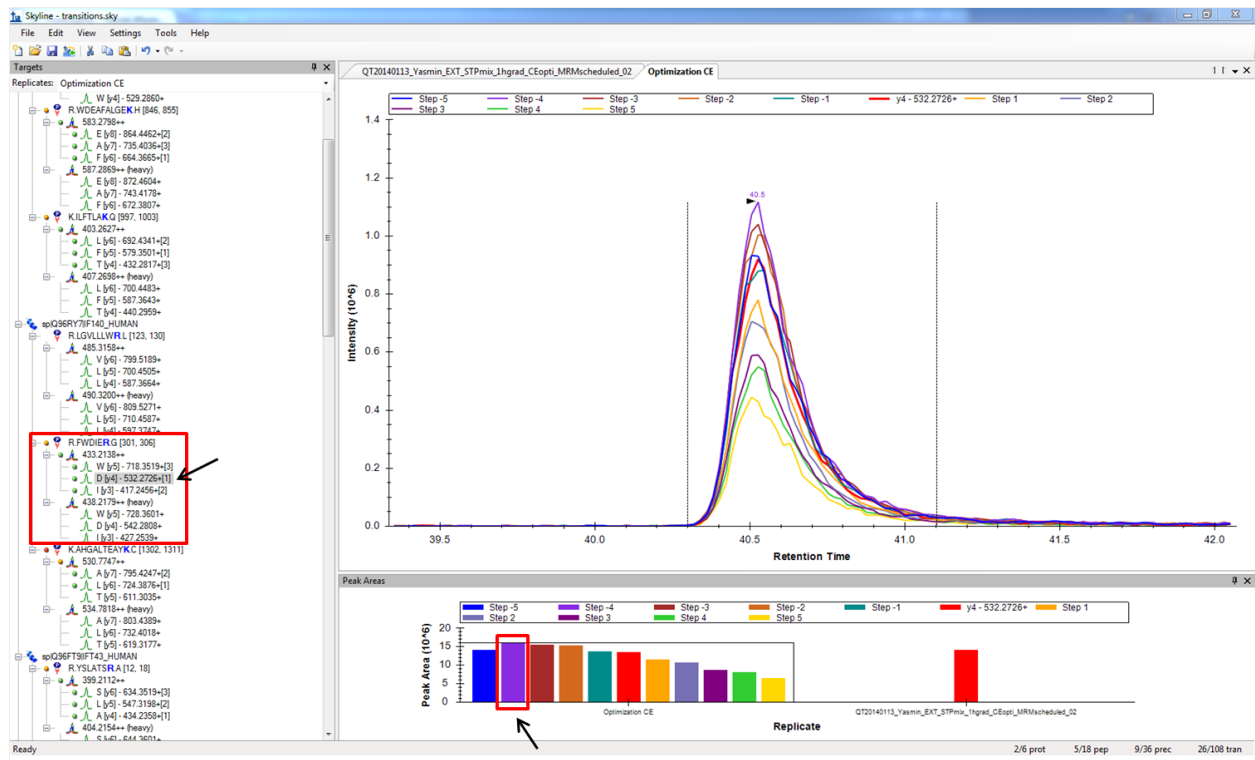
To set up a method for targeted mass spectrometry, the MS profiles of the representative peptides, present in the generated standard mix, were characterized by a data-dependent shotgun run. Therefore, the equimolar standard mix was diluted 1:10 in 0.5% TFA (injected amount of 55fmol of each standard peptide) and used for MS analysis. The obtained rawfile was analyzed using Skyline to evaluate transitions as well as to determine the retention time of each eluting peptide.

##### **4.3.4.1 SRM assay performed on a triple quadrupole**

At the beginning of this study, no mass spectrometer with an option for targeted mass spectrometry was available in our lab. Therefore, the analysis based on the Selected Reaction Monitoring (SRM) protocol was done on a QTrap5500 instrument in the group of Karl Mechtler (Institute of Molecular Pathology, Vienna). For SRM, performed on a triple quadrupole mass spectrometer, the choice of significant transitions (a precursor ion with one corresponding product ion) is essential. The chosen transitions are representatives for the quantification of the peptides. To achieve optimal sensitivity, the collision energy (CE) was optimised for each transition by selecting the most efficient CE value for every ion pair. CE optimisation was performed by measuring each transition with different collision energies. To locate the CE range of each transition, different CE values (10, 20, 30) were tested first. Afterwards, best CE value of the previous experiment was chosen to refine the CE by testing the chosen CE  $\pm$  5 with a step size of  $\pm$  1. Results were visualized in Skyline to identify ideal conditions. As an example, shown in Figure 26, peak areas were analyzed to identify the CE optimum for the ion y4 of the peptide FWDIER. The optimal CE for this transition was identified by the chosen value – 4.

Once optimal CE values were identified for each representative peptide (14 peptides in total as shown in Table 12), a transition list was generated to set parameters for the SRM approach. This list is shown in Table 12 and includes chosen isotopically coded and non-coded peptides ('light': L, 'heavy': H) with corresponding precursor masses as well as chosen product ions. Additionally, optimised CE values are given. To increase speed and selectivity of this approach, the retention time was added. For the analysis a retention time window of  $\pm$  1.5min was selected.





**Figure 26: Determination of the CE optimum for transitions using Skyline**

Ideal fragmentation conditions are essential for a robust and reliable quantification. Therefore, the collision energy value was optimized for each peptide and each transition separately. To identify the optimum of CE, peak areas of each transition, generated using different CE values, were analysed and compared. Therefore, a set of different CE values was tested (10, 20, 30). To refine this optimization the range of tested CE values in the next experiment was defined by the best CE value of the previous experiment  $\pm 5$  with a step size of  $\pm 1$ . Skyline depicts the resulting peak areas to identify the best CE value. As highlighted in red, for the product ion y4 of the double charged precursor of the peptide FWDIER, ten different CE values were tested. Highest peak area was obtained at -4 below the preset CE and used for this very transition.

Table 12: Transition list for SRM with optimised CE values

Protein	Peptide L/H	Precursor [m/z] L/H	Product [m/z] L/H	Transition	RT	CE
IFT122	WDEAFALGEK/K	583.279846/587.286945	864.446159/872.460358	+2y8	43.97	26.9
			735.403566/743.417765	+2y7	43.97	28.9
			664.366452/672.380651	+2y6	43.97	27.9
	ILFTLAK/K	403.262739/407.269839	333.176861/341.19106	+2y3	43.97	34.9
			692.434138/700.448337	+2y6	41.30	19.4
			579.350074/587.364273	+2y5	41.30	18.4
432.28166/440.295859			+2y4	41.30	24.4	
IFT140	FWDIER/R	433.213777/438.217912	331.233982/339.248181	+2y3	41.30	27.4
			718.351865/728.360134	+2y5	40.53	20.5
			532.272552/542.280821	+2y4	40.53	20.5
	AHGALTEAYK/K	530.774731/534.78183	417.245609/427.253878	+2y3	40.53	28.5
			852.446159/860.460358	+2y8	16.47	26.0
			795.424696/803.438895	+2y7	16.47	27.0
			724.387582/732.401781	+2y6	16.47	31.0
			381.213246/389.227445	+2y3	16.47	33.0
			634.351865/644.360134	+2y6	20.45	21.2
			547.319837/557.328106	+2y5	20.45	21.2
IFT43	YSLATSR/R	399.211235/404.215369	434.235772/444.244041	+2y4	20.45	20.2
			363.198659/373.206928	+2y3	20.45	20.2
			808.383559/818.391828	+2y6	30.52	23.6
			679.340966/689.349235	+2y5	30.52	25.6
	ASEEIEDFR/R	548.251285/553.25542	566.256902/576.265171	+2y4	30.52	24.6
			437.214309/447.222578	+2y3	30.52	27.6
			837.421342/847.429611	+2y7	16.38	25.8
			766.384228/776.392497	+2y6	16.38	26.8
	VLAPEHEVR/R	525.290548/530.294682	669.331464/679.339733	+2y5	16.38	32.8
			540.288871/550.29714	+2y4	16.38	32.8
			777.388979/787.397248	+2y7	28.44	21.8
			720.367515/730.375784	+2y6	28.44	21.8
TTC21B	YGSDPVFR/R	470.729792/475.733926	633.335487/643.343756	+2y5	28.44	21.8
			518.308544/528.316813	+2y4	28.44	28.8
			803.425758/813.434027	+2y7	23.56	24.9
	EAILESDAR/R	502.256371/507.260505	690.341694/700.349963	+2y6	23.56	23.9
			577.25763/587.265899	+2y5	23.56	23.9
			448.215037/458.223306	+2y4	23.56	21.9
			615.309666/625.317935	+2y5	18.61	19.0
	LEDVPR/R	364.700503/369.704637	486.267073/496.275342	+2y4	18.61	18.0
			371.24013/381.248399	+2y3	18.61	25.0
			457.22929/457.22929	+2b4	18.61	17.0
WDR19	IFSLLLEK/K	425.257654/429.264753	736.423967/744.438166	+2y6	42.39	20.2
			589.355553/597.369752	+2y5	42.39	20.2
			502.323525/510.337724	+2y4	42.39	26.2
			389.239461/397.25366	+2y3	42.39	28.2
	AASVYIR/R	390.224145/395.22828	708.403901/718.41217	+2y6	21.65	19.9
			637.366787/647.375056	+2y5	21.65	19.9
			550.334758/560.343027	+2y4	21.65	20.9
			451.266344/461.274613	+2y3	21.65	20.9
WDR35	THVIAASK/K	413.742702/417.749802	588.371538/596.385737	+2y6	8.93	22.8
			489.303124/497.317323	+2y5	8.93	21.8
			376.21906/384.233259	+2y4	8.93	23.8
			305.181946/313.196145	+2y3	8.93	23.8
	YASHLLEK/K	480.761092/484.768191	797.451579/805.465778	+2y7	18.88	24.2
			726.414465/734.428664	+2y6	18.88	24.2
			502.323525/510.337724	+2y4	18.88	29.2
			389.239461/397.25366	+2y3	18.88	31.2



#### 4.3.4.2 PRM assay performed on an Orbitrap instrument

Later in this study, an Orbitrap hybrid mass spectrometer (QExactive Plus) was available in our lab to perform targeted mass spectrometry. In comparison to the SRM approach with predefined transitions and transition parameters, parameter setting for a Parallel Reaction Monitoring (PRM) approach which can be performed on an Orbitrap hybrid mass spectrometer, is much less laborious. As depicted in Table 13, only precursor masses for isotopically labelled (highlighted in red) and non-labelled parent ions are predefined. Instead of quantifying only predefined product ions like in SRM, all resulting fragment ions are analyzed and detected in parallel in the Orbitrap. All chosen precursors are 2-fold positively charged and CE values which are essential for a proper fragmentation were universally set to 26. For optimising speed and selectivity, the method was set up in a scheduled manner including a retention time range ( $RT \pm 5\text{min}$ ). For the qualitative identification of the proteins of interest, not only product ions were analysed, but also a targeted  $MS^1$  scan of all included precursor masses was generated.

**Table 13: Parameter settings for PRM**

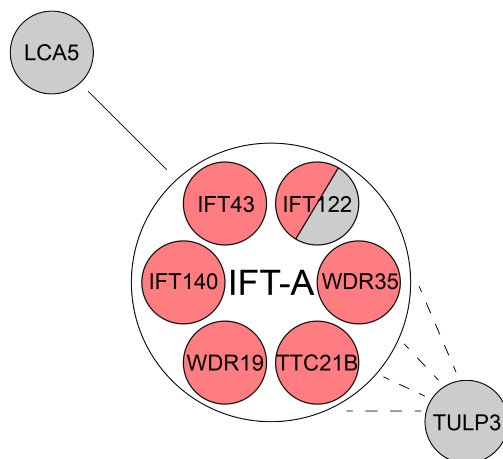
IFT-A component	Peptide L/H	Precursor [m/z] L/H	[z]	Polarity	RT <sub>min</sub>	RT <sub>max</sub>	(N)CE
IFT122	WDEAFALGEK/K	583.27985 / 587.28695	2	+	50.4	60.4	26
	ILFTLAK/K	403.26274 / 407.26984	2	+	47.3	57.3	26
	GEYILLGGSDK/K	576.30078 / 580.30788	2	+	44.2	54.2	26
	DSIGDEDPFTAK/K	647.79351 / 651.80061	2	+	46.4	56.4	26
	YLELISSIEER/R	676.3588 / 681.363	2	+	61.2	71.2	26
	DTPSGISK/K	402.7085 / 406.7156	2	+	13.2	23.2	26
IFT140	FWDIER/R	433.21378 / 438.21791	2	+	44.4	54.4	26
	AHGALTEAYK/K	530.77473 / 534.78183	2	+	16.6	26.6	26
	SHLFVDEGLK/K	572.80349 / 576.81059	2	+	34.6	44.6	26
IFT43	YSLATSR/R	399.21124 / 404.21537	2	+	20.3	30.3	26
	VLAPEHEVR/R	525.29055 / 530.29468	2	+	17.3	27.3	26
	ASEEIEDFR/R	548.25129 / 553.25542	2	+	32.4	42.4	26
TTC21B	EAILESDAR/R	502.25637 / 507.26051	2	+	24.7	34.7	26
	LEDVPR/R	364.7005 / 369.70464	2	+	18.9	28.9	26
	YGSDPVFR/R	470.72979 / 475.73393	2	+	30.1	40.1	26
WDR19	AASVYIR/R	390.22415 / 395.22828	2	+	21.6	31.6	26
	IFSLLEK/K	425.25765 / 429.26475	2	+	48.5	58.5	26
	VGSFLAVGTVK/K	539.31877 / 543.32587	2	+	45.5	55.5	26
	LVFIDEK/K	432.24729 / 436.25439	2	+	40.9	50.9	26
	VGDLLPHVSSPK/K	624.851 / 628.8581	2	+	36.3	46.3	26
	DGDVLAVIAEK/K	565.3086 / 569.3157	2	+	56.2	66.2	26
WDR35	YASHLLEK/K	480.76109 / 484.76819	2	+	19.5	29.5	26
	THVIAASK/K	413.7427 / 417.7498	2	+	9.7	19.7	26
	EIGSLLAR/R	429.75581 / 434.75994	2	+	38.3	48.3	26

#### 4.4 Absolute quantification of purified IFT-A

Absolute quantification of a protein complex of interest consists of different steps including an entire purification of the protein complex, the addition of an internal standard mix and targeted mass spectrometry of each representative peptide (24 peptides in total as shown in Table 13) and its corresponding isotopically labelled peptide. Finally, data analysis is performed to quantify the absolute amount of each representative peptide within the sample.

#### 4.4.1 Validation of entire purification of IFT-A

For the stoichiometry determination of a protein complex of interest, it is necessary to purify the whole complex in its native state without any loss of compounds [58]. Therefore, purification of IFT-A was performed using three different bait proteins. Generated stable cell lines express one of three SF-TAP-tagged baits: IFT122, an internal component of the protein complex of interest; TULP3 which is known to be associated with IFT-A or LCA5 which is described to interact with the complex of interest in a labile and rather transient way (see Figure 27).



**Figure 27: Chosen baits for the purification of IFT-A**

To purify IFT-A, three different baits were chosen for the generation of Flp-In monoclonal lines. Either IFT122, TULP3 or LCA5 were chosen. These three baits are all related to the IFT-A. IFT122 is an integral part of the IFT-A, TULP3 is associated with it without being an integral part of IFT-A and LCA5 is a described interaction partner of IFT-A. All of the chosen baits are containing a SF-TAP tag for affinity purification.

##### 4.4.1.1 One-step affinity purification of IFT-A

To validate a proper and complete purification of IFT-A performing one-step affinity purification using different baits, purified proteins were precipitated and enzymatically cleaved and analysed by LC-MS/MS analysis for the identification of each complex component. As shown in Table 14, no matter which bait was chosen, all six complex components of IFT-A as well as the introduced bait were identified. In this table, the mean of exclusive unique peptide count and the percent of coefficient of variation of three biological replicates were depicted.

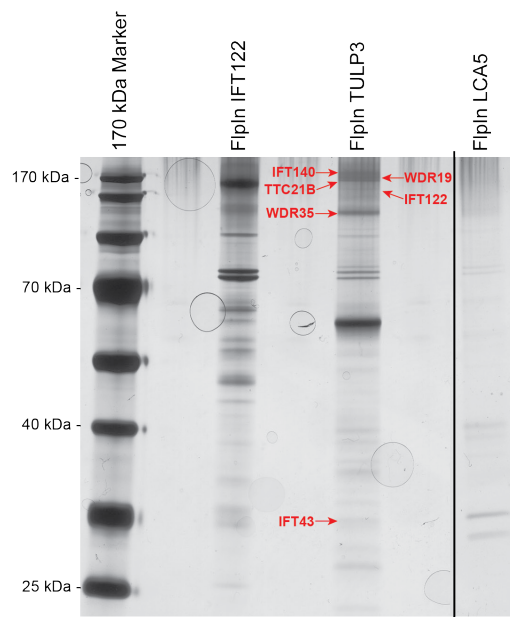
**Table 14: One-step affinity purification of generated Flp-In monoclonal lines**

Identified IFT-A complex component	Gene name	Exclusive Unique Peptide Count					
		Flp-In (N)-SF-IFT122		Flp-In (N)-SF-TULP3		Flp-In (N)-SF-LCA5	
		mean	%CV	mean	%CV	mean	%CV
Tubby-related protein 3	TULP3	0	-	23	4.9	8	25.0
Lebercilin	LCA5	0	-	0	-	46	14.3
Intraflagellar transport protein 122 homolog	IFT122	71	2.1	38	9.2	22	14.8
Intraflagellar transport protein 140 homolog	IFT140	29	38.4	42	10.4	28	20.1
Intraflagellar transport protein 43 homolog	IFT43	10	5.6	8	33.1	5	32.7
Tetratricopeptide repeat protein 21B	TTC21B	48	15.0	33	8.0	23	19.9
WD repeat-containing protein 19	WDR19	40	34.2	50	9.1	32	25.6
WD repeat-containing protein 35	WDR35	52	9.9	37	19.2	28	3.6

%CV: percent coefficient of variation

#### 4.4.1.2 Silver staining and MS analysis of TAP eluates

A second technique for the validation of an entire purification of the IFT-A was performed by SDS-PAGE of TAP-purified IFT-A of monoclonal Flp-In (N)-SF-IFT122, Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5 cells (as described in 3.2.3.3). After SDS-PAGE, silver staining with further in-gel digestion and MS analysis was performed to identify the six described complex components of the protein complex IFT-A (see 3.2.4.6).



**Figure 28: Validation of an entire purification of IFT-A using silver staining**

Purified IFT-A was separated by SDS-PAGE for further silver staining and in-gel digestion followed by MS analysis. As depicted, TAP eluates of Flp-In monoclonal lines, stably expressing either (N)-SF-TAP-tagged IFT122, TULP3 or LCA5 were used to validate the entire purification of all IFT-A complex components. Considering Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5, the protein patterns of purified IFT-A were comparable although the expression level varied. In comparison, the protein pattern of the cell line Flp-In (N)-SF-IFT122, stably expressing an internal part of the IFT-A, differed significantly. MS analysis identified all IFT-A components (highlighted in red).

As depicted in Figure 28, protein patterns of tested Flp-In monoclonal lines varied. While Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5 showed same patterns of purified IFT-A, the expression level of all IFT-A components was much weaker using LCA5 as bait. In comparison, an integral part of the IFT-A as bait (IFT122) resulted in an alteration within this protein pattern as depicted for Flp-In (N)-SF-IFT122. Additionally, all six complex components of the IFT-A were identified using MS analysis (highlighted with red arrows). This successful purification of the whole IFT-A complex, validated in this study, enables an accurate stoichiometry determination of the IFT-A protein complex by absolute quantification.

#### **4.4.2 Differences in complex stoichiometry of IFT-A according to chosen baits**

To achieve a robust and reliable determination of IFT-A composition, the complex stoichiometry from all three generated Flp-In monoclonal cell lines were investigated and compared.

##### **4.4.2.1 SRM assay using an equimolar standard mix of 14 standard peptides**

As an example for the analysis applying the SRM approach, the results of one biological replicate of IFT-A, purified from Flp-In (N)-SF-LCA5 cells, are depicted in Table 15. This sample was measured in three technical replicates. Four transitions of at least two representative peptides were quantified per complex component. Ratios of labelled to unlabelled peptide were calculated using Skyline. As shown in Table 15, median of all representative peptides of one protein was calculated and used for the determination of complex stoichiometry by normalising all values to the value of the lowest abundant protein (TTC21B). As depicted, the generated equimolar standard mix used for this SRM approach consists of 14 standard peptides although three representative peptides per complex component were ordered and synthesized. The poor solubility of four purchased peptides limited the number of available standard peptides to 14: Three representative peptides for the components IFT43 and TTC21B and two representative peptides for IFT122, IFT140, WDR19 and WDR35. For further details like standard deviation (SD) and coefficient of variation (CV), see Table 17 in the annex.

The determined median ratios (H/L) of IFT122, IFT43 and WDR19 showed huge variances within the chosen representative peptides. For WDR19, the representative peptides IFSLLEK and AASVYIR were chosen. While the median ratio (H/L) of all transitions for IFSLLEK was determined with 0.4635, the median ratio (H/L) of all transitions for AASVYIR was 0.2430. A look at two described isoforms of WDR19 (as depicted schematically in Figure 48 and in more detail in Figure 53), identified the presence of the representative peptide IFSLLEK in both known isoforms of WDR19, while the peptide AASVYIR is only present in the larger isoform of WDR19. For IFT43, four different isoforms are already described (see Figure 48 and Figure 51). While two of the chosen representative peptides (YSLATSR and VLAPEHEVR) are existing in

---

all four known isoforms of IFT43, the third chosen peptide (ASEEIDEFR) is only existing in two of the described isoforms. Appropriately, determined median ratios (H/L) of all transitions for YSLATSR and VLAPEHEVR were virtually identical (0.1734 for YSLATSR and 0.1664 for VLAPEHEVR), while the determined value for ASEEIDEFR differed (0.0694). In contrast, both analysed peptides for IFT122 (WDEAFALGEK and ILFTLAK) are existing in all ten described isoforms of IFT122 (for more details see Figure 48 and Figure 49), however the determined median of ratios (H/L) of all transitions for both representative peptides unveiled huge differences (0.3246 for WDEAFALGEK and 1.8846 for ILFTLAK). This significant variety leads to the hypothesis of other existing isoforms which are not described and characterized so far.

Table 15: Absolute quantification of purified IFT-A using Flp-In LCA5 performing SRM

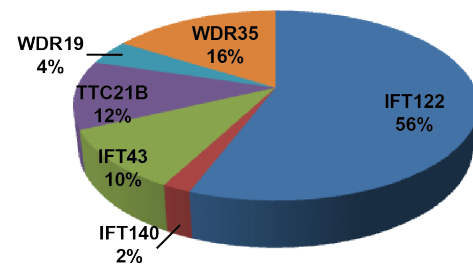
Protein	Peptide	Pre. m/z	Prod. m/z	Trans.	RT	CE	Ratio H/L				
							Trans.	Median <sub>Pep</sub>	Median <sub>Prot</sub>	Stoich.	
IFT122	WDEAFALGEK	583.279846	864.446159	+2y8	43.97	26.9	0.3212	0.3246	1.1046	6.77	
			735.403566	+2y7	43.97	28.9	0.3279				
			664.366452	+2y6	43.97	27.9	0.3043				
	333.176861	+2y3	43.97	34.9	0.3506						
	ILFTLAK	403.262739	692.434138	+2y6	41.30	19.4	2.3033				1.8846
			579.350074	+2y5	41.30	18.4	1.8792				
432.28166			+2y4	41.30	24.4	1.89					
IFT140	FWDIER	433.213777	718.351865	+2y5	40.53	20.5	0.2319	0.2316	0.2031	1.25	
			532.272552	+2y4	40.53	20.5	0.2259				
			417.245609	+2y3	40.53	28.5	0.2316				
	AHGALTEAYK	530.774731	852.446159	+2y8	16.47	26.0	0.1683				0.1747
			795.424696	+2y7	16.47	27.0	0.1652				
			724.387582	+2y6	16.47	31.0	0.2041				
IFT43	YSLATSR	399.211235	381.213246	+2y3	16.47	33.0	0.181	0.1734	0.1664	1.02	
			634.351865	+2y6	20.45	21.2	0.2602				
			547.319837	+2y5	20.45	21.2	0.1663				
	ASEEIEDFR	548.251285	434.235772	+2y4	20.45	20.2	0.1631				0.0694
			363.198659	+2y3	20.45	20.2	0.1805				
			808.383559	+2y6	30.52	23.6	0.073				
TTC21B	YGSDPVFR	470.729792	679.340966	+2y5	30.52	25.6	0.0657	0.1632	0.1632	1.00	
			566.256902	+2y4	30.52	24.6	0.0652				
			437.214309	+2y3	30.52	27.6	0.0819				
	EAILEDAR	502.256371	837.421342	+2y7	16.38	25.8	0.1492				0.1664
			766.384228	+2y6	16.38	26.8	0.1638				
			669.331464	+2y5	16.38	32.8	0.1941				
WDR19	LEDVPR	364.700503	540.288871	+2y4	16.38	32.8	0.1689	0.1609	0.3532	2.17	
			777.388979	+2y7	28.44	21.8	0.1701				
			720.367515	+2y6	28.44	21.8	0.1588				
	IFSLLEK	425.257654	633.335487	+2y5	28.44	21.8	0.1629				0.4635
			518.308544	+2y4	28.44	28.8	0.1634				
			803.425758	+2y7	23.56	24.9	0.2465				
WDR35	AASVYIR	390.224145	690.341694	+2y6	23.56	23.9	0.177	0.2430	0.2231	1.37	
			577.25763	+2y5	23.56	23.9	0.1792				
			448.215037	+2y4	23.56	21.9	0.1906				
	DLAIGLR	379.23197	615.309666	+2y5	18.61	19.0	0.181				0.2337
			486.267073	+2y4	18.61	18.0	0.1609				
			371.24013	+2y3	18.61	25.0	0.1549				
WDR35	YASHLLEK	480.761092	457.22929	+2b4	18.61	17.0	-	0.2124	0.2231	1.37	
			736.423967	+2y6	42.39	20.2	0.4188				
			589.355553	+2y5	42.39	20.2	0.9484				
	DLAIGLR	379.23197	502.323525	+2y4	42.39	26.2	0.4324				0.2337
			389.239461	+2y3	42.39	28.2	0.4946				
			708.403901	+2y6	21.65	19.9	0.3227				
WDR35	YASHLLEK	480.761092	637.366787	+2y5	21.65	19.9	0.2398	0.2124	0.2231	1.37	
			550.334758	+2y4	21.65	20.9	0.2461				
			451.266344	+2y3	21.65	20.9	0.2302				
	DLAIGLR	379.23197	642.429721	+2y6	38.13	22.5	-				0.2337
			529.345657	+2y5	38.13	18.5	0.2222				
			458.308544	+2y4	38.13	18.5	0.2337				
WDR35	YASHLLEK	480.761092	345.22448	+2y3	38.13	17.5	0.239	0.2124	0.2231	1.37	
			797.451579	+2y7	18.88	24.2	0.1976				
			726.414465	+2y6	18.88	24.2	0.2181				
	DLAIGLR	379.23197	502.323525	+2y4	18.88	29.2	0.2067				0.2124
			389.239461	+2y3	18.88	31.2	0.2607				
			797.451579	+2y7	18.88	24.2	0.1976				

Pre. m/z: precursor m/z  
 Prod. m/z: Product m/z  
 Trans.: Transition  
 RT: retention time  
 CE: collision energy  
 Pep: peptide  
 Prot: protein  
 Stoich: stoichiometry  
 Ratio H/L: Ratio labelled peptide to unlabelled peptide

The determined IFT-A stoichiometries using the three different baits are shown in Figure 29. This figure shows the data of four biological replicates, measured in three technical replicates, each. The complex stoichiometry of IFT-A, purified from Flp-In (N)-SF-TULP3 or Flp-In (N)-SF-LCA5 cells, leads to virtually identical complex composition. IFT122 represented the most abundant protein with 48% followed by WDR19 (16%) and WDR35 (11%). TTC21B and IFT43 were less abundant with a proportion of approximately 8%. In contrast and supported by the statement of Gibson et al. in 2013 [58], an introduced overexpression, even though these are no transient overexpressions, of an internal component of the IFT-A (IFT122) leads to a significantly altered complex composition. Analogous to IFT-A complexes isolated with TULP3 and LCA5 as bait, the most abundant protein was as well IFT122 (56%) but with a clearly higher abundance. However, in this experiment WDR19 and IFT140 represented the lowest abundant proteins (4%, 2%) here and clearly differed from Flp-In (N)-SF-LCA5 and Flp-In (N)-SF-TULP3. This demonstrates that even though stable expression was achieved, the artificial expression of an integral component of a protein complex can lead to a severely disturbed stoichiometry.

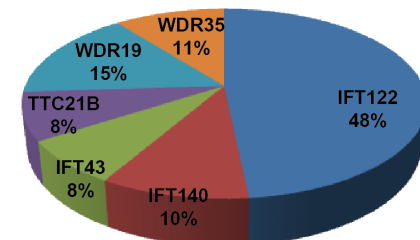
#### Flp-In IFT122

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	28.7	3.8	13.2
IFT140	1.0	0.0	0.0
IFT43	5.3	0.4	8.0
TTC21B	6.0	0.1	2.1
WDR19	2.2	0.1	2.4
WDR35	8.3	0.4	4.3



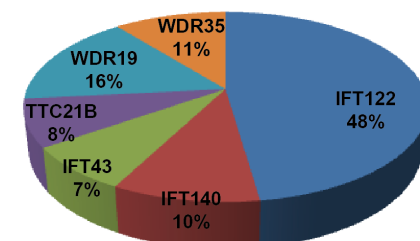
#### Flp-In TULP3

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	6.5	1.8	28.1
IFT140	1.3	0.1	7.9
IFT43	1.0	0.0	0.0
TTC21B	1.1	0.1	6.9
WDR19	2.1	0.4	17.3
WDR35	1.4	0.1	8.7



#### Flp-In LCA5

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	6.3	0.6	9.7
IFT140	1.3	0.1	7.5
IFT43	1.0	0.0	1.0
TTC21B	1.1	0.1	6.7
WDR19	2.1	0.1	4.5
WDR35	1.4	0.1	7.2



**Figure 29: Stoichiometry of purified IFT-A using different baits performing SRM**

To determine the stoichiometry of IFT-A, Flp-In monoclonal lines expressing either (N)-SF-TAP-tagged IFT122, TULP3 or LCA5 were used for one-step affinity purification. Then, absolute quantification was performed using a SRM approach. Using IFT-A-associated proteins or interactors of IFT-A as bait resulted in very similar results, whereas the use of an internal component of the IFT-A resulted in a significantly different composition.

#### 4.4.2.2 PRM Assay using an equimolar standard mix of 24 standard peptides

After a QExactive Plus instrument was available in our lab, a Parallel Reaction Monitoring (PRM) approach was set up. To have at least three representative peptides available for the quantification of each protein, a new equimolar standard mix containing additional representatives was generated. As described in 4.3.4.1, representative peptides which are not existing in all the described isoforms of the corresponding IFT-A component led to significant variances in the determined median ratios (H/L) of the analysed transitions. To circumvent this error-prone quantification, additional representative peptides, ideally present on all described isoforms of the corresponding IFT-A protein, were chosen. This equimolar standard mix consists of three representative peptides for IFT140, IFT43, TTC21B and WDR35 and of six representative peptides for IFT122 and WDR19 (see Table 16). Due to the big difference in determined ratio of unlabelled to labelled peptide of two different representative peptides of the same IFT-A component, shown in further experiments (see Table 15), the number of standard peptides for IFT122 and WDR19 was significantly increased. As an example, detailed information of the analysis of one biological replicate from Flp-In (N)-SF-TULP3 is depicted in Table 16. For further details like standard deviation (SD) and percent coefficient of variation see Table 18. Each IFT-A complex component is depicted in a different colour. Colour gradation highlights different product ions, included in the calculation. Analysis was performed the same way as for the SRM approach. Ratio of labelled to unlabelled transitions was determined using Skyline. After calculating the median of all ratios of fragment ions within a representative peptide, the median of all representative peptides per protein was determined, to diminish the influence of individual representative peptides which are probably not present in all existing isoforms (described as well as not yet described isoforms) of the corresponding IFT-A proteins. The distribution of the representative peptides on described isoforms of the corresponding IFT-A components is depicted in Figure 48. For the stoichiometry determination of IFT-A, all ratios were normalized to the value of the lowest abundant protein (IFT43).



Table 16: Absolute quantification of IFT-A using Flp-In (N)-SF-TULP3 performing PRM

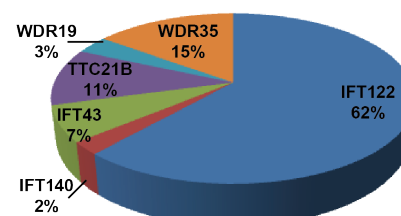
Protein	Peptide	Precursor m/z	Fragment Ion	RT	Ratio H/L			Stoich.	
					Fragment Ion	Median <sub>Peptide</sub>	Median <sub>Protein</sub>		
IFT122	DSIGDEDPFTAK	651.800614	y10	53.71	0.8024	0.7387	0.7074	1.80	
			y6		0.7132				
			y7		0.7132				
			y8		0.7434				
	DTPSGISK	406.715624	y4	21.69	0.3414	0.3414			
			y5		0.3019				
			y6		0.4232				
	GEYILLGGSDK	580.307878	y7	51.66	0.6749	0.6761			
			y8		0.6761				
			y9		0.7025				
	ILFTLAK	407.269839	y4	54.63	0.9516	0.9516			
			y5		1.3372				
y6			0.8298						
WDEAFALGEK	587.286945	y6	57.59	1.1704	1.0113				
		y7		0.9332					
		y8		0.9653					
		y9		1.0574					
YLELISSIEER	681.362959	y6	67.98	0.7501	0.6544				
		y7		0.5174					
		y8		0.5831					
		y9		0.7258					
IFT140	AHGALTEAYK	534.78183	25.59	0.7364	0.6729	0.6729	1.70		
				y6				0.6396	
				y7				0.6606	
				y8				0.6851	
SHLFVDEGLK	576.810587	42.42	51.97	0.7459	0.7393				
				y4		0.7327			
				y5		0.5854			
				y6		0.4631			
IFT43	ASEEIEDFR	553.25542	40.56	0.4969	0.1900	0.3959	1.00		
				y4				0.2023	
				y5				0.1898	
				y6				0.1918	
	VLAPEHEVR	530.294682	26.44	28.77	0.1683			0.3959	
					y4				0.1900
					y5				0.3371
					y6				0.4550
	YSLATSR	404.215369	28.77	33.09	0.3959			0.5853	
					y4				0.3723
					y5				0.4796
					y6				0.5853
TTC21B	EAILES DAR	507.260505	33.09	0.6253	0.6347	0.5743	1.45		
				y4				0.4938	
				y5				0.5960	
				y6				0.6870	
LEDVPR	369.704637	27.9	38.27	0.6347	0.4429				
				y3		0.4036			
				y4		0.4977			
				y5		0.4429			
YGDSPVFR	475.733926	38.27	30.51	0.6111	0.5743				
				y4		0.6012			
				y5		0.5340			
				y6		0.5474			
WDR19	AASVYIR	395.22828	30.51	0.6282	0.6492	0.6002	1.52		
				y3				0.6703	
				y4				0.7151	
				y5				0.4904	
	DGDVLAVIAEK	569.315703	63.31	55.7	0.1573			0.1573	
					y6				0.1573
					y7				0.1592
					y8				0.1408
	IFSLLEK	429.264753	48.6	48.6	0.3657			0.3657	
					y4				0.3210
					y5				0.4165
					y6				0.6151
LVFIDEK	436.254385	48.6	48.6	0.8795	0.6151				
				y4		0.2704			
				y5		0.6151			
				y6		0.8795			
VGDLLPHVSSPK	628.858069	44.1	44.1	0.2704	0.5853				
				y10		0.6379			
				y11		0.4778			
				y6		0.7080			
				y7		0.6302			
				y8		0.5403			
VGSFLAVGTVK	543.325873	52.86	52.86	0.5115	1.0947				
				y10		1.3723			
				y6		1.0876			
				y7		0.9379			
WDR35	EIGSLAR	434.759944	46.17	1.0947	0.6136	0.5381	1.36		
				y4				1.2419	
				y5				0.6749	
				y6				0.5610	
THVIAASK	417.749802	17.93	17.93	0.5268	0.4928				
				y5		0.6136			
				y6		0.4588			
				y4		0.4968			
YASHLLEK	484.768191	28.42	28.42	0.5081	0.5381				
				y5		0.5081			
				y6		0.5681			
				y7		0.5806			

RT: retention time  
Stoich: stoichiometry  
Ratio H/L: Ratio labelled peptide to unlabelled peptide

To validate the result of stoichiometry alterations induced by the use of IFT122 as bait, absolute quantification of IFT-A from all three generated Flp-In clones was performed using PRM and the equimolar standard mixture consisting of 24 representative peptides in total. Four biological replicates were prepared for each cell line and each sample was measured twice (see Figure 30).

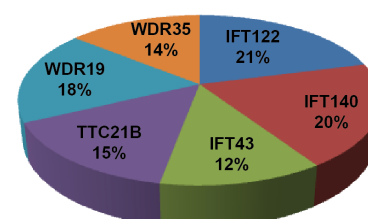
#### Flp-In IFT122

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	26.0	1.1	4.2
IFT140	1.0	0.0	0.0
IFT43	2.9	0.4	13.7
TTC21B	4.7	0.3	6.1
WDR19	1.4	0.3	17.8
WDR35	6.3	0.6	9.9



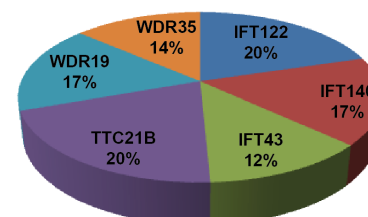
#### Flp-In TULP3

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	1.7	0.2	13.5
IFT140	1.6	0.1	5.4
IFT43	1.0	0.0	0.0
TTC21B	1.2	0.1	9.6
WDR19	1.5	0.0	2.0
WDR35	1.2	0.1	8.9



#### Flp-In LCA5

IFT-A component	Mean <sub>Stoichiometry</sub>	SD	%CV
IFT122	1.7	0.2	12.2
IFT140	1.3	0.1	9.7
IFT43	1.0	0.1	5.8
TTC21B	1.6	1.0	59.2
WDR19	1.4	0.8	58.4
WDR35	1.1	0.2	15.2



**Figure 30: PRM-based stoichiometry of purified IFT-A using three different baits**

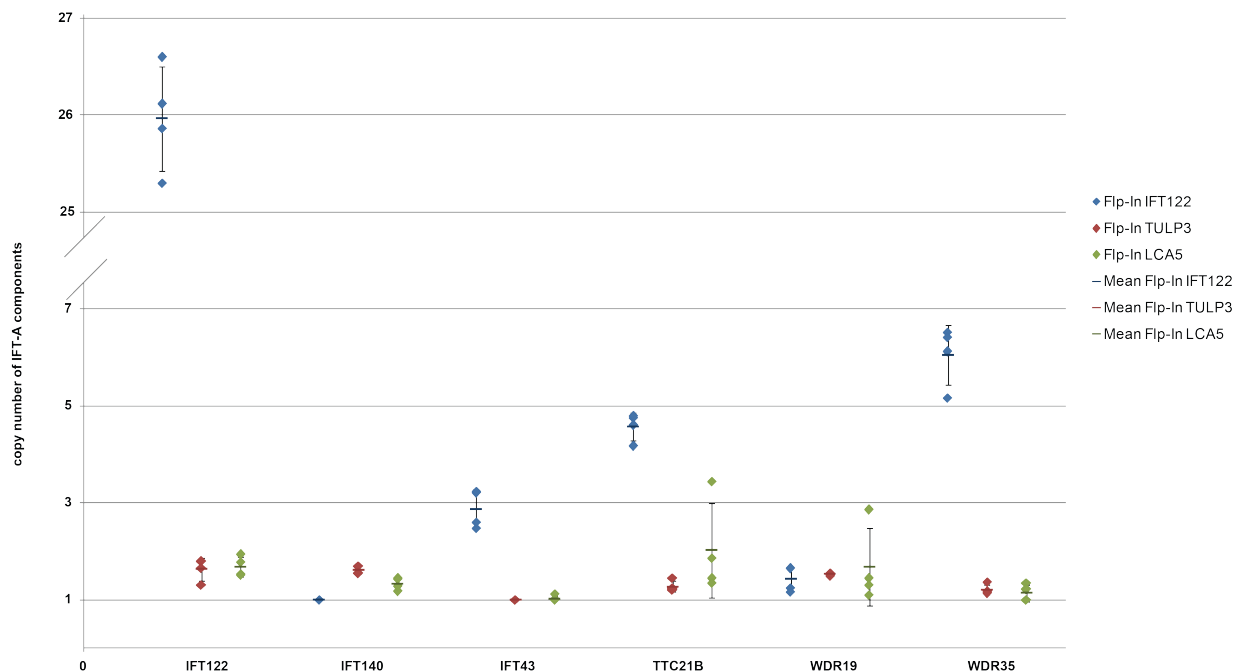
Absolute quantification performing PRM of purified IFT-A from the three Flp-In monoclonal cell lines revealed differences in IFT-A stoichiometry depending on whether an integral part of the protein complex is used as bait (as in case of IFT-A purified from (N)-SF-IFT122) or bait proteins that are associated with the IFT-A ((N)-SF-TULP3) or labile interaction partners ((N)-SF-LCA5) of the protein complex. For the calculation of the stoichiometry, four biological replicates, measured as two technical replicates, were consulted.

Comparing the results from all three cell lines, determined stoichiometries of IFT-A from Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5 were again virtually identical, whereas the composition of IFT-A using IFT122 as bait protein significantly differed. Considering the coefficient of variance (CV) for stoichiometries obtained with IFT122 and TULP3 as baits, CV values were below 20%. Analyzing IFT-A composition, purified from Flp-In (N)-SF-LCA5 cells, the variance especially for WDR19 and TTC21B is quite high. As mentioned before, LCA5 is a rather labile and transiently bound interaction partner of IFT-A, resulting in a lower amount of purified IFT-A using Flp-In (N)-SF-LCA5 cells. Due to this low abundance of IFT-A components, the absolute quantification of the representative peptides is error-prone. Whereas IFT122 predominated in

IFT-A preparations from Flp-In (N)-SF-IFT122 (with a total amount of about 60%), the amount of IFT122, IFT140 and TTC21B was similar (around 20%) in Flp-In cell lines with SF-TAP-tagged TULP3 or SF-TAP-tagged LCA5 as bait. Thereby, the described core complex of IFT-A, formed by IFT122, IFT140 and WDR19 represented the major portion with 59% percent in total. The least abundant protein for IFT-A purified from Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5 was IFT43, whereas IFT140 and WDR19 were the least abundant proteins in purified IFT-A from Flp-In (N)-SF-IFT122 cells. Considering the determined IFT-A stoichiometry in Flp-In (N)-SF-TULP3 cells and the molecular masses of the IFT-A components (IFT122: 141.8kDa, IFT140: 165.2kDa, IFT43: 23.5kDa, TTC21B: 150.9kDa, WDR19: 151.6kDa and WDR35: 133.5kDa), the mass of the IFT-A protein complex is a multiple of 1,098kDa.

#### 4.4.3 Reproducibility

For the study and understanding of the underlying mechanisms of IFT-A-related ciliopathies, it is important to know how reliable the determined protein complex stoichiometry is. Therefore, reproducibility of absolute quantification is a crucial point. To evaluate the reproducibility of this approach, four biological replicates of purified IFT-A, for each of the generated Flp-In monoclonal line, were analysed in duplicates. Results are shown in Figure 31.



**Figure 31: Reproducibility of absolute quantification using different baits**

An important factor to evaluate the determined complex stoichiometry is the reliability of the absolute quantification. Therefore, the reproducibility of four biological replicates per generated cell line, each measured twice, is depicted. Biological replicates of the three chosen Flp-in monoclonal cell lines are depicted as rhombuses and highlighted in different

colours: Flp-In (N)-SF-IFT122 is depicted in blue, Flp-In (N)-SF-TULP3 in red and Flp-In (N)-SF-LCA5 is highlighted in green. The mean copy number of each IFT-A component of analysed biological replicates is depicted as horizontal line. The error bars represent the standard deviation. Biggest variation of biological replicates is apparent in experiments using Flp-In (N)-SF-IFT122 cells. In contrast, results of biological replicates using SF-TAP-tagged TULP3 and SF-TAP-tagged LCA5 are highly reproducible. However, there are outliers for the stoichiometry of TTC21B and WDR19 using Flp-In (N)-SF-LCA5 cells. Flp-In (N)-SF-TULP3 obtained most reliable data performing absolute quantification of IFT-A. Therefore, this cell line was used for further investigations.

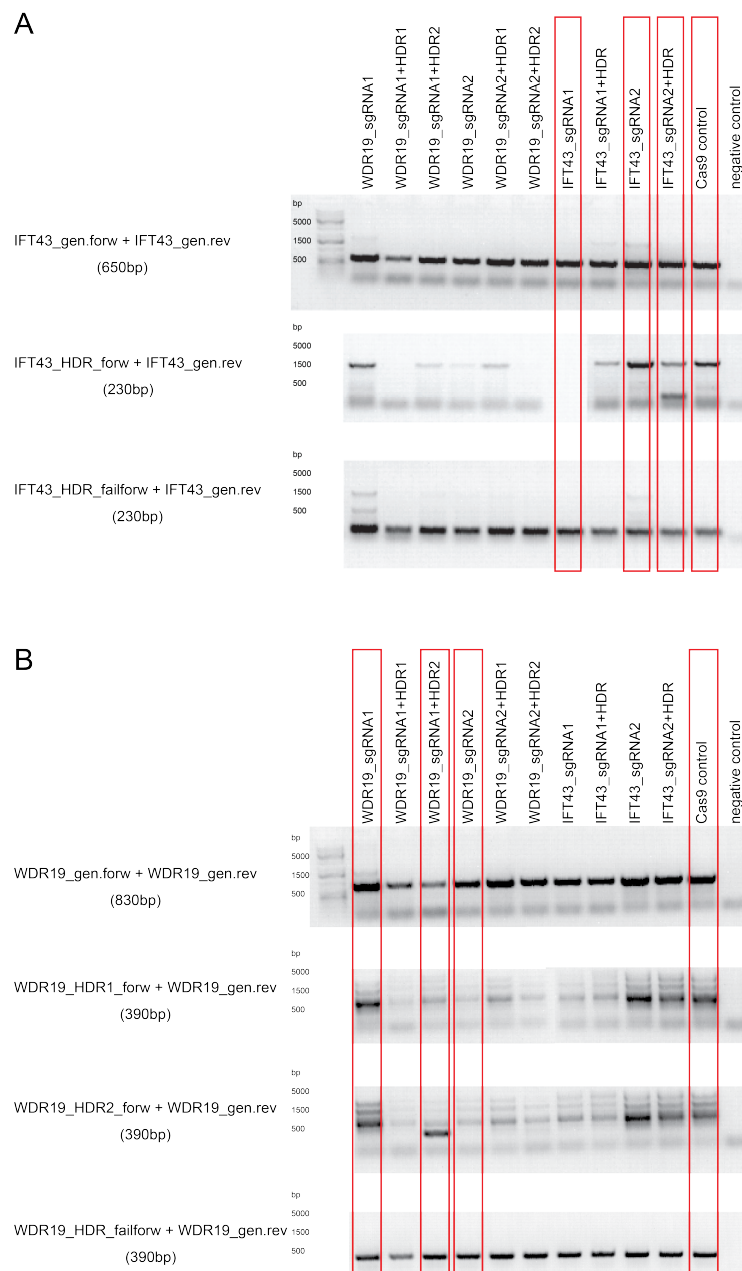
## **4.5 Induced alterations in IFT-A complex stoichiometry**

### **4.5.1 Gene editing of IFT-A components using the CRISPR/Cas9 system**

The advanced CRISPR/Cas9 system was used to generate targeted mutations in genes encoding IFT-A complex components. As shown previously, Flp-In (N)-SF-TULP3 yielded comparable stoichiometries of IFT-A compared to Flp-In (N)-SF-LCA5 but achieved better results regarding reproducibility. Therefore, previously generated Flp-In (N)-SF-TULP3 monoclonal cells were used for gene editing as described in 3.2.2.6.

#### **4.5.1.1 Validation of generated CRISPR clone mixtures using PCR**

Validation of CRISPR/Cas9-induced targeted gene editing was performed by PCR. As depicted in Figure 32, different single guide RNAs (sgRNA) were used to target either WDR19 or IFT43. To enable homologous directed repair (HDR), HDR constructs were also co-transfected in some experiments. Based on expected PCR products of manipulated clones, Flp-In (N)-SF-TULP3 transfected with Cas9 construct bearing an additional cistron for the expression of IFT43\_sgRNA1 or IFT43\_sgRNA2 or with the combination of IFT43\_sgRNA2 and a co-transfected HDR construct showed promising results (present PCR product for IFT43 (650bp) with or without an additional PCR fragment indicating a successful integration of the HDR construct (230bp)) and are highlighted in red boxes in Figure 32. These cell mixtures were selected for single clone selection. Similarly, transfections of Flp-In (N)-SF-TULP3 with constructs bearing WDR19\_sgRNA1, WDR19\_sgRNA2 as well as transfections with the combination of WDR19\_sgRNA1 with HDR construct 2 (HDR2) were selected for single clone selection. To characterize mutations of single clones, DNA-sequencing was performed (see Figure 33 and Figure 34).



**Figure 32: Validation of generated CRISPR clone mixtures via PCR**

(N)-SF-TULP3 cells transfected with a Cas9 construct bearing an additional cistron for various sgRNAs or co-transfected with an additional HDR construct were characterized performing Polymerase Chain Reaction (PCR). Different primer pairs were used depending on applied sgRNA with or without a HDR construct. As expected, PCR fragments for IFT43 and WDR19 are present in every tested mixture. The clone mixture of Flp-In (N)-SF-TULP3 treated with IFT43\_sgRNA2 and the HDR construct showed an additional fragment and was chosen with cells transfected only either with IFT43\_sgRNA1 or IFT43\_sgRNA2 (highlighted in red (A)) for single clone selection. Tested clone mixture of Flp-In (N)-SF-TULP3 transfected with WDR19\_sgRNA1, WDR19\_sgRNA2 and cells transfected with the combination of WDR19\_sgRNA1 and HDR construct showed expected PCR products for a successful editing and were selected for single clone selection. As a control, Flp-In cells transfected with an “empty” Cas9 vector (lacking sgRNA) were also chosen for single clone selection.

#### 4.5.1.2 Validation of CRISPR/Cas9 edited single clones via Sequencing

Flp-In (N)-SF-TULP3 transfected with Cas9/sgRNA constructs were seeded at low density and cultured in DMEM supplemented with puromycin for single clone selection (see 3.2.2.6). Individual clones were then characterized by DNA-sequencing. Therefore, genomic DNA was extracted and PCR products were analysed. As depicted in Figure 33, two different single clones with following mutations in the IFT43 gene were identified: one clone with a single base insertion of an adenine (c.541\_542insA) and one clone with a single base insertion of a guanine (c.582\_583insG). The c.541\_542insA was observed upon the use of IFT43\_sgRNA1 for the gene editing experiments. Due to the frame shift, this mutation results in an amino acid exchange of threonine (T) to asparagines (N) at amino acid position 181 followed by a premature stop (p.T181Nfs\*2) in the reading frame for IFT43.

<b>IFT43_sgRNA1</b>	
→	<b>c.541_542insA</b>
→	<b>p.T181Nfs*2</b>
DNA_wt:	Exon1xxx...xxxExon7xxxGATGATGTCGGCTGGGACTGGGACCATCTGTTCACTGAGGTGTCCTCAGAGGTCCTC ACTGAGTGGGACCCACTGCAGACGGAGAAGGAGGACCCCTGCGGGGCAGGCCAGGCACACCTGA
DNA_mut:	Exon1xxx...xxxExon7xxxGATGATGTCGGCTGGGACTGGGACCATCTGTTCA <b>A</b> CTGAGGTGTCCTCAGAGGTCCTC ACTGAGTGGGACCCACTGCAGACGGAGAAGGAGGACCCCTGCGGGGCAGGCCAGGCACACCTGA
Protein_wt:	Exon1xxx...xxxExon7xxxDDVGWDWDHLFTEVSSEVLTEWDPLQTEKEDPAGQARHT
Protein_mut:	Exon1xxx...xxxExon7xxxDDVGWDWDHLF <b>N*</b>
<b>IFT43_sgRNA2</b>	
→	<b>c.582_583insG</b>
→	<b>p.Q195Afs*22</b>
DNA_wt:	Exon1xxx...xxxExon7xxxGATGATGTCGGCTGGGACTGGGACCATCTGTTCACTGAGGTGTCCTCAGAGGTCCTC ACTGAGTGGGACCCACTGCAGACGGAGAAGGAGGACCCCTGCGGGGCAGGCCAGGCACACCTGA
DNA_mut:	Exon1xxx...xxxExon7xxxGATGATGTCGGCTGGGACTGGGACCATCTGTTCACTGAGGTGTCCTCAGAGGTCCTC ACTGAGTGGGACCCACT <b>G</b> GCAGACGGAGAAGGAGGACCCCTGCGGGGCAGGCCAGGCACACCTGA
Protein_wt:	Exon1xxx...xxxExon7xxxDDVGWDWDHLFTEVSSEVLTEWDPLQTEKEDPAGQARHT*
Protein_mut:	Exon1xxx...xxxExon7xxxDDVGWDWDHLFTEVSSEVLTEWDPL <b>ADGEGGPCGAGQAHLSPSPML*</b>

**Figure 33: Generated mutations in IFT43 using CRISPR/Cas9 system**

CRISPR/Cas9-based targeted editing in Flp-In monoclonal line stably expressing SF-TAP-tagged TULP3 resulted in two different mutants of IFT43: Application of IFT43\_sgRNA1 led to a single base insertion of an adenine (c.541\_542insA). This insertion induces a frame shift and leads to a truncated IFT43 protein with an amino acid exchange of threonine (T) to asparagines (N) and a premature stop (p.T181Nfs\*2). In contrast, using IFT43\_sgRNA2, a mutation (c.582\_583insG) that results in a prolonged protein is predicted. This protein is characterized by an amino acid exchange from glutamine (Q) to alanine (A) followed by 22 additional novel amino acids translated from the altered reading frame (p.Q195Afs\*22).

The presence of a premature stop codon may provoke a known surveillance pathway present in all eukaryotes, the so called nonsense-mediated mRNA decay (NMD). Based on the fact that a translation of a protein with an introduced premature stop codon leads in most instances to a

non-functional or to a down regulated activity of the resulting protein, this mechanism eliminates those altered mRNA constructs preventively [80]. The single base insertion c.582\_583insG obtained with the IFT43\_sgRNA2 results in a frame shift which causes an amino acid exchange from a glutamine (Q) to an alanine (A) residue at amino acid position 195 followed by additional 21 amino acid residues (p.Q195Afs\*22) translated from an alternate reading frame. This actually resulted in a prolonged protein compared to the wildtype IFT43 (Figure 33).

#### WDR19\_sgRNA1

- c.3107\_3108delAT
- p.Y1036Ffs\*5

DNA wt: Exon1xxx...xxxExon26xxxGTTCTGAAGACACTACTAATGAAGACTATCAAAGCATTGCCTTATACTTTGAAGGAGA  
AAAGAGATATCTTCAGGCTGGAAAATTCTTCTTGCTGTGTGGCCAATATTCACGAXXXExon28xxx...xxxExon36

DNA mut: Exon1xxx...xxxExon26xxxGTTCTGAAGACACTACTAATGAAGACTATCAAAGCATTGCCTTATACTTTGAAGGAGA  
AAAGAGATATCTTCAGGCTGGAAAATTCTTCTTGCTGTGTGGCCAAT[---]TCACGAXXXExon28xxx...xxxExon36

Protein wt: Exon1xxx...xxxExon26xxxSEDTTNEYQSIYLFEGEKRYLQAGKFFLLCGQYSRxxxExon28xxx...xxxExon36

Protein mut: Exon1xxx...xxxExon26xxxSENTTNEYQSIYLFEGKRYLQAGKFFLLCGQ**FTST**\*xxxExon28xxx...xxxExon36

#### WDR19\_sgRNA2

- c.3063\_3074delinsAGATC
- p.R1022Dfs\*21

WDR19\_sgRNA2 c.3063\_3074delinsAGATC R1022Dfs\*21

DNA wt: Exon1xxx...xxxExon26xxxGTTCTGAAGACACTACTAATGAAGACTATCAAAGCATTGCCTTATACTTTGAAGGAGA  
AAAGAGATATCTTCAGGCTGGAAAATTCTTCTTGCTGTGTGGCCAATATTCACGAXXXExon28xxx...xxxExon36

DNA mut: Exon1xxx...xxxExon26xxxGTTCTGAAGACACTACTAATGAAGACTATCAAAGCATTGCCTTATACTTTGAAGGAGA  
AAAA**AGATC**[-----]GGCTGGAAAATTCTTCTTGCTGTGTGGCCAATATTCACGAXXXExon28xxx...xxxExon36

Protein wt: Exon1xxx...xxxExon26xxxSEDTTNEYQSIYLFEGEKRYLQAGKFFLLCGQYSRxxxExon28xxx...xxxExon36

Protein mut: Exon1xxx...xxxExon26xxxSENTTNEYQSIYLFEGEK**DRELNSSCCVANIHEHLNTS**\*xxxExon28xxx...xxxExon36

### Figure 34: Generated mutations in WDR19 using CRISPR/Cas9 system

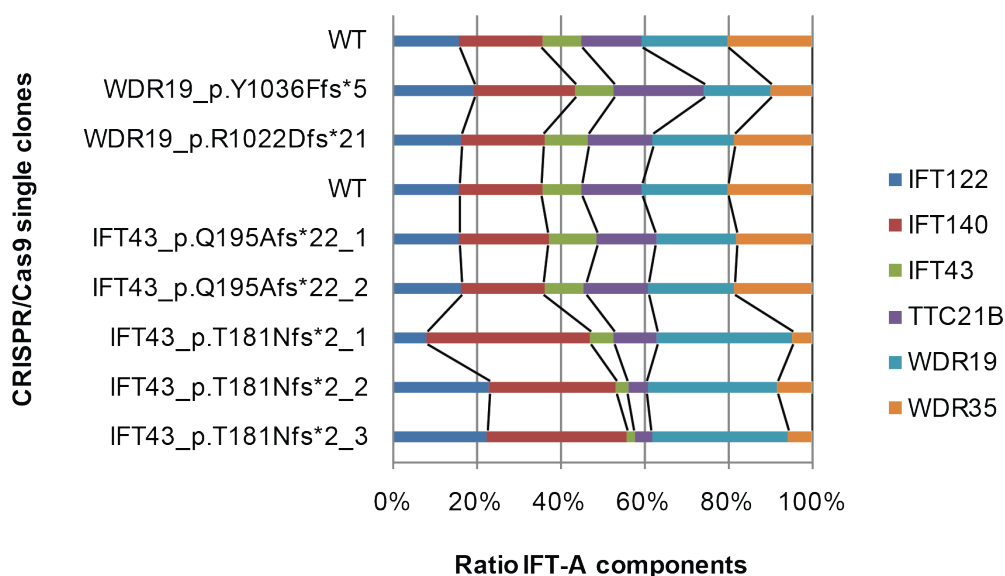
Performing gene editing using CRISPR/Cas9 system resulted in two different mutants of WDR19: Application of WDR19\_sgRNA1 led to a deletion of two bases (c.3107\_3108delAT). Based on an induced frame shift, this leads to an amino acid exchange of tyrosine (Y) to phenylalanine (F) followed by a premature stop (p.Y1036Ffs\*5). The second mutation in WDR19 was obtained applying WDR19\_sgRNA2. Thereby, eleven basepairs were deleted and five basepairs inserted, afterwards (c.3063\_3074delinsAGATC). The resulting frame shift leads to an amino acid exchange from arginine (R) to aspartate (D) at amino acid position 1022 followed by 21 novel amino acid residues (p.R1022Dfs\*21).

Obtained mutations within the WDR19 gene are shown in Figure 34. Again, two mutations within single clones were validated by DNA sequencing. One clone, obtained from gene editing with WDR19\_sgRNA1, had a deletion of two bases (c.3107\_3108delAT) and the second clone, obtained with the WDR19\_sgRNA2, had a complex insertion/ deletion mutation with eleven basepairs deleted and five basepairs inserted (c.3063\_3074delinsAGATC). The c.3107\_3108delAT causes a frame shift and results in an amino acid exchange at position 1036 followed by a premature stop (p.Y1036Ffs\*5). The second identified mutations

(c.3063\_3074delinsAGATC) also introduces a frame shift leading to a protein characterized by an amino acid exchange of an arginine (R) to an aspartate (D) followed by 21 novel amino acid residues (p.R1022Dfs\*21).

#### 4.5.2 Absolute quantification of IFT-A in CRISPR/Cas9-induced mutant cell clones

To determine the potential effect of mutant IFT43 as well as of mutant WDR19 on the composition and stoichiometry of IFT-A, absolute quantification was performed using IFT-A purified from single clones carrying mutations either in WDR19 or IFT43 as well as in control cells. Results from the first pilot experiment are shown in Figure 35. Most mutant clones did not show a significant effect on complex composition of IFT-A, except clones carrying the p.T181Nfs\*2 mutation in IFT43. Sequencing results of all three tested single clones, containing the IFT43\_p.T181Nfs\*2 mutation, were identical. However, differences on protein level were determined.



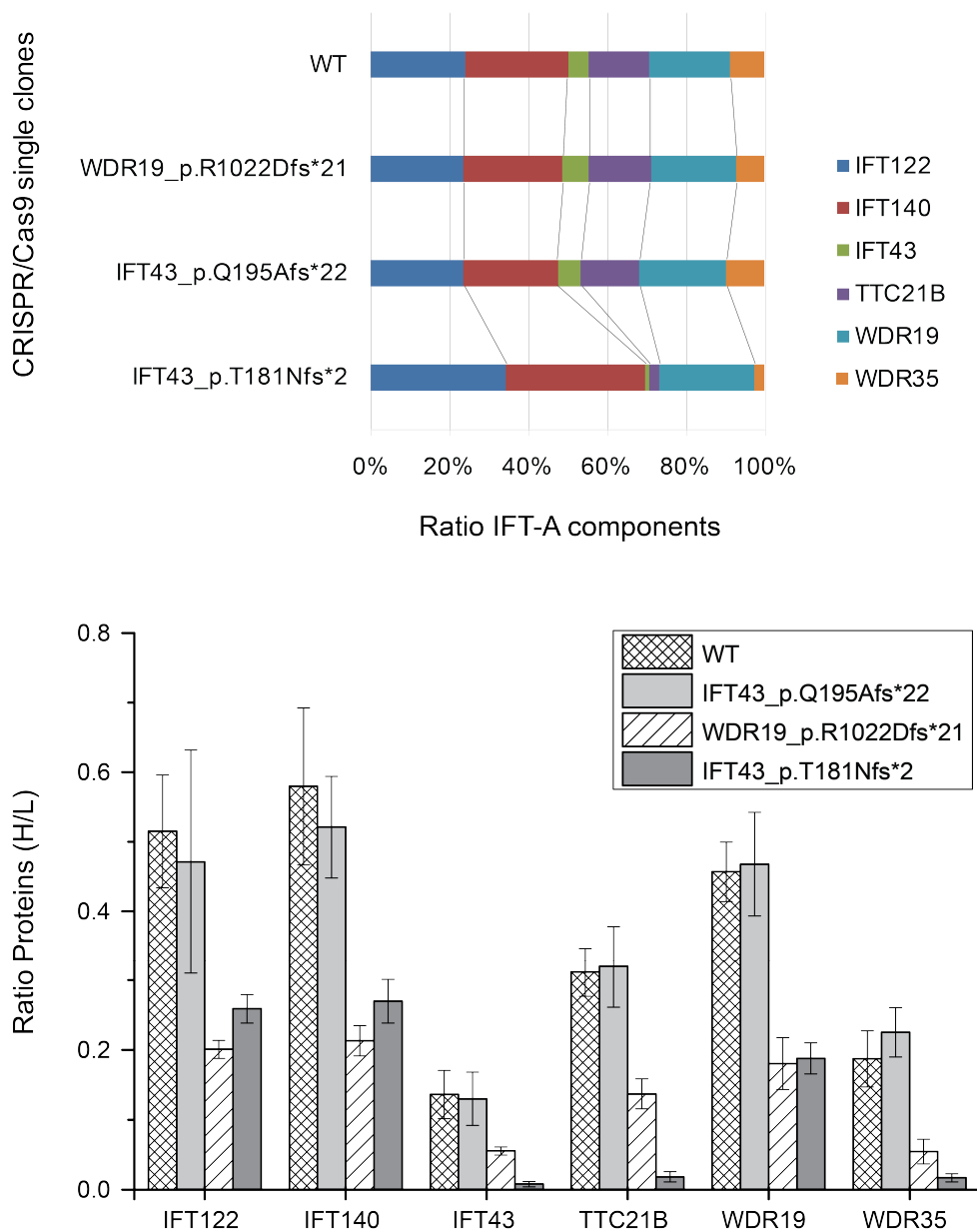
**Figure 35: Effect on IFT-A composition of all identified single clones**

Absolute quantification of CRISPR/Cas9-induced mutant clones was performed to identify effects on the composition and stoichiometry of IFT-A. In comparison to the control (WT), mainly the mutant carrying the p.T181Nfs\*2 in IFT43 showed clear alterations in the stoichiometry of IFT-A components. In comparison to the control, this mutation is characterized by an increased portion of the IFT-A components IFT122, IFT140 and WDR19 in comparison to IFT43, TTC21B and WDR35. Although no difference in DNA sequence of the three tested single clones was identified, a slight alteration in composition of IFT-A can be shown.

To analyse the effect on IFT-A composition upon introduction of a mutant, four different single clones (WT, WDR19\_p.R1022Dfs\*21, IFT43\_p.Q195Afs\*22 and IFT43\_p.T181Nfs\*2) were selected and used to prepare four biological replicates per clone. Absolute quantification was



performed for four biological replicates per single clone and each was measured twice as technical replicate. Results are depicted in Figure 36.



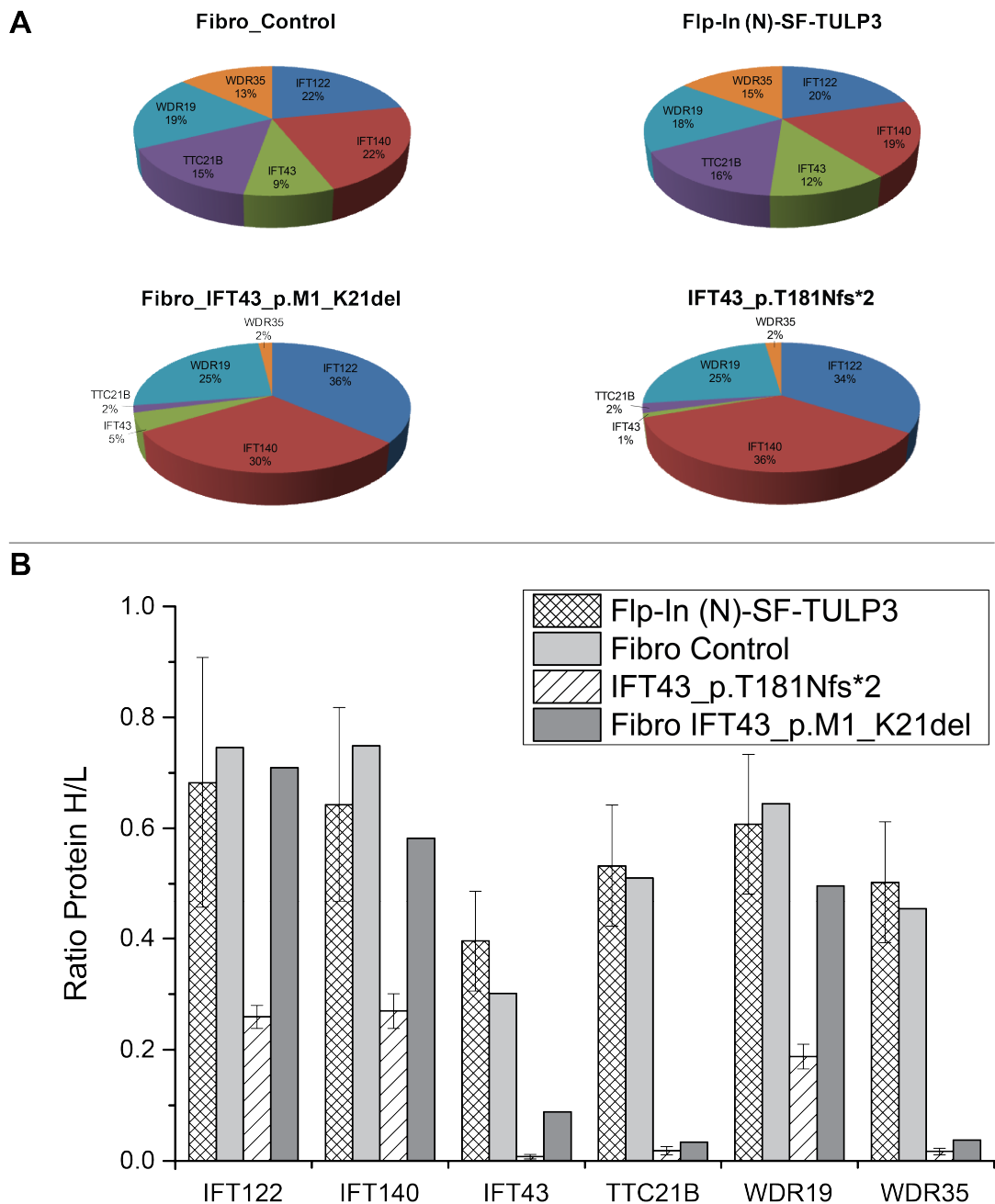
**Figure 36: Effect of selected single clones on complex composition of purified IFT-A**

A strong effect on IFT-A complex stoichiometry was observed in the Flp-In (N)-SF-TULP3 carrying the p.T181Nfs\*2 mutation in IFT43. Cells from clones carrying the p.Q195Afs\*22 mutation in IFT43 or the p.R1022Dfs\*21 mutation in WDR19 showed no alteration in IFT-A composition compared to WT cells. The determined ratio of labelled to unlabelled protein (determined by the representative standard peptides) is depicted in the bar chart. As depicted, protein expression level of the IFT-A components of cells carrying the IFT43\_p.Q195Afs\*22 mutation is similar to the expression level in the control cells. For cells with WDR19\_p.R1022Dfs\*21 the expression level is reduced while the composition of the IFT-A stays the same. In contrast, the overall expression level of IFT-A components in Flp-In (N)-SF-TULP3 cells carrying the IFT43\_p.T181Nfs\*2 mutation is reduced while the composition is characterized by a loss of IFT43, TTC21B and WDR35.

Regarding the bar chart, protein expression level of analysed single clones is depicted. Whereas the expression level of all IFT-A components of Flp-In (N)-SF-TULP3 cells carrying the IFT43\_p.Q195Afs\*22 mutation showed no significant alteration in comparison to the control cells, the expression level of cells either with WDR19\_p.R1022Dfs\*21 mutation or IFT43\_p.T181Nfs\*2 was obviously reduced. While composition of IFT-A in Flp-In (N)-SF-TULP3 with WDR19\_p.R1022Dfs\*21 stayed unaltered, stoichiometry of IFT-A in cells carrying the c.541\_542insA mutation that causes a premature stop in IFT43 (p.T181Nfs\*2) was changed.

#### **4.5.3 IFT-A complex composition within fibroblasts**

This approach was adapted to IFT-A purified from human fibroblasts. Therefore, human fibroblasts of either control persons or of patients carrying mutations related to the ciliopathy Sensenbrenner Syndrome were transfected with TULP3 to pull out the IFT-A complex by one-step affinity purification using anti-FLAG-M2-agarose beads. The cultivation of human fibroblasts as well as the one-step affinity purification was done by Machteld Oud, a PhD student at the Radboud university medical center Department of Human Genetics in Nijmegen (Netherlands). Afterwards, eluates of one-step affinity purification were sent to our lab for protein precipitation and enzymatic cleavage followed by absolute quantification using the equimolar standard mix in combination with PRM (see 3.2.5.1). First results of IFT-A composition of human fibroblasts are highlighted in Figure 37. In this graph, results of IFT-A composition obtained from human fibroblasts are face to face with results obtained with Flp-In (N)-SF-TULP3 monoclonal cells. As depicted in Figure 37A, although only one biological replicate for the human fibroblasts was available for absolute quantification so far, the composition of IFT-A in healthy control fibroblasts was identical to the composition within Flp-In (N)-SF-TULP3 cells. Results for IFT-A composition within human fibroblasts carrying a homozygous mutation in the gene encoding IFT43 (c.1A>G/p.M1\_K21del) were comparable to the composition of Flp-In (N)-SF-TULP3 carrying the IFT43\_p.T181Nfs\*2 mutation. In both cases the portion of IFT43, TTC21B and WDR35 was reduced in contrast to the amount of IFT122, IFT140 and WDR19. Regarding protein expression levels of the IFT-A components, highlighted in Figure 37B, the expression level of the healthy fibroblasts was comparable to the expression level of all components within Flp-In (N)-SF-TULP3 cells. Although the expression level of IFT122, IFT140 and WDR19 within human fibroblasts carrying the homozygous mutation were similar to the control cells, the protein expression level of IFT43, TTC21B and WDR35 was significantly reduced. In contrast, the expression level of the IFT-A components within Flp-In (N)-SF-TULP3 control line with IFT43\_p.T181Nfs\*2 was lower than for the Sensenbrenner Syndrome-related mutation in human fibroblasts.



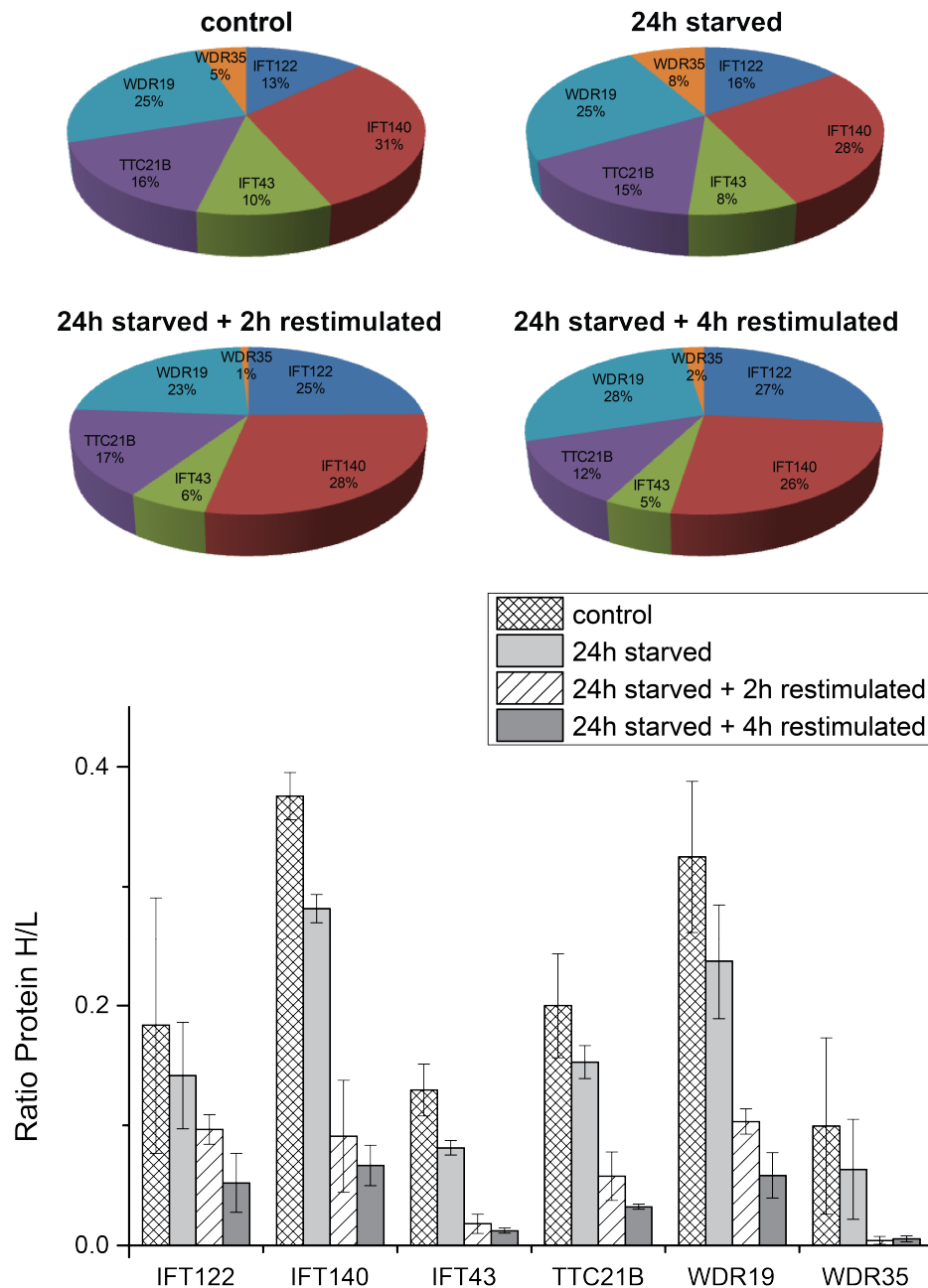
**Figure 37: IFT-A complex composition within human fibroblasts**

To compare complex composition of IFT-A purified from human fibroblasts with IFT-A stoichiometry from Flp-In (N)-SF-TULP3 cells, results of absolute quantification of control cell lines as well as a cell lines carrying mutations within IFT43 are depicted. A: Complex composition of IFT-A in control cell lines was identical for both, human fibroblasts as well as Flp-In cells, although only one biological replicate was available for absolute quantification. Comparing both cell lines carrying mutations within IFT43, the complex composition was virtually identical. B: Regarding both control lines, the protein expression level was similar. For the human fibroblasts carrying the homozygous mutation, expression level of three components (IFT122, IFT140 and WDR19) was comparable to the one obtained in the control lines, while the expression level of IFT43, TTC21B and WDR35 was reduced. In contrast, the protein expression level of all IFT-A components within Flp-In cells carrying the IFT43\_p.T181Nfs\*2 mutation was reduced.

#### 4.5.4 Complex composition of IFT-A during ciliary assembly and disassembly

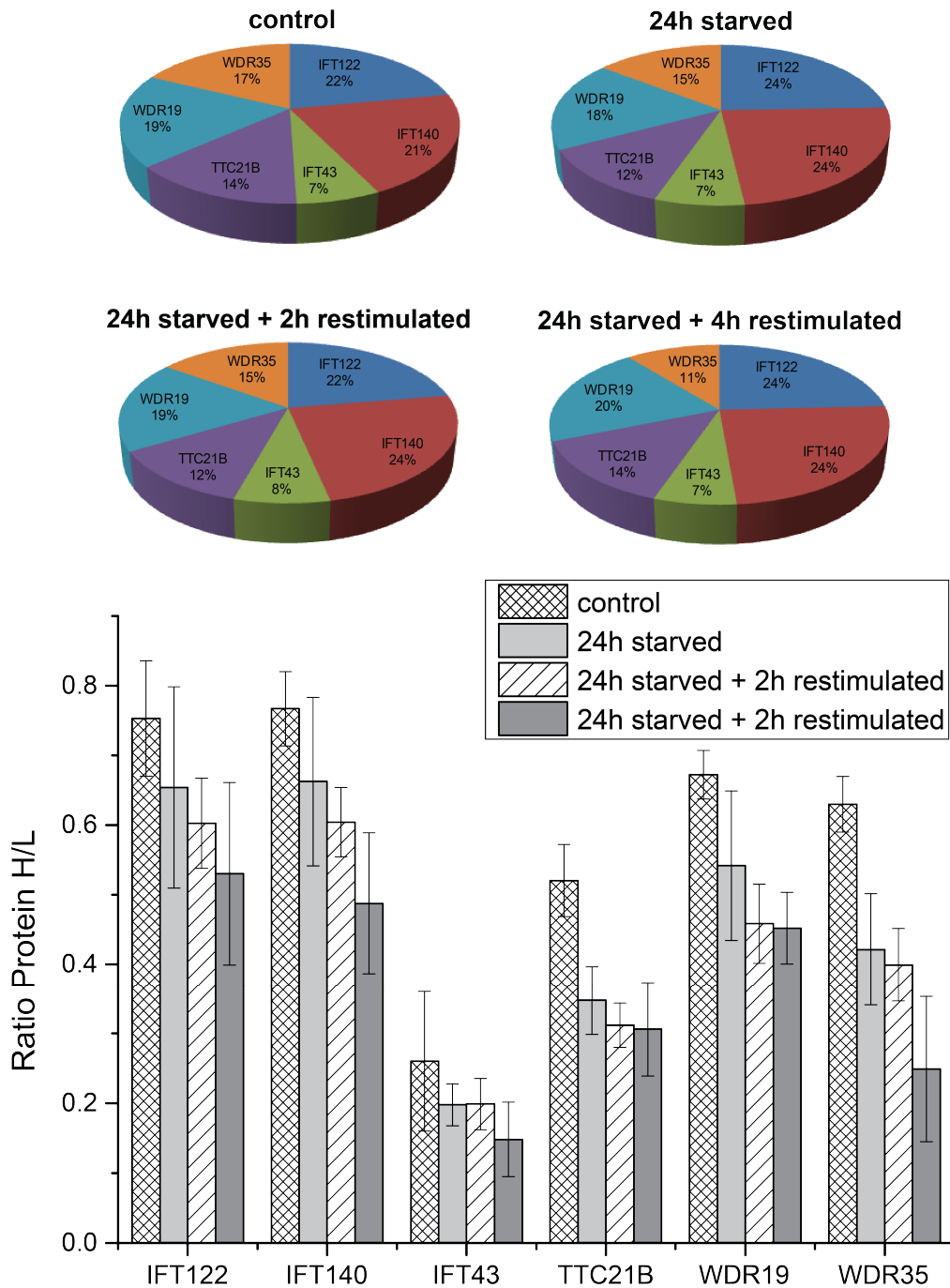
A biological process which might lead to alterations in the stoichiometry of IFT-A is the assembly and disassembly of cilia. As described before (see 1.1.2), IFT-A is involved in retrograde intraflagellar transport (IFT) returning protein cargo from the ciliary tip back to the cell body. To promote different ciliary stages, hTERT-RPE1 cells were cultured in various conditions before absolute quantification of purified IFT-A was performed. To support the assembly of cilia, cells were starved whereas restimulation of starved cells results in the disassembly of cilia.

To compare the alterations within IFT-A complex composition during different ciliary stages, composition of purified IFT-A from hTERT-RPE1 cells, starved for 24h, were compared to the stoichiometry of IFT-A within cells starved for 24h with following restimulation period of either 2h or 4h. As a control, purified IFT-A from hTERT-RPE1 cells, cultured in growth medium, were analysed (see Figure 38). This first experiment unveiled alterations within IFT-A composition during the assembly and disassembly of cilia. As expected, IFT-A composition purified from hTERT-RPE1 cells starved for 24h with following restimulation period showed big differences in complex stoichiometry in contrast to control cells as well as to the cells that were starved for 24h without restimulation. Additionally, control cells showed similar results for IFT-A composition like hTERT-RPE1 with induced assembly of the cilia. Regarding the bar chart, protein expression level of the IFT-A components in control cells as well as in conditions forcing the assembly of cilia is much higher than for IFT-A during ciliary disassembly. Cells, representing the control fraction were seeded one day later than cells of different starvation and restimulation conditions. The similarity of IFT-A composition of the control cells to the composition within cells starved for 24 hours hypothesizes a potential influence of cell confluence on the complex composition of purified IFT-A. Therefore, another experiment investigating complex composition of IFT-A considering higher cell confluence was performed. Different IFT-A stoichiometries considering cell confluence are depicted in Figure 39. Again same growth conditions were tested as in the previous experiment to simulate ciliary assembly as well as the disassembly of cilia. To simulate higher cell confluence, more cells were seeded per 14cm culture dish. As highlighted in Figure 39, no alteration within the complex composition of purified IFT-A was identified. The composition of IFT-A within hTERT-RPE1 cells of high cell confluence was comparable to the stoichiometry of the ciliary assembly of the previous experiment. These results verify the hypothesis that the phenotype of ciliary assembly or disassembly of hTERT-RPE1 cells is depending on cell confluence.



**Figure 38: Complex composition of IFT-A during ciliary assembly and disassembly**

To investigate stoichiometric alterations of IFT-A during different ciliary stages, hTERT-RPE1 cells of different cultivation conditions were analysed by absolute quantification. Control cells were cultured in growth medium, while other ciliary stages were simulated by 24h starvation, 24h starvation followed by 2h of restimulation or 24h starvation followed by 4h of restimulation. Cells, starved for 24h, unveiled different composition of IFT-A conditions in comparison to hTERT-RPE1 cells starved for 24h with following restimulation. Unexpectedly, control cells showed similar complex composition as cells starved for 24h. This may be caused by a high confluence of cells in the control conditions, simulating a similar condition as for starved cells. Regarding the bar charts, protein expression levels of IFT-A components in conditions forcing ciliary disassembly were decreased in comparison to the control cells as well as in the stage of ciliary assembly. The composition of the control cells hypothesizes a potential influence of cell confluence on complex stoichiometry of IFT-A.



**Figure 39: Confluence-related complex composition of IFT-A**

To study the influence of cell confluence on IFT-A complex composition, hTERT-RPE1 cells were cultured under different conditions while more cells were seeded on each culture dish to simulate higher cell confluence. As depicted, no matter which cell condition was analysed by absolute quantification, the complex composition of IFT-A as well as the protein expression level of all tested conditions were comparable. This verifies the hypothesis that the assembly or disassembly phenotype of hTERT-RPE1 cells is depending on cell confluence.

## 4.6 Chemical crosslinking of purified IFT-A

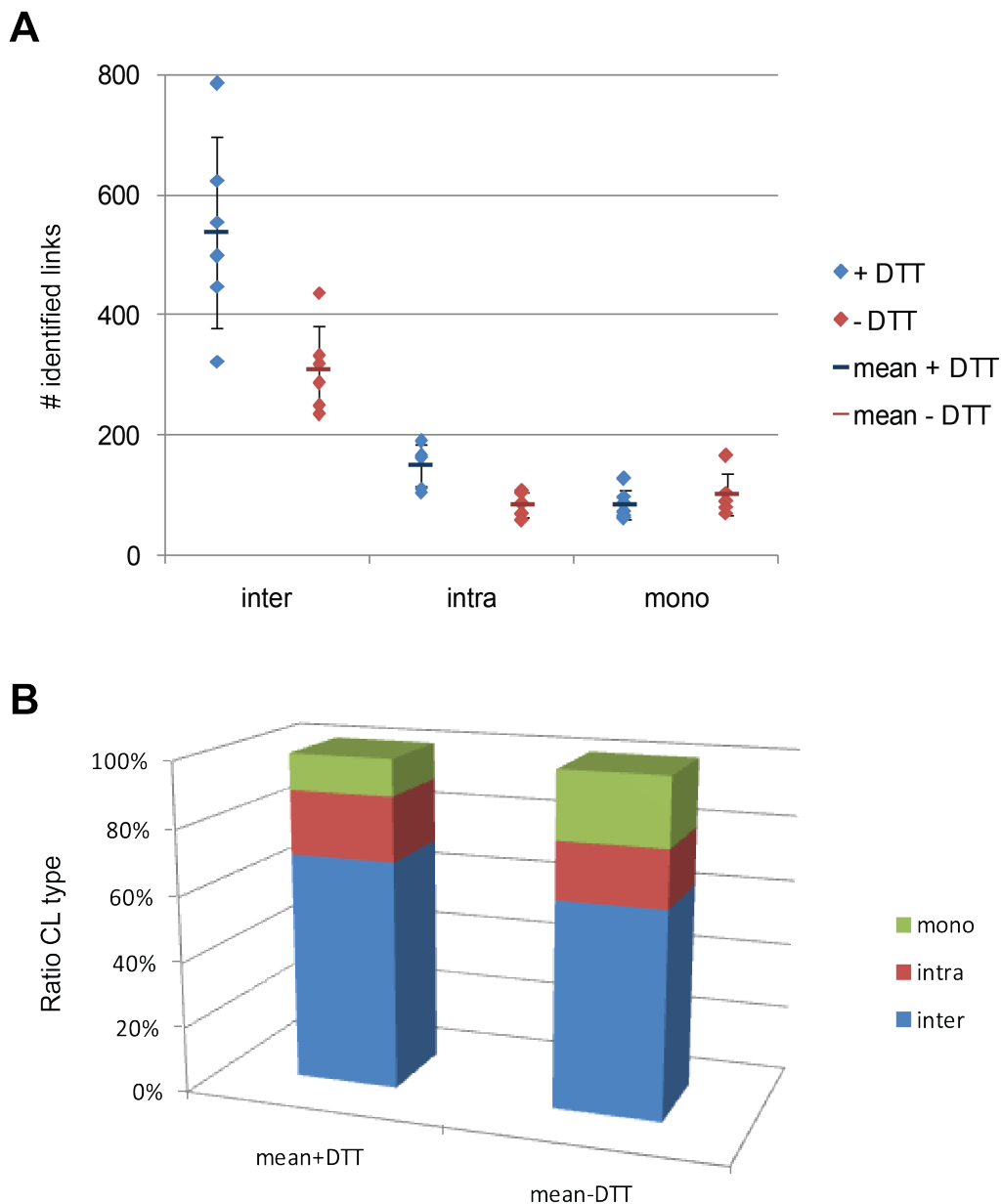
To determine the binding surfaces within the IFT-A and to understand its function, it is necessary to gain structural knowledge. For the structural analysis of IFT-A, a huge amount of the purified protein complex is crucial. Since yields of purified IFT-A were consistently higher using Flp-In (N)-SF-TULP3 cells, this cell line was chosen to investigate the structural interaction of IFT-A components by means of chemical crosslinking combined with mass spectrometry. After cells were harvested, tandem affinity purification (TAP) of IFT-A was performed to purify IFT-A as pure as possible with high and medium affinity interactors only [69, 81]. After chemical crosslinking and proteolysis of purified proteins, cross-linked peptides were enriched before LC-MS/MS analysis to reduce sample complexity and to increase the likelihood of identifying true cross-linked peptides (see 3.2.5.2). For an efficient crosslinking of primary amines of N-termini and lysine residues of purified IFT-A, reaction conditions like the ratio of crosslinker to purified protein and the composition of the reaction buffer had to be chosen carefully.

### 4.6.1 Optimization of the crosslinking workflow

The workflow of a chemical crosslinking approach contains a number of critical parameters like the amount of purified protein, the selection of a suitable crosslinker, the ratio of chemical crosslinker to protein amount, the reaction time and reaction temperature, buffer conditions, the enrichment of cross-linked peptides as well as a suitable method for LC-MS/MS analysis.

#### 4.6.1.1 Presence of dithiotreitol (DTT) during chemical crosslinking

The presence of dithiotreitol (DTT) in reaction buffers stabilizes the native conformation of proteins by preventing cysteine and methionine from oxidation damage [82, 83]. Nevertheless, a drawback of DTT is that it reduces disulfide bridges formed by cysteine residues. Thereby, the protein structure can be severely altered because the tertiary structure is generally stabilized by disulfide bridges [84] and released thiol groups are also accessible to chemical crosslinking with succinimidyl esters [85]. However, this may complicate the further data analysis. To explore whether the use of DTT improves the data quality, crosslinking either with or without DTT in the reaction buffer was performed and identified crosslinks were compared as depicted in Figure 40. Regarding the average amount of identified crosslinks, the number of inter- and intra-links was higher using DTT in the reaction buffer. However, the variance between the biological replicates was significantly higher using DTT, except for identified mono-links. Using DTT during chemical crosslinking, the biggest portion of identified links was represented by inter-links, while mono-links were least abundant. In comparison, without DTT in the reaction buffer, inter-links represented the major portion while intra-links were least abundant.



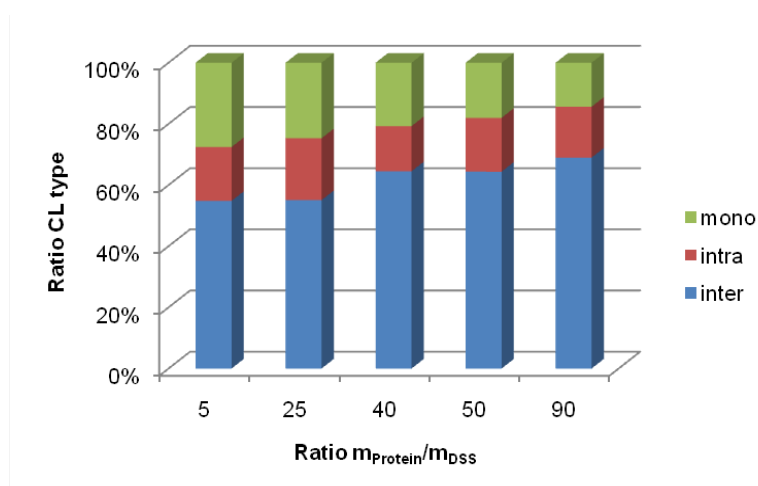
**Figure 40: Number of identified crosslinks with or without DTT**

To compare the influence of dithiotreitol (DTT) on chemical crosslinking, purified IFT-A was cross-linked using a reaction buffer either with or without 2mM DTT. A: The data of six biological replicates are shown by a rhombus either in blue: with DTT or in red: without DTT. The identified amount of crosslinks between proteins (inter) or within proteins (intra) or single peptides attached to the crosslinker (mono) is used to compare the effect of DTT on chemical crosslinking. Mean and standard deviation (SD) of the biological replicates are depicted using a horizontal line. The average amount of inter- and intra- links was higher using DTT in the reaction buffer, although the variance between the biological replicates was much higher using DTT, except for identified mono links. B: To identify differences between both experiments, the fraction of identified crosslink types is depicted. Using DTT during chemical crosslinking, the major fraction of identified links were inter-links while least abundant links were mono-links. In comparison, without DTT in the reaction buffer, inter-links represented the largest portion, while intra-links were least abundant.



#### 4.6.1.2 Concentration of the chemical crosslinker DSS

The ratio of the chemical crosslinker to the protein amount as well as its concentration are critical parameters for a successful crosslinking approach. If the amount of chemical crosslinker is insufficient, the number of available chemically crosslinked peptides may be too low for mass spectrometry analysis. Furthermore, not every accessible lysine residue may be cross-linked resulting in an uneven distribution of the cross-linked residues all over the protein complex. In contrast, an excess of chemical crosslinker may result in a large fraction of mono-linked peptides. As described in previous studies [85, 86], chemical crosslinkers based on succinimidyl ester enable crosslinks between primary amines (N-termini, lysine residues) as well as with other amino acids like cysteine and tyrosine, although its reactivity with primary amines is higher. Certain conditions, like an excess of chemical crosslinker, probably promotes this non-amine reactivity and thereby enables crosslinking of other residues as well. Furthermore, an excess of crosslinker may impede the analysis of low abundant inter- and intra-links due to an overrepresentation of mono-linked peptides that will not react with a second residue because this is already occupied with another crosslinker molecule. To test different concentrations of chemical crosslinker, five different ratios of protein amount to disuccinimidyl suberate (DSS) (5-, 25-, 40-, 50-, 90-fold) were tested. Evaluation was done comparing the fractions of inter-, intra- and mono-links. Results are shown in Figure 41.



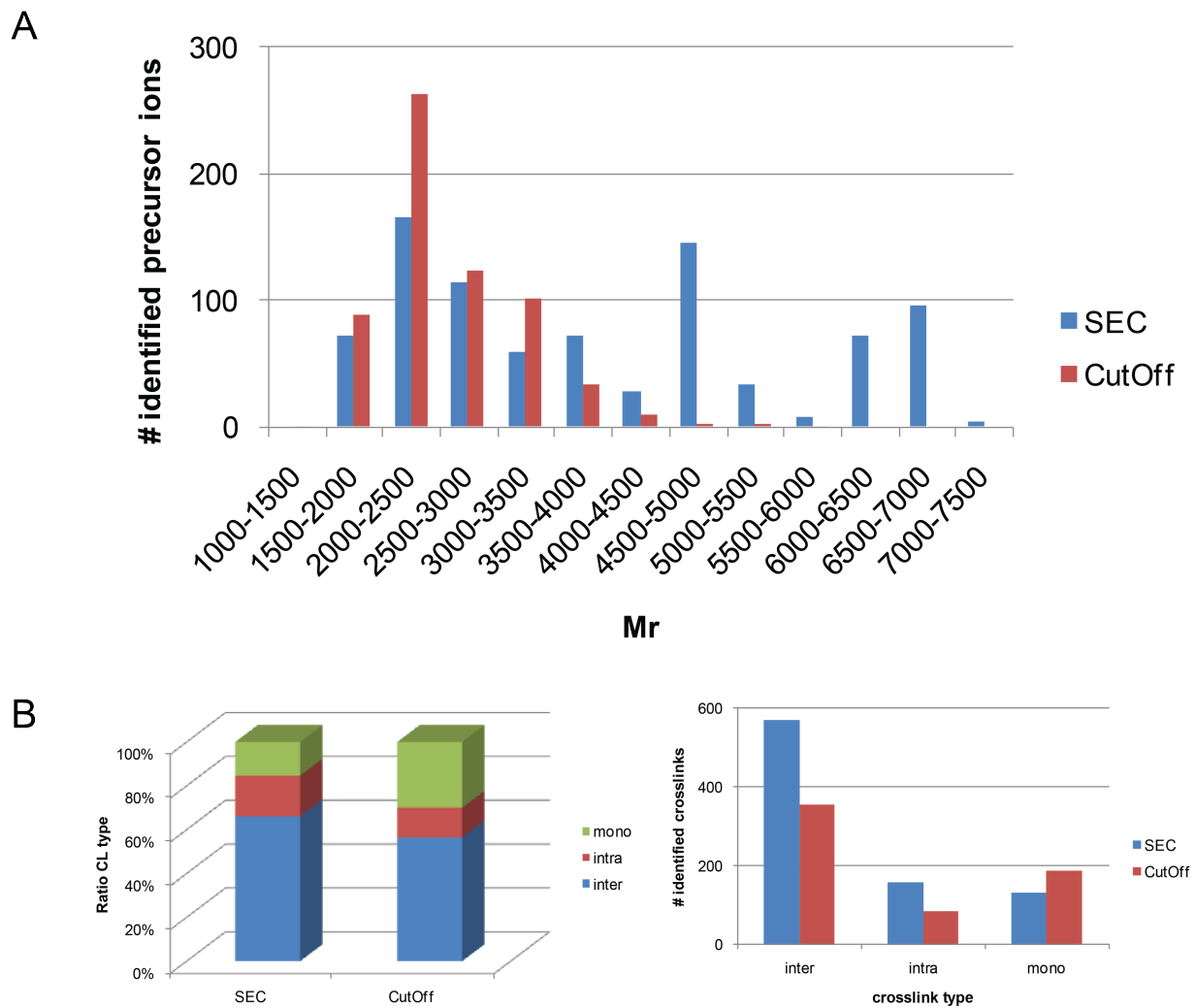
**Figure 41: Optimization of DSS concentration for chemical crosslinking**

Five ratios of protein amount of purified IFT-A to the crosslinker DSS were tested to identify best reaction conditions. As expected, higher amounts of chemical crosslinker increased the fraction of mono-links (depicted in green) and reduced the portion of identified inter-links (blue). In contrast, decreasing the amount of DSS led to an increase in the fraction of identified inter-links. Considering the three experiments with 40-, 50- and 90-fold excess of protein in relation to applied crosslinker, no dramatic increase in intra- and inter-links was achieved. A 50-fold excess of protein is a good ratio because there is enough buffer for technical errors. Therefore, this condition (50-fold protein to crosslinker) was chosen for further structural investigation of IFT-A.

The least amount of identified inter-links (between two different peptides of different proteins) was obtained using a 5-fold or 25-fold excess of protein to chemical crosslinker. For this ratio of protein to crosslinker, highest amount of mono-links was obtained. With decreasing amount of chemical crosslinker the fraction of identified inter-links raised. Considering the three experiments with a 40, 50 and 90-fold excess of protein to DSS, no further dramatic increase in inter- and intra-links was obtained. Based on these results, a 50-fold excess of protein amount in relation to the chemical crosslinker DSS was used for further structural investigation of IFT-A because this ratio offers enough buffer for technical errors.

#### **4.6.1.3 Different methods to enrich cross-linked peptides**

For LC-MS/MS analysis of low abundant peptides, an enrichment step is required. To enrich low abundant cross-linked peptides within a complex mixture of inter-, intra-, mono- and non-linked peptides, size exclusion chromatography (SEC) is a well-established method. This method is based on separation of the peptide mixture by liquid chromatography. Thereby, small molecules are restrained in the pores of the column material, while bigger molecules, for example cross-linked peptides, are able to pass. Thereby, cross-linked peptides are eluting in the first fractions which can then be used for further LC-MS/MS analysis. Although SEC is well described and established for this application, this approach is labour- and time-consuming as well as expensive. Therefore, I tested whether filtration through a 3kDa CutOff spin column is capable to replace SEC for this application. The principle is similar: larger molecules are restrained in the upper part of the filter while small molecules pass through the filter. The results for the comparison of both methods is shown in Figure 42. The amount of each identified crosslink type achieved with SEC or using 3kDa CutOff spin columns is illustrated. For this experiment, Flp-In (N)-SF-TULP3 cells were used. After purification of IFT-A with following chemical crosslinking, cross-linked peptides were either enriched by SEC or 3kDa CutOff spin column filtration to reduce sample complexity. As depicted in Figure 42A, the distribution of identified precursor ions in lower mass ranges (1500-4500) was similar, whereas precursors with higher molecular masses (>4500) were not identified after 3kDa CutOff column filtration. As highlighted in Figure 42B, the proportion of identified crosslink types differed between both methods. After SEC, mono-linked peptides represented the least abundant portion. In contrast, after using CutOff spin columns, the fraction of intra molecular crosslinks was least abundant. Regarding the number of identified crosslinks, SEC of purified samples yielded in significantly more inter- and intra-links than using spin columns.



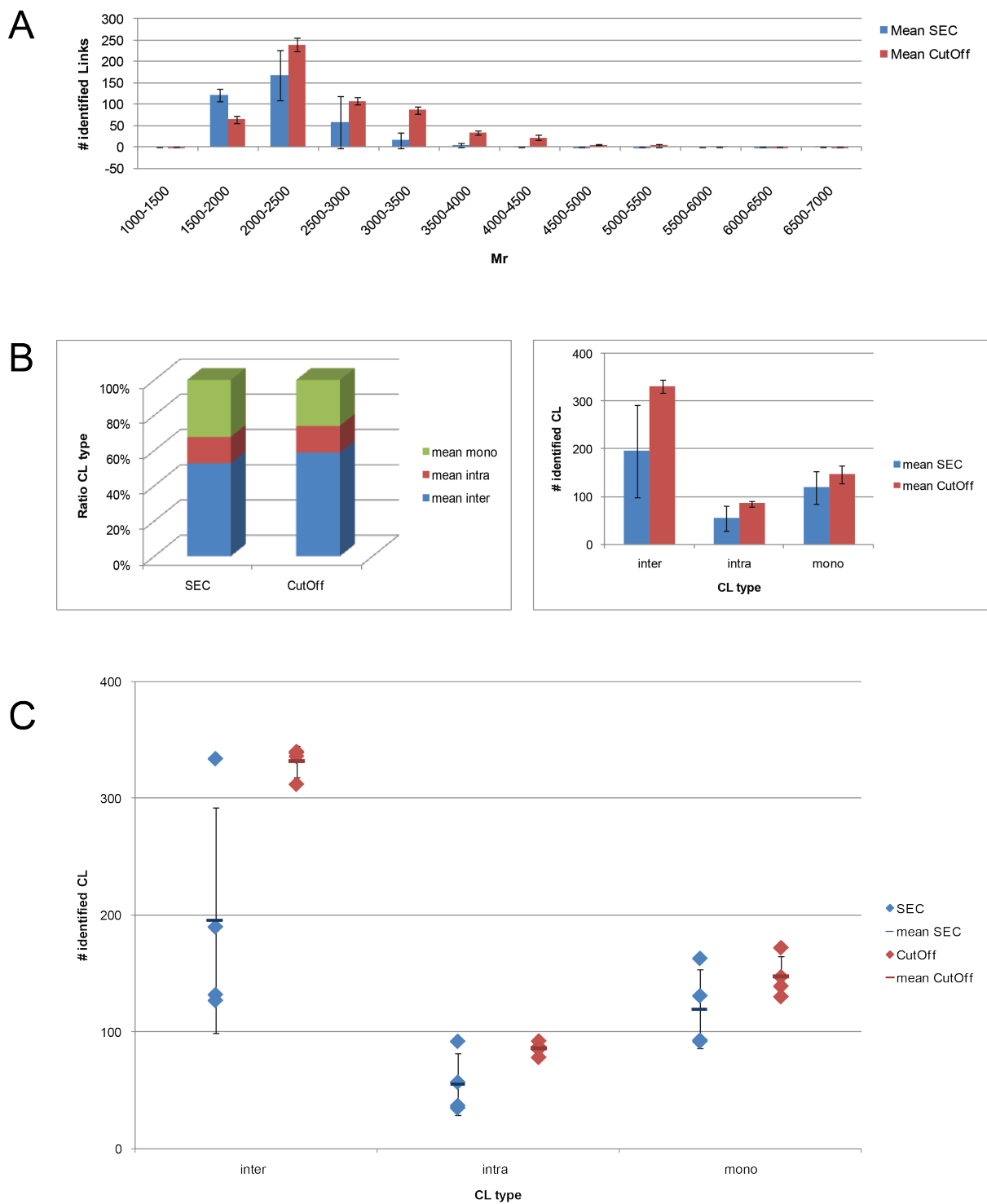
**Figure 42: SEC or CutOff spin column filtration for the purification of cross-linked peptides**

Direct comparison of SEC and 3kDa CutOff spin column filtration for the purification of cross-linked peptide mixture. IFT-A, purified from Flp-In (N)-SF-TULP3 was used for chemical crosslinking. The sample was prefractionated either by SEC or using spin columns followed by LC-MS/MS analysis. A: Comparing the precursor distribution, patterns of both methods were similar in the molecular mass range of 1500 to 4500. For higher  $M_r$ , precursor ions were only identified in the sample purified by SEC. B: The ratio of identified crosslink types differed between both tested methods. Performing SEC, the amount of mono-linked peptides (depicted in green) represented the smallest fraction. Using spin columns resulted in less intra-links (red). Also the overall amount of identified inter- and intra-links was notably higher performing SEC.

---

Another experiment with four biological replicates was performed to compare efficiency and robustness of both enrichment methods in more detail. In this experiment, HEK293T cells transiently transfected with TULP3 were used to generate biological replicates, because the amount of purified IFT-A is higher using overexpression. Again, purification and crosslinking was performed. Then, the cross-linked peptide mix was split to perform both enrichment methods separately. Detailed results are shown in Figure 43. Comparing the distribution of identified precursor ions performing SEC (highlighted in blue) or spin column filtration (depicted in red), the distributions were similar with most identified precursors the molecular mass range of 1500 to 2000 (see Figure 43A). However, standard deviation was much higher performing SEC. Figure 43B gives the average distribution (left) and amount (right) of identified links for both enrichment methods. While no difference within the proportions of crosslink types was apparent, the average number of identified crosslinks was higher using 3kDa CutOff spin columns. Furthermore, this simplified method led to less variations over the four replicates. The distribution of the data for the four replicates is depicted in Figure 43C. Each rhombus represents one biological replicate either performing SEC (blue) or spin column filtration (red). While the identified numbers of crosslinks within the biological replicates using spin columns were stable with only one outlier present, values generated performing SEC were highly dispersed.

The number of identified links within IFT-A, purified from HEK293T cells, was higher after spin column filtration in contrast to the higher number of identified links after SEC in the previous experiment using IFT-A purified from Flp-In (N)-SF-TULP cells (see Figure 42). These variations might be explained by a higher sensitivity of SEC. Thereby, 3kDa Cut Off spin columns are less sensitive with lower amounts of starting material (in experiments purifying IFT-A from Flp-In (N)-SF-TULP3 cells) than SEC. In contrast, with an increase in sample amount (in experiments using HEK293T cells for the purification of IFT-A) this sensitivity gain of SEC might be compensated. To compare both enrichment methods and to evaluate potential effects of the used method on the identification and localisation of crosslinks, identified links for all experiments were visualized using the software tool Xinet (see 4.6.2).



**Figure 43: SEC and spin column filtration of cross-linked peptide mixture from HEK293Ts**

To compare the efficiency of both enrichment methods, four additional biological replicates were generated using HEK293T cells, transiently transfected with (N)-SF-TULP3. After purification and crosslinking, sample was split and enrichment of cross-linked peptides was either performed by SEC or using 3kDa CutOff spin columns. A: Precursor distribution of both tested methods showed no significant difference. With both methods, precursors of a molecular mass within the range of 1500 to 2500 were most

abundant. Variations within analysed replicates were much higher performing SEC. B: The average number of identified crosslinks was higher using spin columns (red), although the distribution of crosslink types was highly similar. Additionally, this simplified method using spin columns led to less variation over the replicates. C: Reproducibility of both enrichment methods. The data points for all biological replicates are shown as rhombuses. The ones obtained after SEC are depicted in blue and the ones using spin columns are highlighted in red. The average number and the standard deviation of identified crosslinks are illustrated by horizontal lines either in blue or red. Identified crosslinks performing SEC were highly disperse, whereas results obtained using spin columns were much more consistent.

#### 4.6.2 Structural investigation of IFT-A using chemical crosslinking

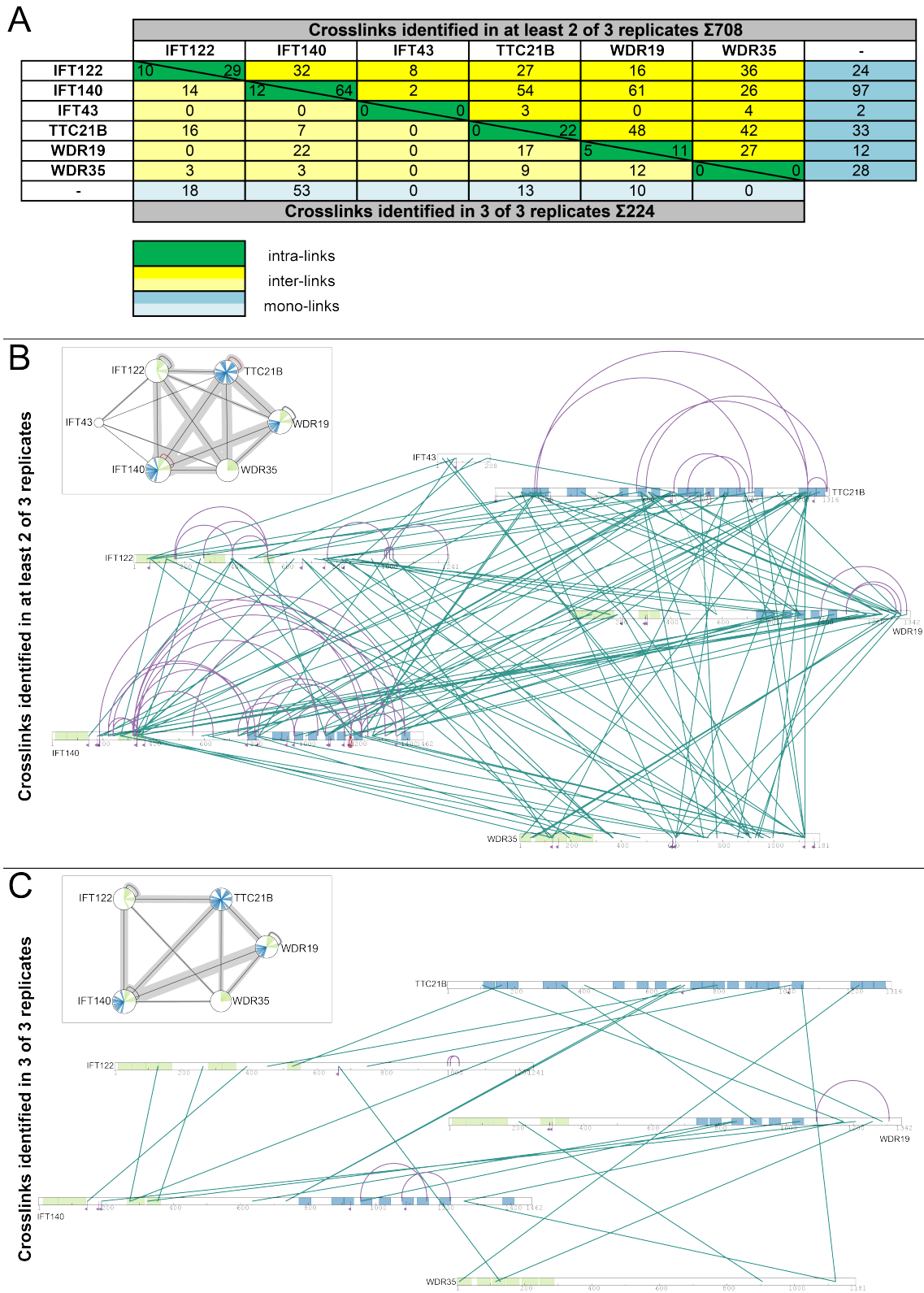
For the extraction of identified links within the biological sample, the Software pipeline xQuest/xProphet was used. The freely available software tool Xinet was applied for the visualization of identified crosslinks.

##### 4.6.2.1 Identified links within IFT-A after SEC

Identified crosslinks, obtained from Flp-In (N)-SF TULP3 cells using a 50-fold excess of protein to DSS, are depicted in Figure 44. In Figure 44A the number of crosslinks either identified in at least two of three biological replicates or in three of three biological replicates are listed. Total number of crosslinks, identified in at least three biological replicates, are visualized in Figure 44B, while crosslinks, identified in all tested biological replicates, are highlighted in Figure 44C. A total number of 708 crosslinks was identified considering only links that were present in at least two of the three tested samples. Thereby, more than 50 percent (54.5%) of the identified crosslinks were inter-links (386 inter-links) followed by 27.7% mono-links (196 mono-links) and 17.8% of intra-links (126 intra-links). In contrast, 224 crosslinks were identified in all three tested biological replicates. Again most portion of identified crosslinks were represented by 103 identified inter-links (46.0%) followed by 42.0% of mono-links (94 mono-links) and 12.1% intra-links (27 identified intra-links). As depicted in Figure 44B, considering crosslinks identified in at least two of the three tested samples, crosslinks within all of the six protein components of IFT-A were identified. The smaller graph on top represents a scheme of the identified crosslinks by correlating the thickness of the lines to the number of identified inter-links. For example in the approach considering only links identified in two of three experiments, only crosslinks between IFT43 and four other components (IFT122, TTC21B, IFT140 and WDR35) were identified and were least abundant (thin lines) in comparison to high abundant links between IFT140 and TTC21B (thick lines). The lower graphs of Figure 44B/C illustrate the identified crosslinks in more detail. This closer look into the localisation of identified links unveiled that most identified crosslinks between proteins and within a protein are present within described interaction domains, like TPR domains (light blue) and WD40 domains (light green). For example, many of the identified inter-links between WDR19 and IFT140 were identified between the TPR regions

between amino acid position 800 and 950 of WDR19 with the WD40 domains between amino acid position 250 and 370 of IFT140 or with the TPR domains localised around amino acid position 1200 of IFT140. As depicted, many links were identified on spots which are not described so far. This data indicates additional interaction domains which are not described and characterized, yet. Similar results were achieved by a more stringent approach, depicting links that were identified in three of three analysed biological replicates, shown in Figure 44C. However, no links within IFT43 were identified considering only crosslinks identified in all three tested replicates. The overall number of identified links was reduced significantly. Nevertheless, many of the crosslinks identified in all three experiments were located within described interaction domains and again not yet described crosslinking “hotspots” were identified, further on.

For detailed information about amino acid position and peptide sequences of identified crosslinks, see Table 19 (for links identified in at least two of three replicates ) and Table 20 ( for crosslinks identified in all tested experiments) in the annex.



**Figure 44: Identified links within IFT-A from Flp-In (N)-SF TULP3 after SEC**

To visualize the extracted information of identified crosslinks, Xinet was used to depict links of Flp-In (N)-SF TULP3 cells cross-linked with DSS in the ratio 50:1 (protein:crosslinker). A: Links either identified in at least two of three or in all three tested



samples are listed. Thereby, the overall number of links identified in at least two of three replicates was 708 in contrast to 224 links identified in all experiments. For both analysis, the portion of inter-links was higher than for intra- as well as for mono-links. The generated model on top of figure B and C illustrates the amount of identified links schematically, while the graph below visualizes the localization of the identified links in detail. Thereby, inter-links between two different IFT-A components are depicted by dark green lines, intra-links are highlighted by purple semicircles, loop-links identified within the same peptide are depicted as red loops and mono-links are highlighted by purple flags. Known and described interaction domains within IFT-A components are highlighted on the protein representation in light green (WD40 domains) and light blue (TPR domains). B: Links identified in at least two of three biological replicates were depicted. Many of the identified links were localized within known interaction domains, for example the identified inter-links between TPR regions (amino acid position 800 to 950) of WDR19 with WD40 domains (amino acid position 250 to 370) or with TPR regions (around amino acid position 1200) of IFT140. Additionally, many links were localized in regions which are not described so far, indicating additional interaction domains (e.g. at amino acid position 600 or 1120 of WDR35). C: Only links identified in all three biological replicates were highlighted. Thereby, the overall amount of identified links was reduced in contrast to links identified in at least two of three experiments. For example no link between IFT43 and any other component of IFT-A was identified in all three experiments. Nevertheless, some links between IFT-A components were quite robust and were identified in all tested samples. Again, some of the robust crosslinks were identified in described interacting domains like in TPR domains of WDR19, whereas some crosslinks were identified in undescribed regions (e.g. around amino acid position 1150 of WDR19) indicating new interaction domains.

#### 4.6.2.2 Identified links within IFT-A using different enrichment methods

As described in 3.2.5, two different methods were used to enrich chemically cross-linked peptides. This fractionation is essential for the identification of low abundant cross-linked peptides via mass spectrometry. To facilitate the enrichment of chemically cross-linked peptides, 3kDa CutOff spin columns were used to reduce sample complexity. This approach was compared to the well described but time-, labour- and cost-intense size exclusion chromatography (SEC). Therefore, a chemically cross-linked sample (as described in 3.2.5.2) was split after proteolysis and either SEC or spin column filtration was performed for fractionation. After MS rawfiles were evaluated by xQuest/xProphet, identified links were visualized using Xinet.

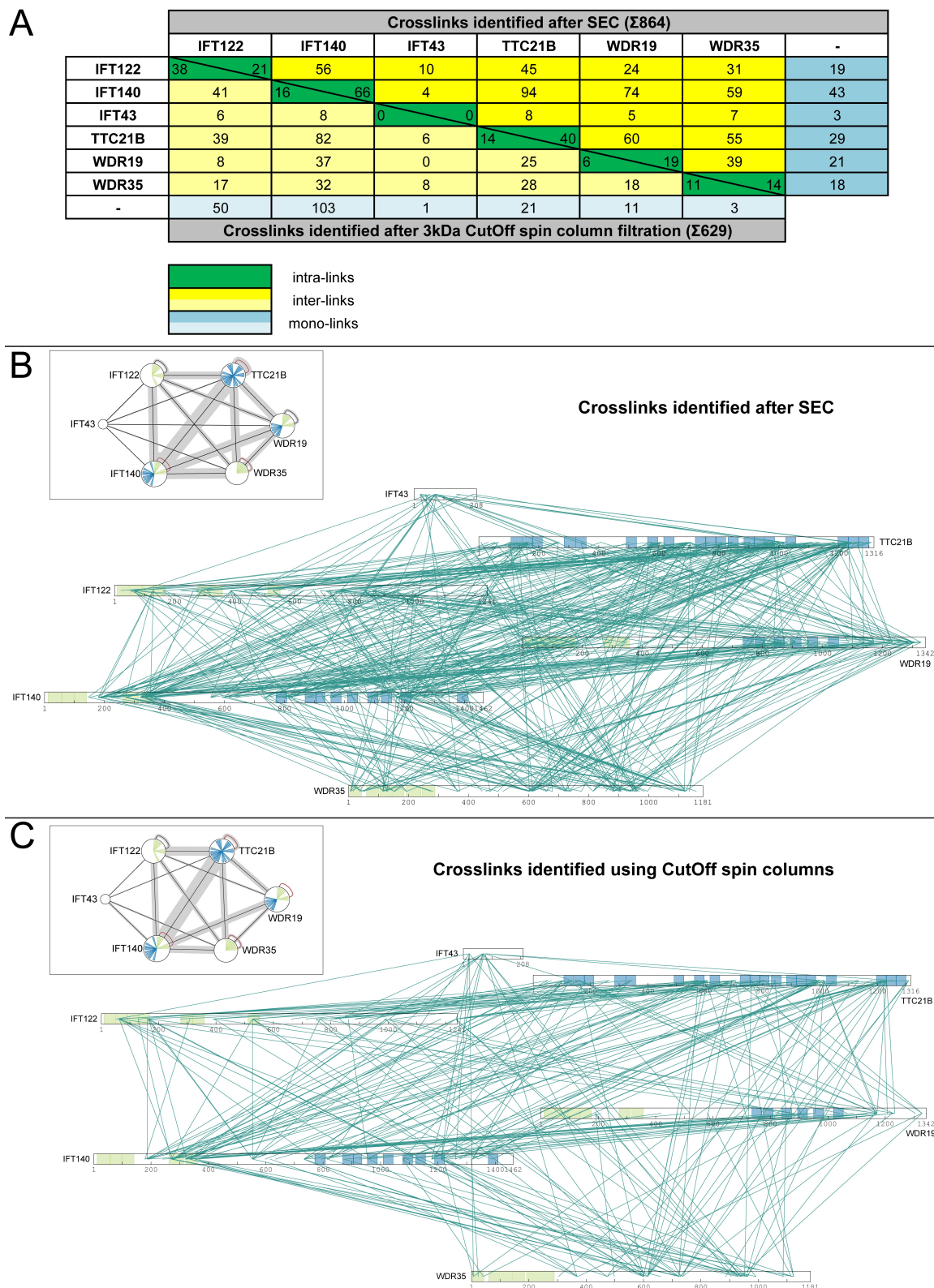
#### Chemical crosslinking of IFT-A purified from Flp-In (N)-SF-TULP3 cells

Results are shown in Figure 45. Total number of identified links of one biological sample obtained either after SEC or 3kDa CutOff spin column filtration are listed in Figure 45A. The total number of links identified after SEC (864 links) was significantly higher than after CutOff column filtration (629 links), although the portion of identified inter-links was higher for both approaches (66.1% for SEC and 56.4% for CutOff spin columns) than for the other link types. Performing SEC resulted in 18.5% intra-links (160 intra-links) followed by 15.4% mono-linked peptides (133 mono-links). In contrast, only 13.5% of identified links using spin columns were intra-links (85 intra-links) while 30.0% were represented by mono-links (189 mono-linked

---

peptides). Localisation of identified links obtained after SEC are depicted in Figure 45B, while position of identified links obtained using spin columns are highlighted in Figure 45C. Considering the schematic models of the identified links (upper graph in figure B and C), links within all six complex components were identified using both enrichment methods. The abundance of inter-links (depicted by the thickness of the lines) was also similar for both methods. Thereby, inter-links between IFT140 and TTC21B were most prevalent while inter-links of IFT43 with the other IFT-A components were least abundant. Using 3kDa CutOff spin column filtration to enrich cross-linked peptides, no inter-links between IFT43 and WDR19 were identified in contrast to SEC. The localization of the identified links was similar for both enrichment methods although the overall number of identified links was higher performing SEC (see Figure 45B/C lower graph). Furthermore, localization of identified links was comparable to previous results highlighting described interaction domains as well as identifying new interaction regions indicated by crosslink “hot-spots” (see Figure 44).

Amino acid position and peptide sequence of the links, identified in both enrichment methods, are shown in the annex (see Table 21).



**Figure 45: Identified inks of IFT-A (Flp-In (N)-SF TULP3) after SEC or spin column filtration**

Results of two different methods for the enrichment of cross-linked peptides are depicted. A: Numbers of links identified either after SEC or after 3kDa CutOff spin column filtration are listed. The overall amount of identified links was higher after SEC (864 links) than

after spin column filtration (629 links). However, the major portion of identified links was represented by inter-links (66.1% after SEC and 56.4% after spin column filtration) for both enrichment methods. B: Visualization of identified links obtained after SEC. C: Visualization of identified links obtained after spin column filtration. Both enrichment methods identified links within all IFT-A complex components, although no inter-link between IFT43 and WDR19 was identified using spin columns. However, the abundances of identified links were similar for both approaches (corresponding to the thickness of the lines in the upper graph). Although the overall amount of identified inter-links was higher after SEC, the localization patterns within the proteins were quite similar for both methods. For example the crosslinking hot-spots all over WDR35 were identified no matter which enrichment method was used.

### **Chemical crosslinking of IFT-A purified from HEK293T**

For the validation of the facilitated enrichment approach using 3kDa CutOff spin columns, biological replicates are crucial. To handle this increase in sample material, Hek293T cells, transiently transfected with (N)-SF-TULP3, were used for the generation of the biological replicates. Samples were generated as described previously in 3.2.5.2 and split before cross-linked peptides were either enriched performing SEC or 3kDa CutOff spin column filtration. Results of links identified in at least two of four IFT-A samples from HEK293T cells are depicted in Figure 46, while links identified in at least three of the four tested replicates are highlighted in Figure 47. For detailed information of crosslinked amino acids and positions of identified links after SEC, see Table 22 (links identified in at least two of four biological replicates) and Table 23 (links identified in at least three of four replicates). Information of identified links after spin column filtration is depicted in Table 24 (identified links within at least two of four biological replicates) and

---

Table 25 (links identified in at least three of four experiments) in the annex.

Considering the number of links identified in at least two of four tested biological replicates from HEK293T cells (Figure 46A), the amount of identified links was significantly higher after spin column filtration (1144 links after spin column filtration and 708 links identified after SEC). After SEC, major portion of link types was represented by mono-linked peptides (47.7% with 338 identified mono-links) followed by 281 inter-links (39.7%) and 89 intra-linked peptides (12.6%). In contrast, major portion after spin columns filtration was represented by 45.6% inter-links (522 identified inter-links) followed by 42.0% mono-linked peptides (481 mono-links) and 12.3% intra-links (141 intra-linked peptides). Although the distribution of crosslink types after both enrichment methods varied, crosslinks within all IFT-A components were identified after both enrichment methods (results after SEC are depicted in Figure 46B and results after spin column filtration are highlighted in Figure 46C). Although the overall amount of links identified in at least two of three replicates of HEK293T cells was higher after 3kDa CutOff spin column filtration, patterns of crosslink localisations after both enrichment methods were comparable. However, some links were only identified either after SEC or after the use of 3kDa CutOff spin columns. For example the inter-link between TTC21B (at amino acid position 1304) and WDR35 (at amino acid position 1119) was only identified after SEC; while the inter-link between TTC21B (at amino acid position 1221) and IFT43 (at amino acid position 21) was only identified after spin column filtration. Again, same crosslinking “hotspots” were detected as in previous experiments using IFT-A purified from Flp-In (N)-SF-TULP3 cells.

The number of links, identified in at least three of four biological replicates, is listed in Figure 47A. Similar to the results obtained considering links identified in at least two of four experiments, the overall number of identified links was higher after 3kDa CutOff spin column filtration (759 identified links after spin column filtration and 423 links after SEC). However, the major portion of link types obtained with both methods was represented by mono-linked peptides (57.7% after SEC and 51.5% after spin column filtration), followed by inter-links (34.8% after SEC and 36.2% after spin column filtration) and least abundant intra-links (7.6% for SEC and 12.3% after spin column filtration). Visualization of identified links is depicted in Figure 47B (SEC) and Figure 47C (3kDa CutOff spin column filtration). Considering only links identified in at least three of four tested samples, no crosslinks within IFT43 were identified after both enrichment methods. However, the abundances of identified links between IFT-A components varied (e.g. inter-links between IFT122 and IFT140 were one of the most prevalent inter-links after SEC, while these inter-links were low abundant after using spin columns for the enrichment of cross-linked peptides). Nevertheless, similar patterns were elucidated considering the localisation of the identified links. Again, as seen in previous experiments some of the

---

identified links were located in described interaction domains while others were identified in yet undescribed regions. These explored hotspots indicate new interacting regions.

Concluding the results of structural information of IFT-A after two different enrichment methods and keeping in mind that SEC is more time-, and cost-intensive than the use of 3kDa CutOff spin columns, this facilitated method is a nice alternative to the well described SEC obtaining similar results of types and positions of identified crosslinks.

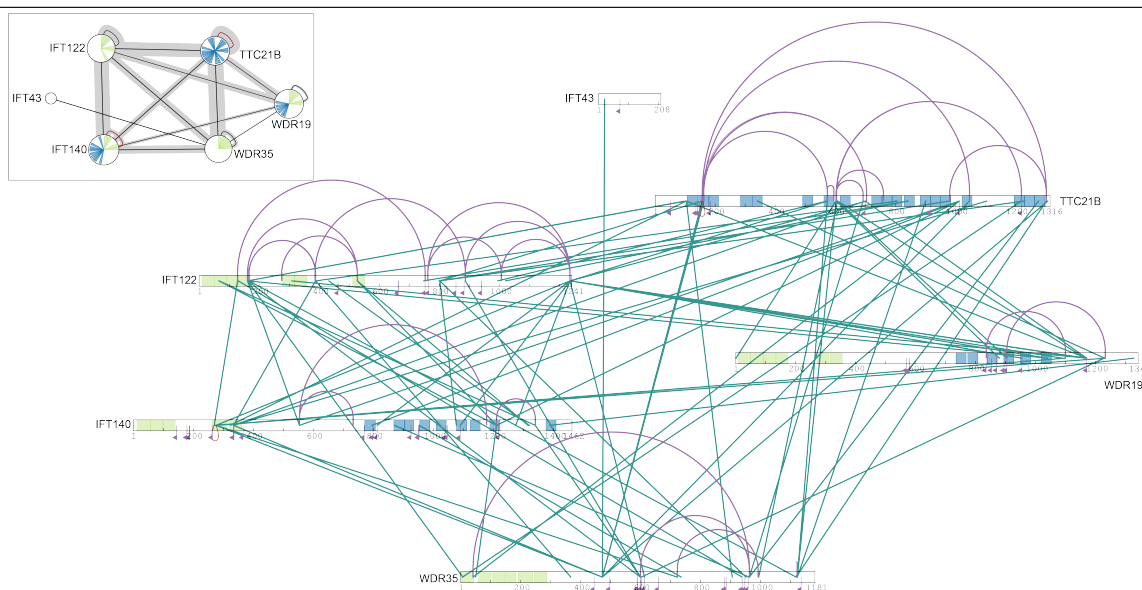
A

	Crosslinks identified after SEC (Σ708)							
	IFT122	IFT140	IFT43	TTC21B	WDR19	WDR35	-	
IFT122	26	25	58	0	47	29	24	87
IFT140	28	33	11	2	20	6	20	127
IFT43	0	10	0	0	0	0	2	2
TTC21B	69	105	2	57	29	31	37	46
WDR19	36	21	6	61	12	5	5	31
WDR35	38	46	5	71	24	11	19	45
-	89	249	0	32	52	59		
Crosslinks identified after 3kDa CutOff spin column filtration (Σ1144)								

	intra-links
	inter-links
	mono-links

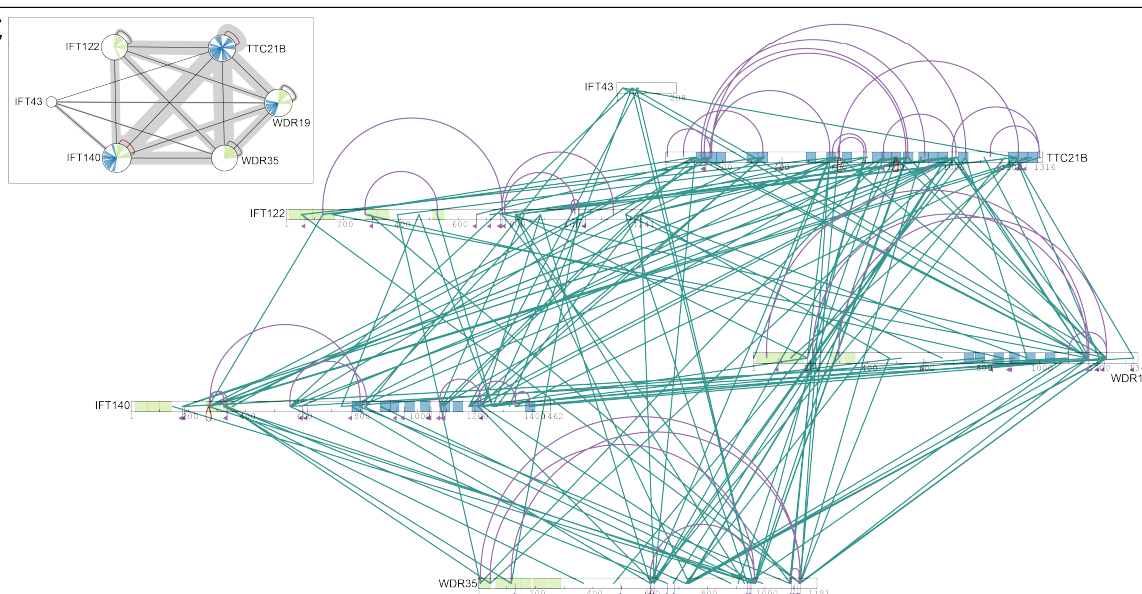
B

Crosslinks identified after SEC



C

Crosslinks identified using CutOff spin columns



**Figure 46: Identified links of IFT-A from HEK293T cells (in at least 2 of 4 replicates)**

Crosslinks within IFT-A (purified from HEK293T cells, transiently transfected with (N)-SF-TULP3) identified in at least two of four biological replicates after SEC or spin column

---

filtration. A: Number of identified links either after SEC or spin column filtration is listed. Thereby, the number of identified links was higher using 3kDa CutOff spin columns for the enrichment of crosslinked peptides (1144 identified links) than after SEC (708 links). After SEC, mono-links (47.7%) represented the major portion of crosslink types, while the major portion after spin column filtration was represented by inter-links (45.6%). B: Visualization of identified crosslinks after SEC. C: Visualization of identified crosslinks after spin column filtration. Considering patterns and localisation of identified crosslinks, results obtained with both enrichment methods were comparable, although the overall amount of identified links was higher after 3kDa CutOff spin columns were used. Visualizing the results of both methods, revealed many links in described interaction domains (WD40 domains in green and TPR domains in blue) as well as unveiled non-described interaction points within IFT-A components.



A

Crosslinks identified after SEC ( $\Sigma$ 423)							
	IFT122	IFT140	IFT43	TTC21B	WDR19	WDR35	-
IFT122	24	13	41	0	31	12	11
IFT140	11	25	0	5	0	10	85
IFT43	0	0	0	0	0	0	0
TTC21B	58	64	0	34	6	14	20
WDR19	26	0	0	31	7	0	3
WDR35	21	11	0	43	9	3	13
-	78	228	0	15	30	41	

Crosslinks identified after 3kDa CutOff spin column filtration ( $\Sigma$ 759)							
	IFT122	IFT140	IFT43	TTC21B	WDR19	WDR35	-
IFT122	24	13	41	0	31	12	11
IFT140	11	25	0	5	0	10	85
IFT43	0	0	0	0	0	0	0
TTC21B	58	64	0	34	6	14	20
WDR19	26	0	0	31	7	0	3
WDR35	21	11	0	43	9	3	13
-	78	228	0	15	30	41	

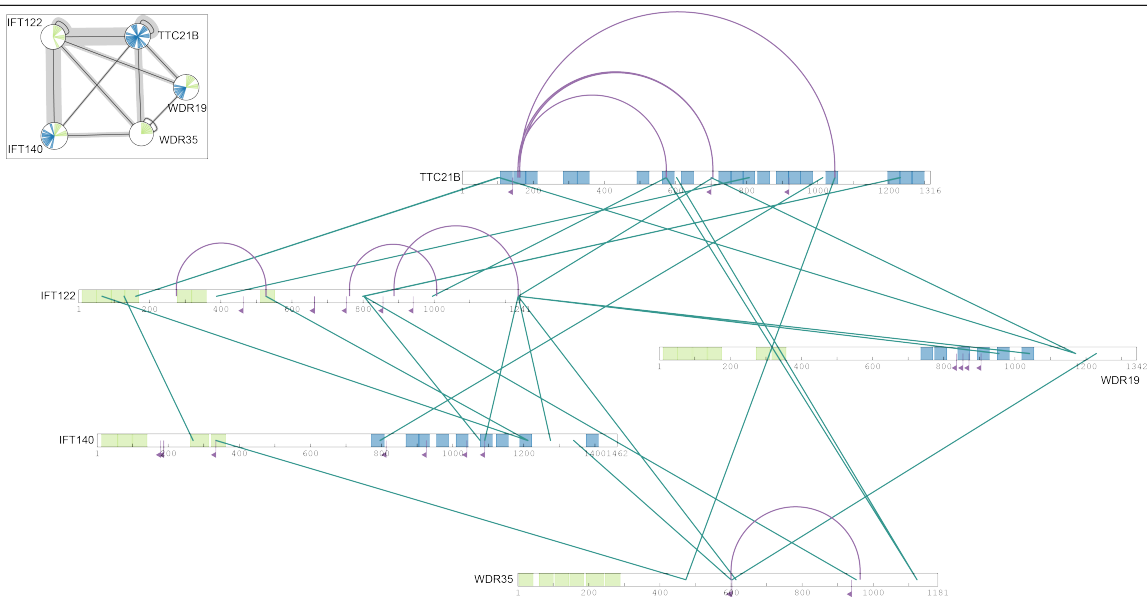
intra-links

inter-links

mono-links

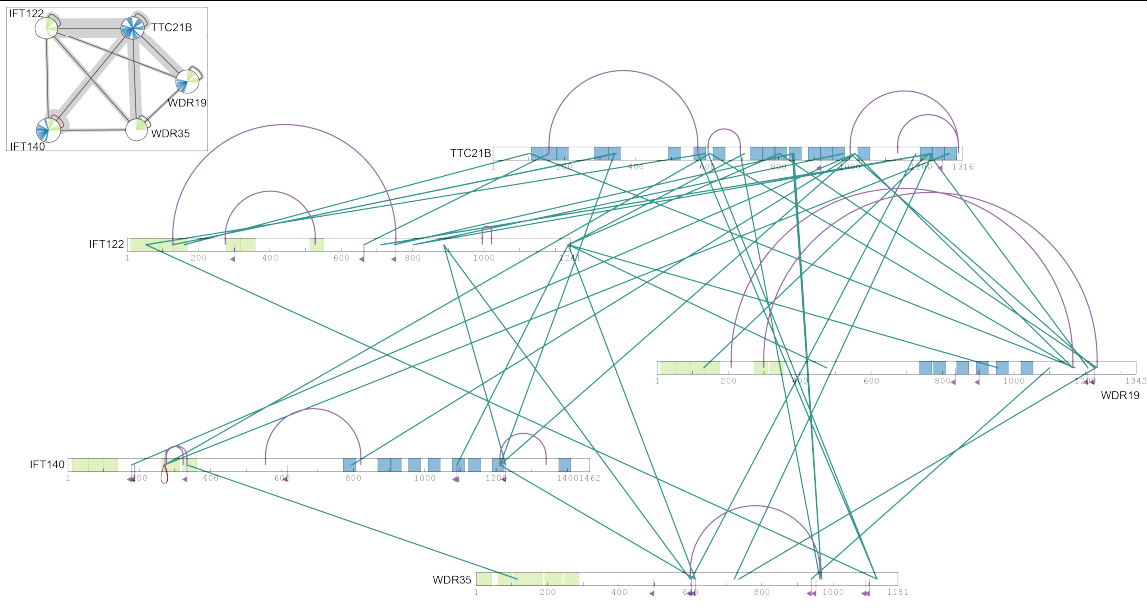
B

Crosslinks identified after SEC



C

Crosslinks identified using CutOff spin columns



**Figure 47: Identified links of IFT-A from HEK293T cells (in at least 3 of 4 replicates)**

Links identified in at least three of four biological replicates of cross-linked IFT-A, purified from transiently transfected HEK293T, are depicted. A: Number of links identified after

---

SEC or after spin column filtration is listed. The number of identified links was higher after spin column filtration (759 identified links) in contrast to 423 identified links after SEC. However, the major portion of link types was represented by mono-linked peptides (57.7% after SEC and 51.5% after spin column filtration) followed by inter- and intra-links for both enrichment methods. B: Visualization of links identified in at least three of four replicates after SEC. C: Visualization of links identified in at least three of four replicates after spin column filtration. For both enrichment methods, no inter-links between IFT43 and other IFT-A components were identified. Considering the abundances of identified links, inter-links between IFT122 and IFT140 were most prevalent in experiments performing SEC, while these inter-links were one of the least prevalent links in experiments performing spin column filtration. Although this distribution varied, both methods were similar regarding localisation of identified crosslinks either within described interaction domains or within undescribed crosslinking “hotspots”.

---

## 5 Discussion

### 5.1 Stoichiometric analysis of IFT-A

The aim of this study is to unveil the stoichiometry as well as the structural composition of IFT-A for the first time which is essential to understand its function in intraflagellar transport as well as its role in IFT-A-related ciliopathies. This study includes the stoichiometry determination of IFT-A in Flp-In monoclonal cell lines under standard growth conditions as well as the investigation of differences in IFT-A composition during different ciliary stages and alterations in stoichiometry caused by mutations introduced in genes encoding individual components of IFT-A. To circumvent the effect of artificial overexpression which may cause protein misfolding as well as complex misassembly [58], this study is based on the generation of Flp-In monoclonal lines stably expressing SF-TAP-tagged baits from a defined genomic integration site. For the purification of IFT-A in its entire composition, three different baits were used: (N)-SF-IFT122, (N)-SF-TULP3 and (N)-SF-LCA5. To eliminate the possible effect of the chosen baits on the assembly of IFT-A which could have violated the stoichiometric investigation of this study, the three baits represent proteins with different relations to IFT-A. While IFT122 is an integral part of the protein complex IFT-A, TULP3 is known to be associated with IFT-A and LCA5 is a labile and rather transient interaction partner of this protein complex [87]. Although the expression level of LCA5 was considerably lower than for the two other baits (IFT122 and TULP3; Figure 21), all six protein complex components were identified after performing one-step affinity purification of extracted proteins of all three cell lines (Table 14) as well as after silver staining with following MS analysis of TAP eluates (Figure 28). For the stoichiometry determination of IFT-A in its naturally occurring composition, mass spectrometry based absolute quantification was used. Therefore, Flp-In monoclonal cell lines were cultured in SILAC 'heavy' medium to introduce a stable isotopic label. To calculate the absolute amount of each proteotypic peptide within purified IFT-A, a known amount of a synthetic standard mixture, containing non-labelled representative peptides, was spiked in. Targeted mass spectrometry was performed to analyse the ratio of non-labelled and labelled representative peptide which was then used to determine the stoichiometry of IFT-A. Other than in previous studies [46, 59] in which spiked standard peptides consisted of corporate stable isotopes, in this study isotopic labels incorporated into the protein complex of interest were used which significantly reduced the costs. To further reduce the costs for the non-labelled standard mixture, the "Equimolarity through Equalizer Peptide" (EtEP) method which was published previously by Holzmann et al. [59] was applied. This method enables the generation of an equimolar standard mix based on representative peptides containing an equalizer peptide sequence at the N-terminus which introduces an

---

additional tryptic cleavage site. Through tryptic proteolysis, an equimolar release of the equalizer peptide and the representative peptide is initiated. Adding a known amount of an isotopically labelled equalizer peptide to the representative peptide enables the absolute quantification of the released equalizer peptide as well as of the representative peptide (see Figure 13). This method also allows the generation of a standard mixture with variable and adjusted amounts of each representative peptide. Although the idea of adjusting the internal standard mix according to the amount of the identified complex components seemed more appropriate regarding the measurement accuracy of mass spectrometer in its dynamic range [49], the variability of data using an adjusted standard mix was obviously higher due to a more complex preparation procedure. At the beginning of this study, three representative peptides were chosen for each complex component. Since four of the synthesized peptides were not dissolvable, additional new representative peptides were chosen to create another equimolar standard mix. To reduce the high variances between the chosen representative peptides for IFT122 as well as for WDR19 (Table 15), more than three representative peptides were used for the new equimolar mix. The presence of some variances between different representative peptides of the same protein, as unveiled for IFT122 and WDR19 performing SRM with an equimolar standard mix of 14 representative peptides (Table 15), may be explained by the presence of different isoforms, cleavage products or post-translational modifications of these IFT-A proteins. For example, the determined ratio (H/L) of the peptides IFSLLEK and AASVYIR which were chosen as representatives for WDR19 showed huge variances. A closer look on described isoforms of WDR19 unveiled that IFSLLEK is present in both known isoforms, while AASVYIR is only present in the longer isoform of WDR19 (Figure 48 and Figure 53). The existence of protein isoforms which may differ in purpose and property enables the functional diversity [88]. However, the general task to identify the present isoforms involved in protein complexes as well as their effect is still challenging [89].

Targeted mass spectrometry is a versatile tool to perform absolute quantification of a protein complex of interest. Different targeted mass spectrometry approaches are available for the different types of mass spectrometers. The first stoichiometry determination of this study using Selected Reaction Monitoring (SRM) was performed with the group of Karl Mechtler at the Institute of Molecular Pathology in Vienna on a QTrap5500. SRM is a well-described method to perform targeted mass spectrometry using a triple quadrupole instrument. In comparison to Parallel Reaction Monitoring (PRM), performed on a hybrid mass spectrometer comprising a quadrupole as mass filter and an Orbitrap as mass analyzer, more intense method development is necessary for a SRM approach. As described previously by Ronsein et al. in 2015 [90], both methods are comparable regarding linearity, dynamic range, precision and reproducibility while

---

the analysis of many product ions using PRM results in a gain in specificity. The comparison of the determined complex stoichiometry performing either PRM or SRM is shown in Figure 29 and Figure 30. For these data, biological replicates of each of the three bait-expressing monoclonal Flp-In cell lines were analysed performing absolute quantification using either SRM or PRM for targeted mass spectrometry analysis. However, two different equimolar standard mixtures were used. As mentioned before, the first equimolar standard mix, used for the SRM approach in this study, was consisting of only 14 representative peptides resulting in highly diverse calculated ratios (H/L) of some representative peptides of the same protein (Table 15). Thereby, only the percent coefficient of variation (%CV) of the representative peptides for three IFT-A components (IFT140, TTC21B and WDR35) was below 20% (Table 17). Further PRM approaches were performed using the new equimolar amount containing 24 representative peptides in total. This resulted in notably differences of IFT-A stoichiometry determined either by SRM or PRM (Figure 29 and Figure 30) based on more robust data with %CV value below 20% (Table 18) for all six IFT-A complex components. Nevertheless, both methods identified a huge effect on the complex composition of IFT-A using an integral component of IFT-A (IFT122) as bait. No matter which targeted mass spectrometry method was used, IFT-A composition in Flp-In (N)-SF-TULP3 and Flp-In (N)-SF-LCA5 were virtually identical, while stoichiometry of IFT-A in Flp-In (N)-SF-IFT122 differed. Thereby, not only a predominant overrepresentation of IFT122 (56% by SRM; 62% by PRM) was identified, also the portion of IFT140 (2% by SRM; 2% by PRM) and WDR19 (4% by SRM; 3% by PRM) were notably reduced (Figure 29 and Figure 30). This leads to the conclusion that the selection of an appropriate bait is a crucial point for affinity purification of the protein complex of interest without disrupting its complex composition which is in agreement with Gibson et al. [58]. Although not much is known about the composition and the molecular function of IFT-A, it consists of a core complex which is formed by three IFT-A proteins: IFT122, IFT140 and WDR19 as described previously [29, 91]. The mass of IFT-A was previously predicted with around 750kDa [91] which conforms an IFT-A stoichiometry of one single copy per IFT-A component. In contrast, considering the determined complex composition of IFT-A in Flp-In (N)-SF-TULP3 (IFT122:IFT140:IFT43:TTC21B:WDR19:WDR35 – 1.7:1.6:1.0:1.2:1.5:1.2) (see Figure 30) and keeping in mind that IFT-A consists of repetitive units [92] which hampers the determination of an exact mass of the existing protein complex, the predicted mass of this protein complex is a multiple of 1,098kDa. Based on the results presented here, the major portion of the determined IFT-A composition was represented by the IFT-A core complex with an amount of 59% (IFT122: 21%, IFT140: 20% and WDR19: 18%).

One way to improve the understanding of this IFT subcomplex is to investigate changes in complex compositions during ciliary assembly and disassembly as well as to analyse the protein

---

complex of interest if one of the components is lacking or mutated. The formation and disassembly of cilia includes many consecutive steps as depicted in Figure 2. It has been shown that IFT-A which represents a subcomplex of the intraflagellar transport machinery (IFT) is involved in retrograde ciliary transport [1, 13, 28]. This transport of protein cargo from the ciliary tip back to the cell body is indispensable for ciliary disassembly and turnover which includes many consecutive steps as depicted in Figure 2. This poses the question, if the composition of IFT-A and its complex stoichiometry changes during the assembly and the disassembly of a cilium. Based on this hypothesis, hTERT-RPE1 cells were cultivated in SILAC 'heavy' medium for absolute quantification of IFT-A during ciliogenesis. Other than HEK293T based cell lines, hTERT-RPE1 cells enable a controlled triggering of different ciliary cell stages. Starvation of hTERT-RPE1 using growth medium without fetal bovine serum (FBS) as supplement induces the formation of a cilium, whereas restimulating the cells using DMEM containing FBS promotes the disassembly of cilia [93-95]. For the investigation of stoichiometry changes during different ciliary cell stages of IFT-A in hTERT-RPE1 cells, cells were cultured under normal growth conditions, under starvation conditions and under starvation conditions followed by restimulation periods. In cilia-induced cell stages (starvation of cells), the major portion of the IFT-A complex was represented by the core complex (69%) formed by IFT140 (28%), WDR19 (25%) and IFT122 (16%). In comparison, the portion of IFT-A core increased slightly to 71% after 2h of restimulation and to 76% after a restimulation period of 4h. However, the portion of two IFT-A components (WDR35 and IFT43) decreased notably (Figure 38). These results conclude that the core complex of the IFT-A remains stable, while the amount of the other components may vary during ciliary assembly and maintenance. In contrast, the complex stoichiometry of IFT-A from hTERT-RPE1 cells cultured under the same conditions but with an increased amount of cells per culture dish, was not changed (Figure 39) and further stresses the point that cellular conditions are a crucial and sensitive factor for this experimental setup. To improve reproducibility of such sensitive experiments, cell counting may be an effective tool for future investigations of stoichiometry changes of IFT-A during different ciliary stages [96]. The second type of experiments to study induced changes of IFT-A complex stoichiometry applied in this work was targeted gene editing of genes encoding IFT-A proteins. The CRISPR/Cas9 system is a highly efficient and advanced method to create either repair-induced mutations or specific mutations at desired gene loci. In this study, mutations within IFT43 or WDR19 were generated using Flp-In (N)-SF-TULP3 cells in combination with different designed single guide RNAs (sgRNAs) as well as homology directed repair (HDR) constructs. Sequencing results of the mutated single clones identified two repair-induced mutations in IFT43 (Figure 33) and two repair-induced mutations in WDR19 (Figure 34). In contrast, the introduction of defined

---

mutations using HDR constructs failed. Although two repair mechanisms occur in dividing cells, the error-prone non-homologous end joining (NHEJ) is more prominent than the directed HDR [97, 98]. As observed in prior studies, the efficiency of HDR is depending on the concentration and length of the HDR construct, on the cell cycle as well as on the activity of the NHEJ within this cell line [99, 100]. As a consequence, the introduction of specific and disease-associated mutations is very time consuming. Using CRISPR/Cas9-based gene editing, four identified single clones including mutations either in IFT43 or WDR19 were characterized and used for one-step affinity purification of IFT-A. Only one out of the four mutant lines showed significant changes in IFT-A composition in comparison to control cells (Figure 35). This mutation is characterized by a single base insertion of an adenine (c.541\_542insA) resulting in a frameshift and premature termination of the translation of IFT43 (p.T181Nfs\*2). Considering the composition of IFT-A, purified from Flp-In cells carrying the p.T181Nfs\*2 mutation, the overall expression level of IFT-A was decreased in comparison to the control cells. Furthermore, the amount of three IFT-A components (IFT43, TTC21B and WDR35) was decreased tremendously. Although disease-causing mutations in IFT43 are already described [33], none of the published ciliopathy-related mutation was identified in or nearby this location. One of the CRISPR/Cas9-induced clones which included a mutation in WDR19 carried a deletion of two bases (c.3107\_3108delAT) and is characterized by an amino acid exchange of tyrosine to phenylalanine (Y1036Ffs\*5) causing a frame shift. This mutant is only one codon apart from a known ciliopathy-related mutation (p.S1037L). This described mutation WDR19\_p.S1037L [57] causes an amino acid substitution from serine to leucine at position 1037 of the protein. Although both variations are within the same region, no significant change in IFT-A complex stoichiometry was observed in this CRISPR/Cas9-induced clone. This allows to draw the conclusion that the amino acid sequence affected in those mutants is important to bind motor proteins or for specific cargo binding, Due to that, mutations within this domain may result in an altered binding surface of the IFT-A component WDR19. The ciliopathy-related mutation p.S1037L is caused by the exchange of a cytosine with thymine. Two-thirds of all existing single-nucleotide polymorphisms (SNPs) represents an exchange of C with T and are based on an epigenetic mechanism: Cytosine (C) methylation is performed using the enzyme methyltransferase. This enzyme transfers a methyl group from S-adenosyl-L-methionine to cytosine. Based on deamination of methylcytosine, thymine (T) is formed. Therefore, this mechanism is essential for many different functions. For example to control gene expression as well as to protect the genome against selfish DNA [101-103]. To demonstrate the potential of this method for medical as well as clinical applications, the stoichiometry of IFT-A from human fibroblast cell lines generated from two patients with Sensenbrenner Syndrome in comparison to

IFT-A purified from fibroblasts of one healthy donor, was determined in a pilot experiment within this study. Thereby, the cultivation of human fibroblasts as well as the one-step affinity purification was performed by Machteld Oud, a PhD student at the Radboud University Medical Center Department of Human Genetics in Nijmegen (Netherlands). Preliminary data of determined IFT-A stoichiometry unveiled obvious effects on IFT-A complex stoichiometry induced by both Sensenbrenner Syndrome associated mutations (one fibroblast cell line per Sensenbrenner Syndrome associated mutation) in comparison to the control fibroblast cell line. The complex composition of IFT-A purified from the control fibroblast cell line was virtually identical to the composition of IFT-A purified from Flp-In (N)-SF-TULP3 cells (Figure 37). Comparing both cells lines carrying mutations in the gene encoding IFT43, the induced effect on the composition of IFT-A was similar: Whereas the amount of WDR35, TTC21B and IFT43 was obviously decreased, the portion of the IFT-A core complex (IFT122, IFT140 and WDR19) was increased obviously. However, these results were obtained from only one biological replicate per cell line.

## 5.2 Structural investigations of IFT-A

Previous studies unveiled the train-like structure of IFT particles. Thereby, two distinct train types were identified: A long and narrow train-like structure involved in anterograde IFT and a short and compact structure of particles which is important for the retrograde ciliary movement [92]. Structural architecture of IFT-B, especially of the core complex of IFT-B was already characterized by crystal structure analysis [104, 105] whereas, only little is known about the structure and the assembly of IFT-A, so far. To generate structural information of this ciliopathy-related protein complex for the first time, chemical crosslinking of TAP-purified IFT-A was performed. Chemical crosslinking of lysine residues within protein complexes is an efficient tool to gain proximity as well as structural information about involved proteins [51, 106, 107]. According to the defined length of the used chemical crosslinker disuccinimidyl suberate (DSS) of 11.4Å and considering the length of the lysine side chains (6-6.5Å), the distance of crosslinked residues is predefined [50]. A 1:1 mixture of isotopically coded and non-coded homobifunctional disuccinimidyl suberate (DSS) was used in this study, to facilitate computational analysis using the software pipeline xQuest/xProphet, as described previously [52, 108]. Chemical crosslinking was performed with IFT-A, purified by tandem affinity purification (TAP) to obtain the complex of interest with little contaminants [69, 81]. For a successful chemical crosslinking, the ratio of protein to chemical crosslinker is important. Therefore, the experimental setup was optimised applying different crosslinker concentrations as well as different reaction buffers. Thereby, the presence of dithiotreitol (DTT) in the reaction buffer increased the number of identified crosslinks. However, the variance within tested



---

biological replicates was lower using the reaction buffer without DTT (Figure 40). The distribution of identified crosslink types differed only slightly. Nevertheless, one needs to consider that the presence of dithiotreitol within the reaction buffer reduces disulfide bonds and enables thiol groups of cysteines to be linked by DSS as well [85]. Thereby, the amount of crosslinker available for chemical crosslinking of lysine residues is effectively reduced. Based on these results and considerations, a reaction buffer without DTT was chosen for further experiments. To investigate an optimal ratio of chemical crosslinker to protein amount, 5-, 25-, 40-, 50-, or 90-fold excess of protein to chemical crosslinker were tested. While a decrease of chemical crosslinker resulted in an increased proportion of inter-molecular crosslinks and a decreased proportion of mono-linked peptides, only a slight further increase in proportion of identified inter-links was achieved upon increasing the ratio of purified protein to chemical crosslinker from 50-fold to 90-fold excess (Figure 41). Further increase in amount of DSS may result in an overrepresentation of mono-linked residues which are not able to react with a second residue because this is occupied already by another crosslinker molecule, as well. Therefore, a 50-fold excess of protein to DSS was chosen for the following experiments. While the number of identified inter-links between TTC21B and IFT140 were most abundant, only few inter-links were observed between IFT43 and other IFT-A proteins. Moreover, applying a more stringent approach illustrating only links that were identified in all three tested replicates, no crosslinks between IFT43 and other IFT-A components were depicted (Figure 44). The low number of identified crosslinks with IFT43 supports the results of determined stoichiometry of IFTA, where IFT43 is the least abundant protein in the complex. A closer look into the cross-linked positions within the IFT-A proteins unveiled the presence of crosslinking “hotspots” where crosslinks were identified more frequently. These “hotspots” were located in either known and described interaction domains like WD40 and TPR regions [109-111] or indicate novel interacting domains which are not described yet. As described previously [91], four of the six IFT-A components (IFT122, IFT140, WDR19 and WDR35) carry a similar domain distribution comprising multiple WD40 domains at the N-terminus. These interaction domains (WD40 and TPR) are estimated to stabilize the protein complex on the one hand and to undertake the selective transport of ciliary cargoes on the other hand [112]. The WD40 domain within Clathrin, a complex involved in the formation of coated vesicles [112] for example is known to selectively bind unique cargo [113]. Another example is the protein IFT46 which is a known component of the intraflagellar transport protein B (IFT-B). This protein is part of the core complex of IFT-B [114]. Its C-terminal domain is important for the complex stability [104, 115], while its N-terminal domain is essential for the ciliary transport [116-118]. Many inter-links were detected close to the C-terminus of WDR19 and WDR35. These parts of the proteins are not characterized as

interaction domains, so far. Regarding intra-links identified within TTC21B, almost every detected intra-molecular crosslink links two lysine residues within two described and characterized interaction domains (TPR). Illustrating only more reliable links, identified in all tested biological replicates of Flp-In (N)-SF-TULP3 cells, the prevalence of identified inter-links differed. Thereby, the amount of inter-links between IFT140 and WDR19 was most prevalent. No matter which stringency was applied, both indicated the existence of new non-characterized interaction domains within IFT-A complex components (Figure 44). Similar to previous experiments using IFT-A, purified from Flp-In (N)-SF-TULP3, chemical crosslinking of IFT-A, purified from HEK293T cells, identified many already described interaction domains within IFT-A components. Most of the depicted links were identified within characterized WD40- and TPR domains (Figure 46B/C). However, many crosslinks were identified in “hotspots” which are not described so far. For example, only WD40 domains at the N-terminus of WDR35 are described, but identified crosslinks were covering the whole WDR35 sequence. The same pattern was shown for IFT122. Also many of the intra-links were identified in at least one described interaction domain Figure 46B. As previously described by siRNA studies, three IFT-A components (IFT122, IFT140 and WDR19) form a core-complex of the protein complex of interest [29]. Based on the identified crosslinks, obtained in this study and illustrated using Xinet, TTC21B may be another part of this core complex, as well. The amount of identified inter-links of TTC21B with other complex components was as abundant as for IFT122, IFT140 and WDR35 (Figure 44-Figure 47).

To understand the effect of CRISPR/Cas9-induced mutations in IFT-A proteins, localisation of identified crosslinks in IFT-A, purified from wildtype cells, was considered. Some of the identified inter-links of IFT43 were close to the C-terminus of this protein (Figure 44B). Two of the CRISPR/Cas9-induced mutant clones cover mutations at the C-terminus of IFT43. One clone with a single base insertion of an adenine (c.541\_542insA/ p.T181Nfs\*2) resulted in an amino acid exchange of threonine (T) to asparagines (N) at amino acid position 181 with a further premature stop in the translation of IFT43. This mutant reduces the size of the protein IFT43 notably. In contrast, the second mutation caused by a single base insertion (c.582\_583insG/ p.Q195Afs\*22) resulted in a slightly prolonged form of IFT43 (Figure 33). As depicted in Figure 35, a tremendous effect on the complex composition of IFT-A was only obtained in IFT-A purified from the clone carrying the mutation resulting in a shortened IFT43 (IFT43\_ c.541\_542insA/ p.T181Nfs\*2). The loss in binding of IFT43, TTC21B and WDR35 can be explained considering the identified crosslinks of IFT43. Most of the inter-links, identified close to the C-terminus of IFT43, were linked to TTC21B. A loss of the C-terminal half of IFT43, caused by the c.541\_542insA/ p.T181Nfs\*2 mutation, resulted in a reduction of TTC21B and

---

proteins linked to TTC21B like WDR35, although there were still many other cross-links present. This allows to draw the conclusion that IFT43 may function as a stabilizer of the repetitive structure of IFT-A. Regarding the second mutation c.582\_583insG/ p.Q195Afs\*22, the extension of IFT43 did not cause an apparent effect on complex composition of IFT-A. This allows to draw the conclusion that the amino acid sequence around the mutation site of the clone is crucial to stabilize IFT-A. Both generated CRISPR/Cas9-induced clones mutated in genes encoding WDR19 (p.Y1036Ffs\*5 and p.R1022Dfs\*22), resulted in an alteration within a described TPR domain. Analysing the complex stoichiometry of both mutations identified no alteration in comparison to the control (Figure 35). Regarding identified inter-links within the TPR domain of WDR19 (around amino acid position 1020), only one inter-link within this domain of WDR19 with the N-terminal half of IFT140 was identified. This refers to the conclusion that this identified linkage between the identified domains of IFT140 and WDR19 is not essential to keep the composition of IFT-A in its naturally occurring state, however it might be important for binding of external proteins like specific cargoes.

As described previously [51], for LC-MS/MS analysis of low abundant cross-linked peptides a previous enrichment step of these molecules is crucial. Size exclusion chromatography (SEC) is a well-described and commonly performed prefractionation method to enrich cross-linked peptides, although SEC is time-, labour- and cost-intensive. To facilitate this prefractionation, an economic method for the reduction of sample complexity is described in this study. This easy-to-handle method is characterized by the use of 3kDa CutOff spin columns. This part of the study includes the comparison of both prefractionation methods with regard to the amount and the localisation of identified crosslinks as well as the applicability of this innovative enrichment method for the following LC-MS/MS analysis of cross-linked peptides. Therefore, purified IFT-A was chemically cross-linked using DSS. After proteolysis of cross-linked proteins, sample was split and prefractionation was performed either by SEC or spin column filtration. The precursor distribution of all identified precursor ions showed same patterns for both methods considering lower mass ranges (1500-4500). Using spin columns, however no precursor ions with relative molecular mass ( $M_r$ ) higher than 5500 were detected, although 3kDa CutOff spin columns should restrain peptides with a molecular mass >3kDa. The overall amount of crosslinks was higher performing SEC, except the amount of mono-links. Considering the proportion of identified link types, the fractions of identified crosslinks varied between both enrichment methods (Figure 42). Since the available amount of IFT-A purified from Flp-In (N)-SF-TULP3 cells is limited, following experiment was performed using transiently transfected HEK293T cells to compare SEC with spin column filtration. Again, sample was split after proteolysis and either spin column filtration or SEC was performed to enrich cross-linked peptides. Using HEK293T

---

cells, transiently transfected with (N)-SF-TULP3, precursor distribution patterns of both prefractionation methods were comparable (Figure 43). In contrast to previous experiments using IFT-A purified from Flp-In cells, the overall amount of identified precursors was higher using 3kDa CutOff spin columns except in the relative molecular mass range of 1500-2000. Furthermore, the obtained standard deviation of the four tested biological replicates was significantly higher performing SEC, especially for the amount of identified inter-links. While the overall amount of identified crosslinks (no matter which crosslink type) was higher using 3kDa CutOff spin columns, the portion of mono-links was higher after SEC (Figure 43). One needs to consider that the amount of IFT-A, purified from Flp-In cells is lower than purified from transiently transfected HEK293T cells which may explain the differences in both experiments regarding different cell types. For example the material of spin columns may bind a certain amount of peptides irreversibly which may be irrelevant if the amount of starting material exceeds a certain amount of applied peptides. In contrast to SEC, spin column filtration is an easy-to-handle and much faster approach which does not require a lot of LC experience. Due to this facilitated method, every lab without the required equipment for SEC can easily perform chemical crosslinking with further mass spectrometric analysis.

### 5.3 Modelling IFT-A

As validated in previous studies, chemical crosslinking together with mass spectrometry analysis enables structural modelling of a protein complex of interest comparable with investigated crystal structures [119-121]. This combination of chemical crosslinking with mass spectrometry provides great potential in elucidating and characterizing structural information of a protein complex of interest [122]. Thereby, generated mass spectrometry data has to be incorporated into modelling algorithms as constraint for potential models. Until now, this data integration constitutes a big bottleneck [50]. The combination of determined complex stoichiometry of IFT-A with crosslinking positions, identified in this study, can be used to generate a structural model of the protein complex IFT-A which may facilitate the understanding of its function during ciliary assembly and disassembly as well as its role in ciliopathies. Complex composition of IFT-A identified three IFT-A proteins (IFT122, WDR19 and IFT140) as most prevalent complex components (Figure 30). This promotes the presence of a core-complex of IFT-A [29]. TTC21B was more prevalent in the determined IFT-A composition than WDR35 and IFT43 which was furthermore indicated by the amount of identified crosslinks between TTC21B and other IFT-A components (Figure 44). However, chemical crosslinking has a great disadvantage: The irregular distribution of lysine residues all over the proteins leads to a patchy data generation [50] using a lysine-reactive crosslinker like DSS for chemical crosslinking. Thereby, only positive crosslinking results can be recognised for the structural

---

modelling of IFT-A because the absence of crosslinks may be due to sterical as well as chemical reasons [41]. Prior studies unveiled two different train-like IFT particles using electron-tomographic analysis of the IFT trains. These train-like IFT particles consist of repetitive units: The long and narrow trains for the anterograde transport as well as the short and compact particle trains involved in retrograde IFT [92]. Putting all these outcomes together, a detailed and meaningful model of the IFT-A can be generated to create the foundation for further studies of IFT-A-related ciliopathies. Nevertheless, a less complex assembly and stoichiometry of IFT-A was expected. Considering the obtained information of IFT-A, modelling of this protein complex represents a massive task because so many combinations of the structural composition of IFT-A are possible. Additionally, the intra-links can be inter-links between two copies of a protein. This makes it even more complex.

## 5.4 Perspectives

The stoichiometric and structural investigations of the protein complex IFT-A, unveiled in this study, provided new insight into the composition and the stoichiometry alterations of IFT-A during different ciliary stages. However, these obtained results allow to draw new conclusions and propose new questions. For example, the influence of cell confluence on the complex composition of IFT-A unveiled the complex stoichiometry as sensitive and flexible at least during ciliogenesis or ciliary disassembly. To study stoichiometry alterations induced by targeted CRISPR/Cas9 mutants, further optimisation of this approach are crucial to promote the homology directed repair (HDR) mechanism. Based on the designed HDR constructs, there is an opportunity to simulate the IFT-A composition present in identified IFT-A-related ciliopathies like Sensenbrenner Syndrome. The attempt to introduce the described amino acid exchange of serine to leucine at amino acid position 1037 in WDR19 (p.S1037L) using a designed single guide RNA (sgRNA) only resulted in a repair-induced mutation (WDR19\_p.Y1036F). Although non-homologous end joining (NHEJ) is the predominate DNA repair mechanism, the homology directed repair (HDR) mechanism can be encouraged by selected reaction conditions like cell cycle or the concentration and length of the designed HDR construct [99, 100]. The comparison of the IFT-A stoichiometry from a targeted ciliopathy-related mutant within an IFT-A protein with the complex composition of IFT-A from control cells might shed light upon the underlying molecular mechanisms of ciliopathies by identifying the role of IFT-A within the intraflagellar transport. Although, generated Flp-In cells are of human origin, the study of stoichiometry alterations of IFT-A in human fibroblasts of patients suffering from ciliopathy diseases as well as from healthy donors would be a great benefit for medical and clinical applications. The first pilot experiment with one replicate per fibroblast cell line showed virtually identical results compared to IFT-A from Flp-In cell lines (Figure 37). However, more biological replicates have to be

---

analysed to promote this conclusion. To find out if there is a common underlying mechanism of IFT-A-related ciliopathies, IFT-A composition within more patients has to be analysed. The approach of absolute quantification of IFT-A, presented in this study, is a robust tool that enables the exact determination of alterations in complex composition and also to distinguish between alternative mechanisms.

Chemical crosslinking of CRISPR/Cas9-induced mutant Flp-In (N)-SF-TULP3 monoclonal cells could be the basis for further investigations of disease-related structural changes in IFT-A composition. Thereby, the structural investigation can shed light on the function and role of each IFT-A complex component. For this task it would be great to have tools at hand that allow the quantification of identified crosslinks. However, available software tools are limited to identification of crosslinks, so far. To circumvent the patchy nature of chemical crosslinking considering the irregular distribution of lysine residues all over the surface of purified IFT-A, the use of other crosslinker types (varying in reactive site or spacer length) represents a promising tool to gain additional structural information about the protein complex of interest [50]. For the evaluation of the facilitated enrichment method of cross-linked peptides, portrayed in this study, another approach using other protein complexes to unveil its efficiency may be set up. However, this includes much optimization effort to enable an efficient chemical crosslinking of the purified protein complex of interest.

This study presented an effective and cost-efficient approach to study the stoichiometric as well as the structural composition of IFT-A. Based on the novel results, obtained in this study, other protein complexes of interest can be studied the same way to identify their role and molecular function.

---

## 6 References

1. Ishikawa, H. and W.F. Marshall, *Ciliogenesis: building the cell's antenna*. Nat Rev Mol Cell Biol, 2011. **12**(4): p. 222-34.
2. Hildebrandt, F., T. Benzing, and N. Katsanis, *Ciliopathies*. N Engl J Med, 2011. **364**(16): p. 1533-43.
3. Avasthi, P. and W.F. Marshall, *Stages of ciliogenesis and regulation of ciliary length*. Differentiation, 2012. **83**(2): p. S30-42.
4. Ishikawa, H., et al., *Proteomic analysis of mammalian primary cilia*. Curr Biol, 2012. **22**(5): p. 414-9.
5. Berbari, N.F., et al., *The primary cilium as a complex signaling center*. Curr Biol, 2009. **19**(13): p. R526-35.
6. Baker, K. and P.L. Beales, *Making sense of cilia in disease: the human ciliopathies*. Am J Med Genet C Semin Med Genet, 2009. **151C**(4): p. 281-95.
7. Williams, C.L., et al., *MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis*. J Cell Biol, 2011. **192**(6): p. 1023-41.
8. Kee, H.L., et al., *A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia*. Nat Cell Biol, 2012. **14**(4): p. 431-7.
9. Hoyle, H.D., et al., *Tubulin sorting during dimerization in vivo*. Mol Biol Cell, 2001. **12**(7): p. 2185-94.
10. Satir, P. and S.T. Christensen, *Structure and function of mammalian cilia*. Histochem Cell Biol, 2008. **129**(6): p. 687-93.
11. Reiter, J.F., O.E. Blacque, and M.R. Leroux, *The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization*. EMBO Rep, 2012. **13**(7): p. 608-18.
12. Afzelius, B.A., *Cilia-related diseases*. J Pathol, 2004. **204**(4): p. 470-7.
13. Yildiz, O. and H. Khanna, *Ciliary signaling cascades in photoreceptors*. Vision Res, 2012. **75**: p. 112-6.
14. Sorokin, S., *Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells*. J Cell Biol, 1962. **15**: p. 363-77.
15. Anderson, R.G., *The three-dimensional structure of the basal body from the rhesus monkey oviduct*. J Cell Biol, 1972. **54**(2): p. 246-65.
16. Pearson, C.G., T.H. Giddings, Jr., and M. Winey, *Basal body components exhibit differential protein dynamics during nascent basal body assembly*. Mol Biol Cell, 2009. **20**(3): p. 904-14.
17. Wheatley, D.N., *Cilia in cell-cultured fibroblasts. I. On their occurrence and relative frequencies in primary cultures and established cell lines*. J Anat, 1969. **105**(Pt 2): p. 351-62.

18. Masland, R.H., *The functional architecture of the retina*. Sci Am, 1986. **255**(6): p. 102-11.
19. Wang, J.S. and V.J. Kefalov, *The cone-specific visual cycle*. Prog Retin Eye Res, 2011. **30**(2): p. 115-28.
20. Rachel, R.A., T. Li, and A. Swaroop, *Photoreceptor sensory cilia and ciliopathies: focus on CEP290, RPGR and their interacting proteins*. Cilia, 2012. **1**(1): p. 22.
21. Liu, Q., et al., *RP1 is required for the correct stacking of outer segment discs*. Invest Ophthalmol Vis Sci, 2003. **44**(10): p. 4171-83.
22. Sahel, J.A., *Spotlight on childhood blindness*. J Clin Invest, 2011. **121**(6): p. 2145-9.
23. Emmer, B.T., D. Maric, and D.M. Engman, *Molecular mechanisms of protein and lipid targeting to ciliary membranes*. J Cell Sci, 2010. **123**(Pt 4): p. 529-36.
24. Sung, C.H. and M.R. Leroux, *The roles of evolutionarily conserved functional modules in cilia-related trafficking*. Nat Cell Biol, 2013. **15**(12): p. 1387-97.
25. Nachury, M.V., E.S. Seeley, and H. Jin, *Trafficking to the ciliary membrane: how to get across the periciliary diffusion barrier?* Annu Rev Cell Dev Biol, 2010. **26**: p. 59-87.
26. Benzing, T. and B. Schermer, *Transition zone proteins and cilia dynamics*. Nat Genet, 2011. **43**(8): p. 723-4.
27. Rosenbaum, J.L. and G.B. Witman, *Intraflagellar transport*. Nat Rev Mol Cell Biol, 2002. **3**(11): p. 813-25.
28. Taschner, M., S. Bhogaraju, and E. Lorentzen, *Architecture and function of IFT complex proteins in ciliogenesis*. Differentiation, 2012. **83**(2): p. S12-22.
29. Mukhopadhyay, S. and P.K. Jackson, *The tubby family proteins*. Genome Biol, 2011. **12**(6): p. 225.
30. Anand, M. and H. Khanna, *Ciliary transition zone (TZ) proteins RPGR and CEP290: role in photoreceptor cilia and degenerative diseases*. Expert Opin Ther Targets, 2012. **16**(6): p. 541-51.
31. Walczak-Sztulpa, J., et al., *Cranioectodermal Dysplasia, Sensenbrenner syndrome, is a ciliopathy caused by mutations in the IFT122 gene*. Am J Hum Genet, 2010. **86**(6): p. 949-56.
32. Gilissen, C., et al., *Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome*. Am J Hum Genet, 2010. **87**(3): p. 418-23.
33. Arts, H.H., et al., *C14ORF179 encoding IFT43 is mutated in Sensenbrenner syndrome*. J Med Genet, 2011. **48**(6): p. 390-5.
34. Bredrup, C., et al., *Ciliopathies with skeletal anomalies and renal insufficiency due to mutations in the IFT-A gene WDR19*. Am J Hum Genet, 2011. **89**(5): p. 634-43.
35. Davis, E.E., et al., *TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum*. Nat Genet, 2011. **43**(3): p. 189-96.



- 
36. Mill, P., et al., *Human and mouse mutations in WDR35 cause short-rib polydactyly syndromes due to abnormal ciliogenesis*. Am J Hum Genet, 2011. **88**(4): p. 508-15.
  37. Eggenschwiler, J.T. and K.V. Anderson, *Cilia and developmental signaling*. Annu Rev Cell Dev Biol, 2007. **23**: p. 345-73.
  38. Mukhopadhyay, S., et al., *TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia*. Genes Dev, 2010. **24**(19): p. 2180-93.
  39. Williamson, S.M., et al., *Probing the role of IFT particle complex A and B in flagellar entry and exit of IFT-dynein in Chlamydomonas*. Protoplasma, 2012. **249**(3): p. 851-6.
  40. Perrault, I., et al., *Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations*. Am J Hum Genet, 2012. **90**(5): p. 864-70.
  41. Lottspeich, F., *Bioanalytik*. 2. Aufl., aktualisierte u. erw. Neuaufl. ed. SAV Biowissenschaften. 2006, München ; Heidelberg: Spektrum. 1119.
  42. Yates lii, J.R., *A century of mass spectrometry: from atoms to proteomes*. Nat Meth, 2011. **8**(8): p. 633-637.
  43. Thomson, J.J., XL. *Cathode Rays*. Philosophical Magazine Series 5, 1897. **44**(269): p. 293-316.
  44. Rehm, H. and T. Letzel, *Der Experimentator: Proteinbiochemie, Proteomics*. 6. Aufl. ed. Der Experimentator. 2010, Heidelberg: Spektrum Akademischer Verlag. 390.
  45. Domon, B. and R. Aebersold, *Mass spectrometry and protein analysis*. Science, 2006. **312**(5771): p. 212-7.
  46. Gerber, S.A., et al., *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS*. Proc Natl Acad Sci U S A, 2003. **100**(12): p. 6940-5.
  47. Gygi, S.P., et al., *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology*. Proc Natl Acad Sci U S A, 2000. **97**(17): p. 9390-5.
  48. Jarnuczak, A.F., et al., *Analysis of Intrinsic Peptide Detectability via Integrated Label-Free and SRM-Based Absolute Quantitative Proteomics*. J Proteome Res, 2016.
  49. Ludwig, C. and R. Aebersold, *CHAPTER 4 Getting Absolute: Determining Absolute Protein Quantities via Selected Reaction Monitoring Mass Spectrometry*, in *Quantitative Proteomics*. 2014, The Royal Society of Chemistry. p. 80-109.
  50. Rappsilber, J., *The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes*. J Struct Biol, 2011. **173**(3): p. 530-40.
  51. Leitner, A., T. Walzthoeni, and R. Aebersold, *Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-*

- MS/MS and the xQuest/xProphet software pipeline*. Nat Protoc, 2014. **9**(1): p. 120-37.
52. Rinner, O., et al., *Identification of cross-linked peptides from large sequence databases*. Nat Methods, 2008. **5**(4): p. 315-8.
53. Hull, S., et al., *Nonsyndromic Retinal Dystrophy due to Bi-Allelic Mutations in the Ciliary Transport Gene IFT140*. Invest Ophthalmol Vis Sci, 2016. **57**(3): p. 1053-62.
54. Katoh, Y., et al., *Overall Architecture of the Intraflagellar Transport (IFT)-B Complex Containing Cluap1/IFT38 as an Essential Component of the IFT-B Peripheral Subcomplex*. J Biol Chem, 2016. **291**(21): p. 10962-75.
55. Perrault, I., et al., *IFT81, encoding an IFT-B core protein, as a very rare cause of a ciliopathy phenotype*. J Med Genet, 2015. **52**(10): p. 657-65.
56. Taschner, M. and E. Lorentzen, *Recombinant Reconstitution and Purification of the IFT-B Core Complex from Chlamydomonas reinhardtii*. Methods Mol Biol, 2016. **1454**: p. 69-82.
57. Boldt, K., et al., *An organelle-specific protein landscape identifies novel diseases and molecular mechanisms*. Nat Commun, 2016. **7**: p. 11491.
58. Gibson, T.J., M. Seiler, and R.A. Veitia, *The transience of transient overexpression*. Nat Methods, 2013. **10**(8): p. 715-21.
59. Holzmann, J., et al., *Stoichiometry determination of the MP1-p14 complex using a novel and cost-efficient method to produce an equimolar mixture of standard peptides*. Anal Chem, 2009. **81**(24): p. 10254-61.
60. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system*. Nat Protoc, 2013. **8**(11): p. 2281-308.
61. Landy, A., *Dynamic, structural, and regulatory aspects of lambda site-specific recombination*. Annu Rev Biochem, 1989. **58**: p. 913-49.
62. Bendall, S.C., et al., *Prevention of amino acid conversion in SILAC experiments with embryonic stem cells*. Mol Cell Proteomics, 2008. **7**(9): p. 1587-97.
63. Lossner, C., et al., *Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions*. Anal Biochem, 2011. **412**(1): p. 123-5.
64. Kim, T.K. and J.H. Eberwine, *Mammalian cell transfection: the present and the future*. Anal Bioanal Chem, 2010. **397**(8): p. 3173-8.
65. O'Gorman, S., D.T. Fox, and G.M. Wahl, *Recombinase-mediated gene activation and site-specific integration in mammalian cells*. Science, 1991. **251**(4999): p. 1351-5.
66. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications*. Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4350-4.

- 
67. Gloeckner, C.J., et al., *Tandem affinity purification of protein complexes from mammalian cells by the Strep/FLAG (SF)-TAP tag*. *Methods Mol Biol*, 2009. **564**: p. 359-72.
  68. Gloeckner, C.J., et al., *A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes*. *Proteomics*, 2007. **7**(23): p. 4228-34.
  69. Gloeckner, C.J., K. Boldt, and M. Ueffing, *Strep/FLAG tandem affinity purification (SF-TAP) to study protein interactions*. *Curr Protoc Protein Sci*, 2009. **Chapter 19**: p. Unit19 20.
  70. Wessel, D. and U.I. Flugge, *A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids*. *Anal Biochem*, 1984. **138**(1): p. 141-3.
  71. Rappsilber, J., Y. Ishihama, and M. Mann, *Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics*. *Anal Chem*, 2003. **75**(3): p. 663-70.
  72. Picotti, P. and R. Aebersold, *Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions*. *Nat Methods*, 2012. **9**(6): p. 555-66.
  73. Olsen, J.V., et al., *Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap*. *Mol Cell Proteomics*, 2005. **4**(12): p. 2010-21.
  74. Lange, V., et al., *Selected reaction monitoring for quantitative proteomics: a tutorial*. *Mol Syst Biol*, 2008. **4**: p. 222.
  75. Keller, A., et al., *Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search*. *Anal Chem*, 2002. **74**(20): p. 5383-92.
  76. Nesvizhskii, A.I., et al., *A statistical model for identifying proteins by tandem mass spectrometry*. *Anal Chem*, 2003. **75**(17): p. 4646-58.
  77. Leitner, A., et al., *Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics*. *Mol Cell Proteomics*, 2010. **9**(8): p. 1634-49.
  78. MacLean, B., et al., *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments*. *Bioinformatics*, 2010. **26**(7): p. 966-8.
  79. Makarov, A., et al., *Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer*. *J Am Soc Mass Spectrom*, 2006. **17**(7): p. 977-82.
  80. Brogna, S. and J. Wen, *Nonsense-mediated mRNA decay (NMD) mechanisms*. *Nat Struct Mol Biol*, 2009. **16**(2): p. 107-13.
  81. Oeffinger, M., *Two steps forward--one step back: advances in affinity purification mass spectrometry of macromolecular complexes*. *Proteomics*, 2012. **12**(10): p. 1591-608.

- 
82. Cleland, W.W., *Dithiothreitol, a New Protective Reagent for Sh Groups*. Biochemistry, 1964. **3**: p. 480-2.
  83. Coiner, H., et al., *Methylation of sulfhydryl groups: a new function for a family of small molecule plant O-methyltransferases*. Plant J, 2006. **46**(2): p. 193-205.
  84. Anfinsen, C.B., *The formation and stabilization of protein structure*. Biochem J, 1972. **128**(4): p. 737-49.
  85. Leavell, M.D., et al., *Strategy for selective chemical cross-linking of tyrosine and lysine residues*. J Am Soc Mass Spectrom, 2004. **15**(11): p. 1604-11.
  86. Anjaneyulu, P.S. and J.V. Staros, *Reactions of N-hydroxysulfosuccinimide active esters*. Int J Pept Protein Res, 1987. **30**(1): p. 117-24.
  87. Texier, Y., et al., *Elution profile analysis of SDS-induced subcomplexes by quantitative mass spectrometry*. Mol Cell Proteomics, 2014. **13**(5): p. 1382-91.
  88. Kirkwood, K.J., et al., *Characterization of native protein complexes and protein isoform variation using size-fractionation-based quantitative proteomics*. Mol Cell Proteomics, 2013. **12**(12): p. 3851-73.
  89. Angelini, C., D. De Canditiis, and I. De Feis, *Computational approaches for isoform detection and estimation: good and bad news*. BMC Bioinformatics, 2014. **15**: p. 135.
  90. Ronsein, G.E., et al., *Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics*. J Proteomics, 2015. **113**: p. 388-99.
  91. Behal, R.H., et al., *Subunit interactions and organization of the Chlamydomonas reinhardtii intraflagellar transport complex A proteins*. J Biol Chem, 2012. **287**(15): p. 11689-703.
  92. Pigino, G., et al., *Electron-tomographic analysis of intraflagellar transport particle trains in situ*. J Cell Biol, 2009. **187**(1): p. 135-48.
  93. Spalluto, C., D.I. Wilson, and T. Hearn, *Evidence for reciliation of RPE1 cells in late G1 phase, and ciliary localisation of cyclin B1*. FEBS Open Bio, 2013. **3**: p. 334-40.
  94. Tucker, R.W., A.B. Pardee, and K. Fujiwara, *Centriole ciliation is related to quiescence and DNA synthesis in 3T3 cells*. Cell, 1979. **17**(3): p. 527-35.
  95. Lim, Y.C., et al., *Culture and detection of primary cilia in endothelial cell models*. Cilia, 2015. **4**: p. 11.
  96. Ongena, K., et al., *Determining cell number during cell culture using the Scepter cell counter*. J Vis Exp, 2010(45).
  97. Mao, Z., et al., *Comparison of nonhomologous end joining and homologous recombination in human cells*. DNA Repair (Amst), 2008. **7**(10): p. 1765-71.
  98. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway*. Annu Rev Biochem, 2010. **79**: p. 181-211.

- 
99. Heyer, W.D., K.T. Ehmsen, and J. Liu, *Regulation of homologous recombination in eukaryotes*. *Annu Rev Genet*, 2010. **44**: p. 113-39.
  100. Lin, S., et al., *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery*. *Elife*, 2014. **3**: p. e04766.
  101. Jeltsch, A., *Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases*. *Chembiochem*, 2002. **3**(4): p. 274-93.
  102. Seisenberger, S., et al., *Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers*. *Philos Trans R Soc Lond B Biol Sci*, 2013. **368**(1609): p. 20110330.
  103. Marinus, M.G. and J. Casadesus, *Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more*. *FEMS Microbiol Rev*, 2009. **33**(3): p. 488-503.
  104. Taschner, M., et al., *Biochemical mapping of interactions within the intraflagellar transport (IFT) B core complex: IFT52 binds directly to four other IFT-B subunits*. *J Biol Chem*, 2011. **286**(30): p. 26344-52.
  105. Taschner, M., et al., *Crystal structures of IFT70/52 and IFT52/46 provide insight into intraflagellar transport B core complex assembly*. *J Cell Biol*, 2014. **207**(2): p. 269-82.
  106. Iglesias, A.H., L.F. Santos, and F.C. Gozzo, *Collision-induced dissociation of Lys-Lys intramolecular crosslinked peptides*. *J Am Soc Mass Spectrom*, 2009. **20**(4): p. 557-66.
  107. Merkley, E.D., et al., *Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances*. *Protein Sci*, 2014. **23**(6): p. 747-59.
  108. Petrotchenko, E.V., J.J. Serpa, and C.H. Borchers, *An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics*. *Mol Cell Proteomics*, 2011. **10**(2): p. M110 001420.
  109. Cole, D.G., *The intraflagellar transport machinery of Chlamydomonas reinhardtii*. *Traffic*, 2003. **4**(7): p. 435-42.
  110. Jekely, G. and D. Arendt, *Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium*. *Bioessays*, 2006. **28**(2): p. 191-8.
  111. van Dam, T.J., et al., *Evolution of modular intraflagellar transport from a coatomer-like progenitor*. *Proc Natl Acad Sci U S A*, 2013. **110**(17): p. 6943-8.
  112. Bhogaraju, S., B.D. Engel, and E. Lorentzen, *Intraflagellar transport complex structure and cargo interactions*. *Cilia*, 2013. **2**(1): p. 10.
  113. ter Haar, E., S.C. Harrison, and T. Kirchhausen, *Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin*. *Proc Natl Acad Sci U S A*, 2000. **97**(3): p. 1096-100.
  114. Lucker, B.F., et al., *Characterization of the intraflagellar transport complex B core: direct interaction of the IFT81 and IFT74/72 subunits*. *J Biol Chem*, 2005. **280**(30): p. 27688-96.

- 
115. Lucker, B.F., et al., *Direct interactions of intraflagellar transport complex B proteins IFT88, IFT52, and IFT46*. J Biol Chem, 2010. **285**(28): p. 21508-18.
  116. Hou, Y., et al., *Functional analysis of an individual IFT protein: IFT46 is required for transport of outer dynein arms into flagella*. J Cell Biol, 2007. **176**(5): p. 653-65.
  117. Ahmed, N.T., et al., *ODA16 aids axonemal outer row dynein assembly through an interaction with the intraflagellar transport machinery*. J Cell Biol, 2008. **183**(2): p. 313-22.
  118. Ahmed, N.T. and D.R. Mitchell, *ODA16p, a Chlamydomonas flagellar protein needed for dynein assembly*. Mol Biol Cell, 2005. **16**(10): p. 5004-12.
  119. Rappsilber, J., et al., *A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry*. Anal Chem, 2000. **72**(2): p. 267-75.
  120. Heymann, M., et al., *MSX-3D: a tool to validate 3D protein models using mass spectrometry*. Bioinformatics, 2008. **24**(23): p. 2782-3.
  121. Guaitoli, G., et al., *Structural model of the dimeric Parkinson's protein LRRK2 reveals a compact architecture involving distant interdomain contacts*. Proc Natl Acad Sci U S A, 2016. **113**(30): p. E4357-66.
  122. Zelter, A., et al., *The molecular architecture of the Dam1 kinetochore complex is defined by cross-linking based structural modelling*. Nat Commun, 2015. **6**: p. 8673.

## 7 Annex

### 7.1 Additional Figures and Tables

Table 17: Absolute quantification of IFT-A using Flp-In (N)-SF-LCA5 performing SRM

Protein	Peptide	Pre. m/z	Prod. m/z	Trans.	RT	CE	Ratio H/L									
							Trans.	SD <sub>Trans</sub>	%CV <sub>Trans</sub>	Median <sub>Peptide</sub>	SD <sub>Peptide</sub>	%CV <sub>Peptide</sub>	Median <sub>Protein</sub>	SD <sub>Protein</sub>	%CV <sub>Protein</sub>	Stoich.
IFT122	WDEAFALGEK	583.279846	864.446159	+2y6	43.97	26.9	<b>0.3212</b>	0.0058	1.8	<b>0.3246</b>	0.0192	5.9	<b>1.1046</b>	1.1031	99.9	<b>6.77</b>
			735.403566	+2y7	43.97	28.9	<b>0.3279</b>	0.0263	8.0							
			664.366452	+2y6	43.97	27.9	<b>0.3043</b>	0.0110	3.6							
			333.176861	+2y3	43.97	34.9	<b>0.3506</b>	0.0199	5.7							
	ILFTLAK	403.262739	692.434139	+2y6	41.30	19.4	<b>2.3033</b>	0.0637	2.8	<b>1.8846</b>	0.2362	12.5				
579.350074	+2y5	41.30	18.4	<b>1.8792</b>	0.0436	2.3										
432.28166	+2y4	41.30	24.4	<b>1.89</b>	0.1177	6.2										
331.233982	+2y3	41.30	27.4	<b>1.7648</b>	0.0995	5.6										
IFT140	FWDIER	433.213777	718.351865	+2y5	40.53	20.5	<b>0.2319</b>	0.0082	3.6	<b>0.2316</b>	0.0034	1.5	<b>0.2031</b>	0.0403	19.8	<b>1.25</b>
			532.272552	+2y4	40.53	20.5	<b>0.2259</b>	0.0046	2.0							
			417.245609	+2y3	40.53	28.5	<b>0.2316</b>	0.0073	3.1							
	AHGALTEAYK	530.774731	852.446159	+2y8	16.47	26.0	<b>0.1683</b>	0.0062	3.7	<b>0.1747</b>	0.0177	10.1				
	795.424696	+2y7	16.47	27.0	<b>0.1652</b>	0.0256	15.5									
724.387582	+2y6	16.47	31.0	<b>0.2041</b>	0.0265	13.0										
381.213246	+2y3	16.47	33.0	<b>0.181</b>	0.0120	6.6										
IFT43	YSLATSR	399.211235	634.351865	+2y6	20.45	21.2	<b>0.2602</b>	0.0086	3.3	<b>0.1734</b>	0.0457	26.4	<b>0.1664</b>	0.0581	35.0	<b>1.02</b>
			547.319837	+2y5	20.45	21.2	<b>0.1663</b>	0.0054	3.2							
			434.235772	+2y4	20.45	20.2	<b>0.1631</b>	0.0026	1.6							
	ASEEIEDFR	548.251285	363.198659	+2y3	20.45	20.2	<b>0.1805</b>	0.0050	2.8	<b>0.0694</b>	0.0078	11.3				
	808.383559	+2y6	30.52	23.6	<b>0.073</b>	0.0038	5.2									
679.340966	+2y5	30.52	25.6	<b>0.0657</b>	0.0016	2.5										
566.256902	+2y4	30.52	24.6	<b>0.0652</b>	0.0038	5.8										
437.214309	+2y3	30.52	27.6	<b>0.0819</b>	0.0050	6.1										
VLAPEHEVR	525.290548	837.421342	+2y7	16.38	25.8	<b>0.1492</b>	0.0052	3.5	<b>0.1664</b>	0.0187	11.2					
		766.384228	+2y6	16.38	26.8	<b>0.1638</b>	0.0030	1.8								
		669.331464	+2y5	16.38	32.8	<b>0.1941</b>	0.0066	3.4								
		540.288871	+2y4	16.38	32.8	<b>0.1689</b>	0.0091	5.4								
		777.388979	+2y7	28.44	21.8	<b>0.1701</b>	0.0090	5.3								
YGSDPVFR	470.729792	720.367515	+2y6	28.44	21.8	<b>0.1588</b>	0.0009	0.6	<b>0.1632</b>	0.0047	2.9					
		633.335487	+2y5	28.44	21.8	<b>0.1629</b>	0.0179	11.0								
		518.308544	+2y4	28.44	28.8	<b>0.1634</b>	0.0025	1.5								
		803.425758	+2y7	23.56	24.9	<b>0.2465</b>	0.0310	12.6								
		690.341694	+2y6	23.56	23.9	<b>0.177</b>	0.0026	1.5								
EAILES DAR	502.256371	577.25763	+2y5	23.56	23.9	<b>0.1792</b>	0.0020	1.1	<b>0.1849</b>	0.0327	17.7					
		448.215037	+2y4	23.56	21.9	<b>0.1906</b>	0.0008	0.4								
		615.309666	+2y5	18.61	19.0	<b>0.181</b>	0.0175	9.7								
		486.267073	+2y4	18.61	18.0	<b>0.1609</b>	0.0029	1.8								
		371.24013	+2y3	18.61	25.0	<b>0.1549</b>	0.0044	2.9								
457.22929	+2b4	18.61	17.0	-	-	-										
IFSLLEK	425.257654	736.423967	+2y6	42.39	20.2	<b>0.4188</b>	0.0087	2.1	<b>0.4635</b>	0.2521	54.4	<b>0.353225</b>	0.1560	44.2	<b>2.165032</b>	
		589.355553	+2y5	42.39	20.2	<b>0.9484</b>	0.0827	8.7								
		502.323525	+2y4	42.39	26.2	<b>0.4324</b>	0.0905	20.9								
		389.239461	+2y3	42.39	28.2	<b>0.4946</b>	0.0245	4.9								
		708.403901	+2y6	21.65	19.9	<b>0.3227</b>	0.0388	12.0								
AASVYIR	390.224145	637.366787	+2y5	21.65	19.9	<b>0.2398</b>	0.0088	3.7	<b>0.2430</b>	0.0425	17.5					
		550.334758	+2y4	21.65	20.9	<b>0.2461</b>	0.0032	1.3								
		451.266344	+2y3	21.65	20.9	<b>0.2302</b>	0.0059	2.6								
		642.429721	+2y6	38.13	22.5	-	-	-								
DLAIGLR	379.23197	529.345657	+2y5	38.13	18.5	<b>0.2222</b>	0.0118	5.3	<b>0.2337</b>	0.0086	3.7	<b>0.22305</b>	0.0151	6.8	<b>1.367147</b>	
		458.308544	+2y4	38.13	18.5	<b>0.2337</b>	0.0067	2.9								
		345.22448	+2y3	38.13	17.5	<b>0.239</b>	0.0064	2.7								
		797.451579	+2y7	18.88	24.2	<b>0.1976</b>	0.0297	15.0								
		726.414465	+2y6	18.88	24.2	<b>0.2181</b>	0.0065	3.0								
YASHLLEK	480.761092	502.323525	+2y4	18.88	29.2	<b>0.2067</b>	0.0348	16.9	<b>0.2124</b>	0.0279	13.1					
		389.239461	+2y3	18.88	31.2	<b>0.2607</b>	0.0038	1.4								

**Table 18: Absolute quantification of IFT-A using Flp-in (N)-SF-TULP3 performing PRM**

Protein	Peptide	Precursor m/z	Fragment Ion	RT	Fragment Ion	Median <sub>Pep</sub>	SD <sub>Pep</sub>	Ratio H/L				
								%CV <sub>Pep</sub>	Median <sub>Prot</sub>	SD <sub>Prot</sub>	%CV <sub>Prot</sub>	Stoich.
IFT122	DSIGDEDPFTAK	651.800614	y10	53.71	0.8024	0.7387	0.0365	4.9	0.7074	0.0894	12.6	1.80
			y6		0.7132							
			y7		0.7132							
			y8		0.7434							
	DTPSGISK	406.715624	21.69	y9	0.7387							
				y4	0.3414							
				y5	0.3019							
				y6	0.4232							
	GEYILLGGSDK	580.307878	51.66	y7	0.6749							
				y8	0.6761							
				y9	0.7025							
				y4	0.9516							
ILFTLAK	407.269839	54.63	y5	1.3372								
			y6	0.8298								
			y6	1.1704								
			y7	0.9332								
WDEAFALGEK	587.286945	57.59	y8	0.9653								
			y9	1.0574								
			y6	0.7501								
			y7	0.5174								
YLELISSIEER	681.362959	67.98	y8	0.5831								
			y9	0.7258								
			y5	0.7364								
			y6	0.6396								
IFT140	AHGALTEAYK	534.78183	y7	0.6606								
			y8	0.6851								
			y4	0.7459								
			y5	0.7327								
FWDIER	438.217912	51.97	y6	0.5854								
			y6	0.4631								
			y7	0.4969								
			y8	0.4969								
SHLFVDEGLK	576.810587	42.42	y4	0.2023								
			y5	0.1898								
			y6	0.1918								
			y7	0.1683								
IFT43	ASEEIEDFR	553.25542	y8	0.1900								
			y4	0.3371								
			y5	0.4550								
			y6	0.3959								
	VLAPEHEVR	530.294682	26.44	y7	0.3723							
				y8	0.4796							
				y4	0.5853							
				y5	0.6253							
	YSLATSR	404.215369	28.77	y6	0.4938							
				y5	0.5960							
				y6	0.6870							
				y7	0.6347							
TTC21B	EAILESDAR	507.260505	y3	0.4036								
			y4	0.4977								
			y5	0.4429								
			y4	0.6111								
LEDVPR	369.704637	27.9	y5	0.6012								
			y6	0.5340								
			y7	0.5474								
			y3	0.6282								
YGSDPVFR	475.733926	38.27	y4	0.6703								
			y5	0.7151								
			y6	0.4904								
			y6	0.1573								
WDR19	AASVYIR	395.22828	y7	0.1592								
			y8	0.1408								
			y4	0.3657								
			y5	0.3210								
	DGDVLAVIAEK	569.315703	63.31	y6	0.4165							
				y4	0.6151							
				y5	0.8795							
				y6	0.2704							
	IFSLEK	429.264753	55.7	y6	0.6379							
				y7	0.4778							
				y6	0.7080							
				y7	0.6302							
LVFIDEK	436.254385	48.6	y8	0.5403								
			y9	0.5115								
			y10	1.3723								
			y6	1.0876								
VGDLLPHVSSPK	628.858069	44.1	y7	0.9379								
			y8	1.0947								
			y9	1.2419								
			y4	0.6749								
VGSFLAVGTVK	543.325873	52.86	y5	0.6136								
			y6	0.5610								
			y5	0.5268								
			y6	0.4588								
WDR35	EIGLLAR	434.759944	y4	0.4968								
			y5	0.5081								
			y6	0.5681								
			y7	0.5806								
THVIAASK	417.749802	17.93	y4	0.4968								
			y5	0.5081								
			y6	0.5681								
			y7	0.5806								
YASHLLEK	484.768191	28.42	y4	0.4968								
			y5	0.5081								
			y6	0.5681								
			y7	0.5806								



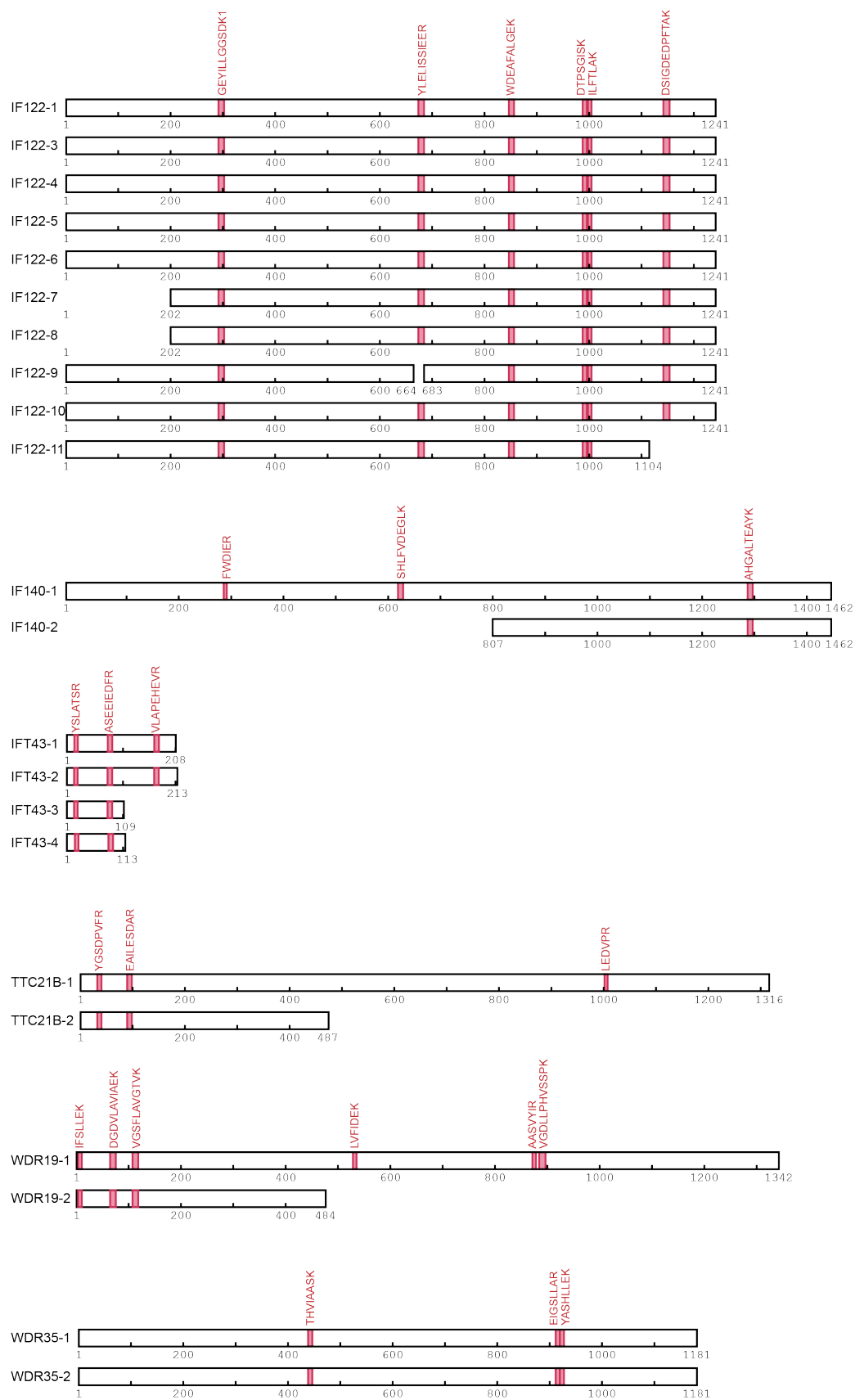


Figure 48: Scheme of known isoforms of IFT-A components

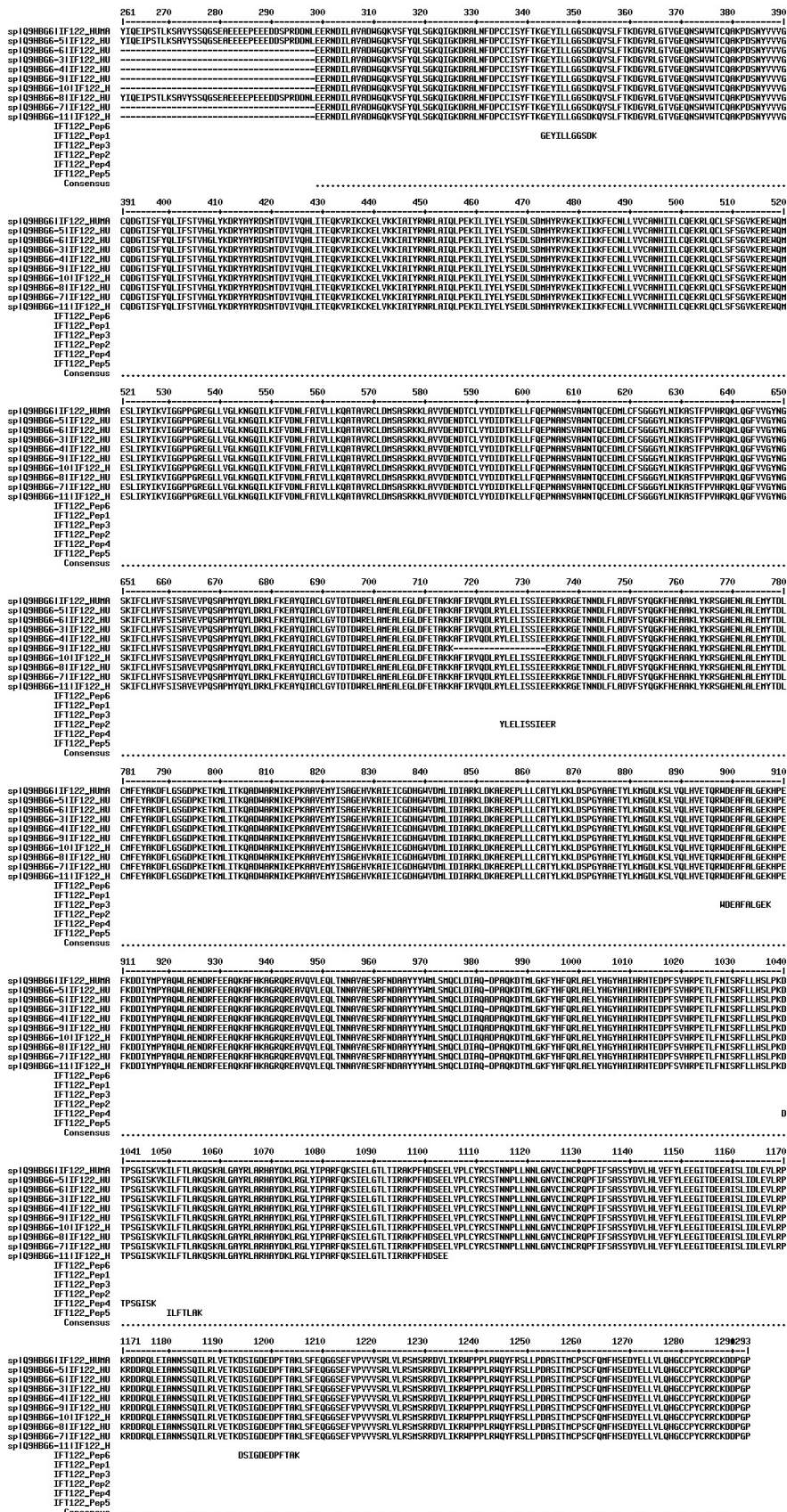


Figure 49: Alignment of IFT122 isoforms and chosen representative peptides

```

1      10      20      30      40      50      60      70      80      90      100     110     120     130
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
131    140    150    160    170    180    190    200    210    220    230    240    250    260
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
261    270    280    290    300    310    320    330    340    350    360    370    380    390
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
                                FMDIER
.....
391    400    410    420    430    440    450    460    470    480    490    500    510    520
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
521    530    540    550    560    570    580    590    600    610    620    630    640    650
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
                                SHLFVDEGLK
.....
651    660    670    680    690    700    710    720    730    740    750    760    770    780
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
781    790    800    810    820    830    840    850    860    870    880    890    900    910
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
911    920    930    940    950    960    970    980    990    1000   1010   1020   1030   1040
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
1041  1050  1060  1070  1080  1090  1100  1110  1120  1130  1140  1150  1160  1170
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
1171  1180  1190  1200  1210  1220  1230  1240  1250  1260  1270  1280  1290  1300
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
1301  1310  1320  1330  1340  1350  1360  1370  1380  1390  1400  1410  1420  1430
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....
1431  1440  1450  1462
spIQ96RY7|IF140_HUMA |-----|
spIQ96RY7-2|IF140_HU |-----|
IFT140_Pep3
IFT140_Pep2
IFT140_Pep1
Consensus
.....

```

Figure 50: Alignment of IFT140 isoforms and chosen representative peptides

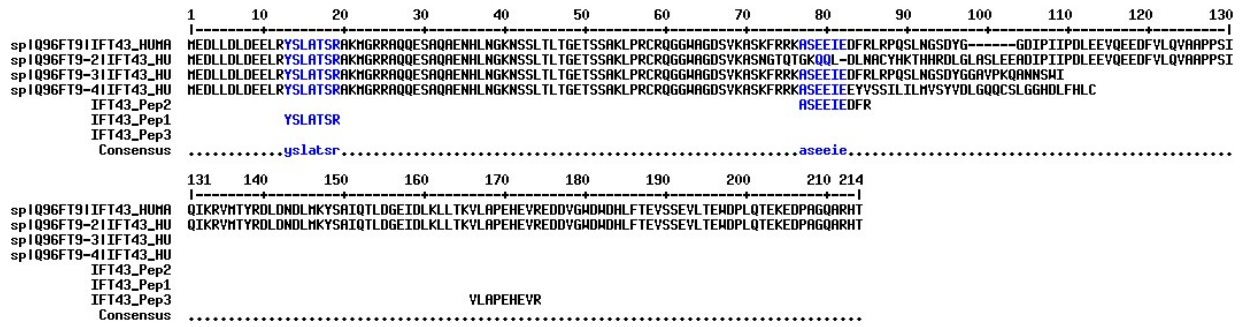


Figure 51: Alignment of IFT43 isoforms and chosen representative peptides

```

1      10     20     30     40     50     60     70     80     90     100    110    120    130
sp1Q7Z4L5|TT21B_HUMA MDSEQLKTLINYYCQERYFHHVLLVASEGKRYGSDPVFRFYHRYGTLMEGKTQEQALREFEAIKKNQDVSLSCLLALIIYAHKMSPNPDREAITLESQARVKEQRKGAGEKALYHAGFLWHIGRHDKAREY
sp1Q7Z4L5-2|TT21B_HU MDSEQLKTLINYYCQERYFHHVLLVASEGKRYGSDPVFRFYHRYGTLMEGKTQEQALREFEAIKKNQDVSLSCLLALIIYAHKMSPNPDREAITLESQARVKEQRKGAGEKALYHAGFLWHIGRHDKAREY
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
131    140    150    160    170    180    190    200    210    220    230    240    250    260
sp1Q7Z4L5|TT21B_HUMA IDRHKISDGSQGHVLAHLDIRGKEPYTKKALKYFEELQDGNDFALLGKAQCLEMRQNSGALETYNQIIVNFPFSLPAFVKKMKLQLALQDQDQTVETAQRLLLDQSQNVEALRHQALYYVCRE
sp1Q7Z4L5-2|TT21B_HU IDRHKISDGSQGHVLAHLDIRGKEPYTKKALKYFEELQDGNDFALLGKAQCLEMRQNSGALETYNQIIVNFPFSLPAFVKKMKLQLALQDQDQTVETAQRLLLDQSQNVEALRHQALYYVCRE
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
261    270    280    290    300    310    320    330    340    350    360    370    380    390
sp1Q7Z4L5|TT21B_HUMA GDIEKASTKLEMLGNTLDAMEPQNAQLFYNIILAFSRTCGRSGLILQKIQTLLERAFSLNPQQSEFATELGYQMIILQGRVKEALKAYKTAMTLDTSYSALVGFVQCLIEGQLQDADQQLLEFLNETIQS
sp1Q7Z4L5-2|TT21B_HU GDIEKASTKLEMLGNTLDAMEPQNAQLFYNIILAFSRTCGRSGLILQKIQTLLERAFSLNPQQSEFATELGYQMIILQGRVKEALKAYKTAMTLDTSYSALVGFVQCLIEGQLQDADQQLLEFLNETIQS
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
391    400    410    420    430    440    450    460    470    480    490    500    510    520
sp1Q7Z4L5|TT21B_HUMA IGKSARELIYLAHVLAHKKKRQEEVINLLNDVLDTHFSQLEGLPLGIQYFEKLNPDFLLEIYMEYLSFCPHQSPGQPLCPLLRRCISVLETVVRYTPGLLQTVFLIARVYKLSGDIARFNNLQKCLE
sp1Q7Z4L5-2|TT21B_HU IGKSARELIYLAHVLAHKKKRQEEVINLLNDVLDTHFSQLEGLPLGIQYFEKLNPDFLLEIYMEYLSFCPHQSPGQPLCPLLRRCISVLETVVRYTPGLLQTVFLIARVYKLSGDIARFNNLQKCLE
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
521    530    540    550    560    570    580    590    600    610    620    630    640    650
sp1Q7Z4L5|TT21B_HUMA HNPSYADAHLLLAQVYLSQEKYKLCQSLELCLSYDFKYRDPYLYHLIKAQSQKNGEIDAIKTLHHAHSLPGHKRIGASTSKORKTEYDTSRLSIFLELDIVHRLNGEQHEATKYLQDAIHEFSGT
sp1Q7Z4L5-2|TT21B_HU HNPSYADAHLLLAQVYLSQEKYKLCQSLELCLSYDFKYRDPYLYHLIKAQSQKNGEIDAIKTLHHAHSLPGHKRIGASTSKORKTEYDTSRLSIFLELDIVHRLNGEQHEATKYLQDAIHEFSGT
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
651    660    670    680    690    700    710    720    730    740    750    760    770    780
sp1Q7Z4L5|TT21B_HUMA SEEVRYTIANADLALAQDIERALSILQNVYTAEQPYFEAREKHADIYLAHKKRKHLYITCFREIARHMANPFRSLLLDGAYMNIPEEAIIVAYEQALNQMPKDGTLASKGKALIKTHNYSRAITYYE
sp1Q7Z4L5-2|TT21B_HU SEEVRYTIANADLALAQDIERALSILQNVYTAEQPYFEAREKHADIYLAHKKRKHLYITCFREIARHMANPFRSLLLDGAYMNIPEEAIIVAYEQALNQMPKDGTLASKGKALIKTHNYSRAITYYE
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
781    790    800    810    820    830    840    850    860    870    880    890    900    910
sp1Q7Z4L5|TT21B_HUMA AALKTGQKNYLCYDLAELLKLYDKREKVLQHALAHEPYNELSALNEDGRCQVLLAKYYSKMEKLDGIDITALQQARELQARVLRKRYQMEQPDAYPAQKHLAELCAEIAKHSVAQDQYKRIKIFYREA
sp1Q7Z4L5-2|TT21B_HU AALKTGQKNYLCYDLAELLKLYDKREKVLQHALAHEPYNELSALNEDGRCQVLLAKYYSKMEKLDGIDITALQQARELQARVLRKRYQMEQPDAYPAQKHLAELCAEIAKHSVAQDQYKRIKIFYREA
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
911    920    930    940    950    960    970    980    990    1000   1010   1020   1030   1040
sp1Q7Z4L5|TT21B_HUMA LVHCETONKIHLELARLYLAQDDPDSCLRQCALLLQSDQDNEARATHMAADLHFRKQDYEQAVFHLQQLLERKPDNYHTLSRLIDLLRRCGLKEDVPRFFSHAEKRNRAKLEPGFYCKGLYLYHTGEPH
sp1Q7Z4L5-2|TT21B_HU LVHCETONKIHLELARLYLAQDDPDSCLRQCALLLQSDQDNEARATHMAADLHFRKQDYEQAVFHLQQLLERKPDNYHTLSRLIDLLRRCGLKEDVPRFFSHAEKRNRAKLEPGFYCKGLYLYHTGEPH
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
1041  1050   1060   1070   1080   1090   1100   1110   1120   1130   1140   1150   1160   1170
sp1Q7Z4L5|TT21B_HUMA DALRHFNKARKDRDHGQNALYNIIEICLNPDNETYGGVEFENLDGDLGNSTKEQESVQLAVRTAEKLLKELKPQTVQGHVQLRIMENYCLMATKQKSNVEQALNTFTEIARSEKEHLPALLGHATAYMIL
sp1Q7Z4L5-2|TT21B_HU DALRHFNKARKDRDHGQNALYNIIEICLNPDNETYGGVEFENLDGDLGNSTKEQESVQLAVRTAEKLLKELKPQTVQGHVQLRIMENYCLMATKQKSNVEQALNTFTEIARSEKEHLPALLGHATAYMIL
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
1171  1180   1190   1200   1210   1220   1230   1240   1250   1260   1270   1280   1290   1300
sp1Q7Z4L5|TT21B_HUMA KQTPRARNLKRIAKHNNNAIDAEFEKSLLLADIYIQSAKYDHAEDLLKRCRLHNRSCCKAYEYHGYTHEKEQAYDAALNYENAAKYSNRTPAVGYKLAFFNYLAKKRYDSDICHQVLEAHPTYP
sp1Q7Z4L5-2|TT21B_HU KQTPRARNLKRIAKHNNNAIDAEFEKSLLLADIYIQSAKYDHAEDLLKRCRLHNRSCCKAYEYHGYTHEKEQAYDAALNYENAAKYSNRTPAVGYKLAFFNYLAKKRYDSDICHQVLEAHPTYP
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....
1301  1310   1316
sp1Q7Z4L5|TT21B_HUMA KIRKQILDKARRSLRP
sp1Q7Z4L5-2|TT21B_HU
TTC21B_Pep2
TTC21B_Pep1
TTC21B_Pep3
Consensus
.....

```

Figure 52: Alignment of TTC21B isoforms and chosen representative peptides

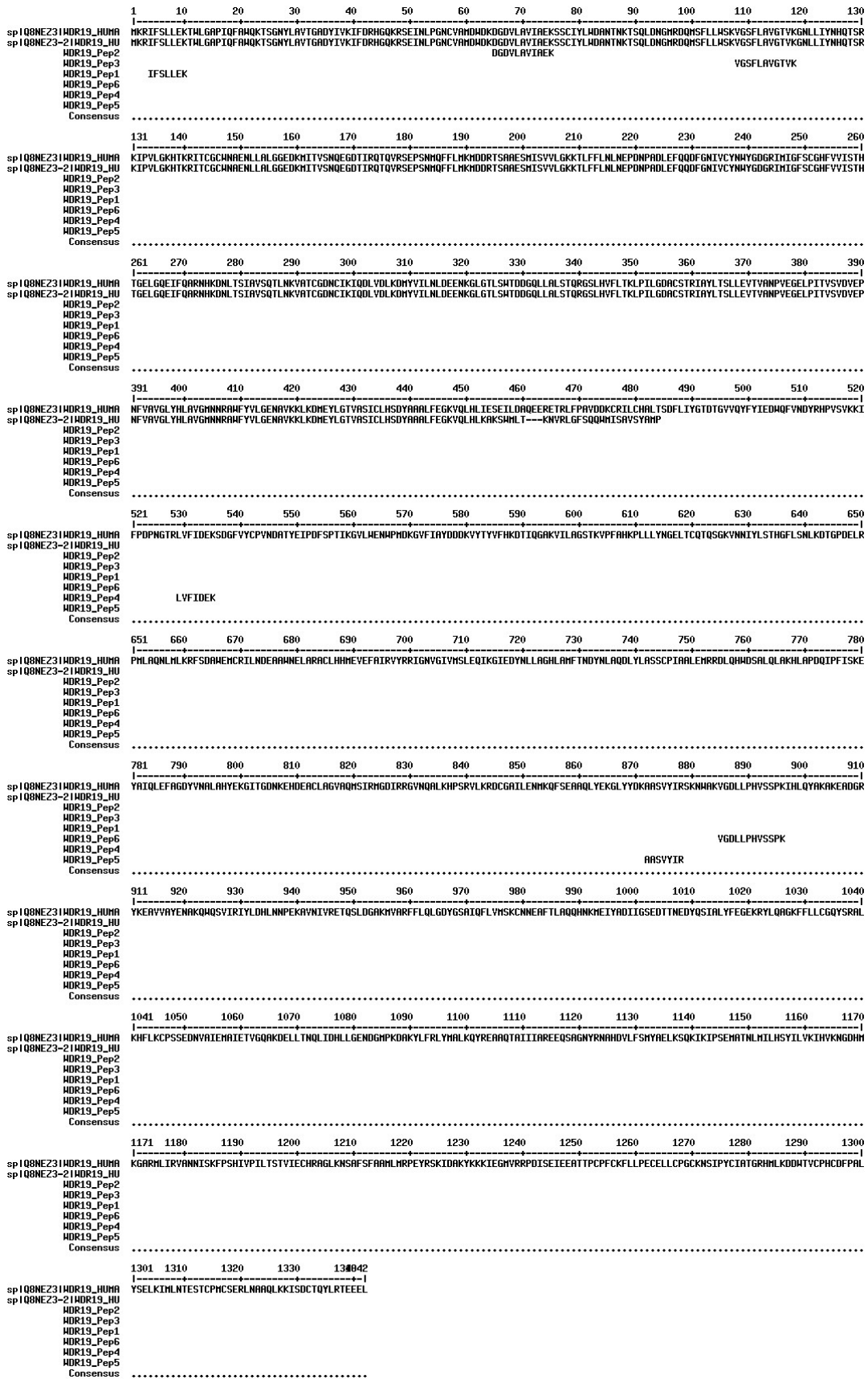


Figure 53: Alignment of WDR19 isoforms and chosen representative peptides

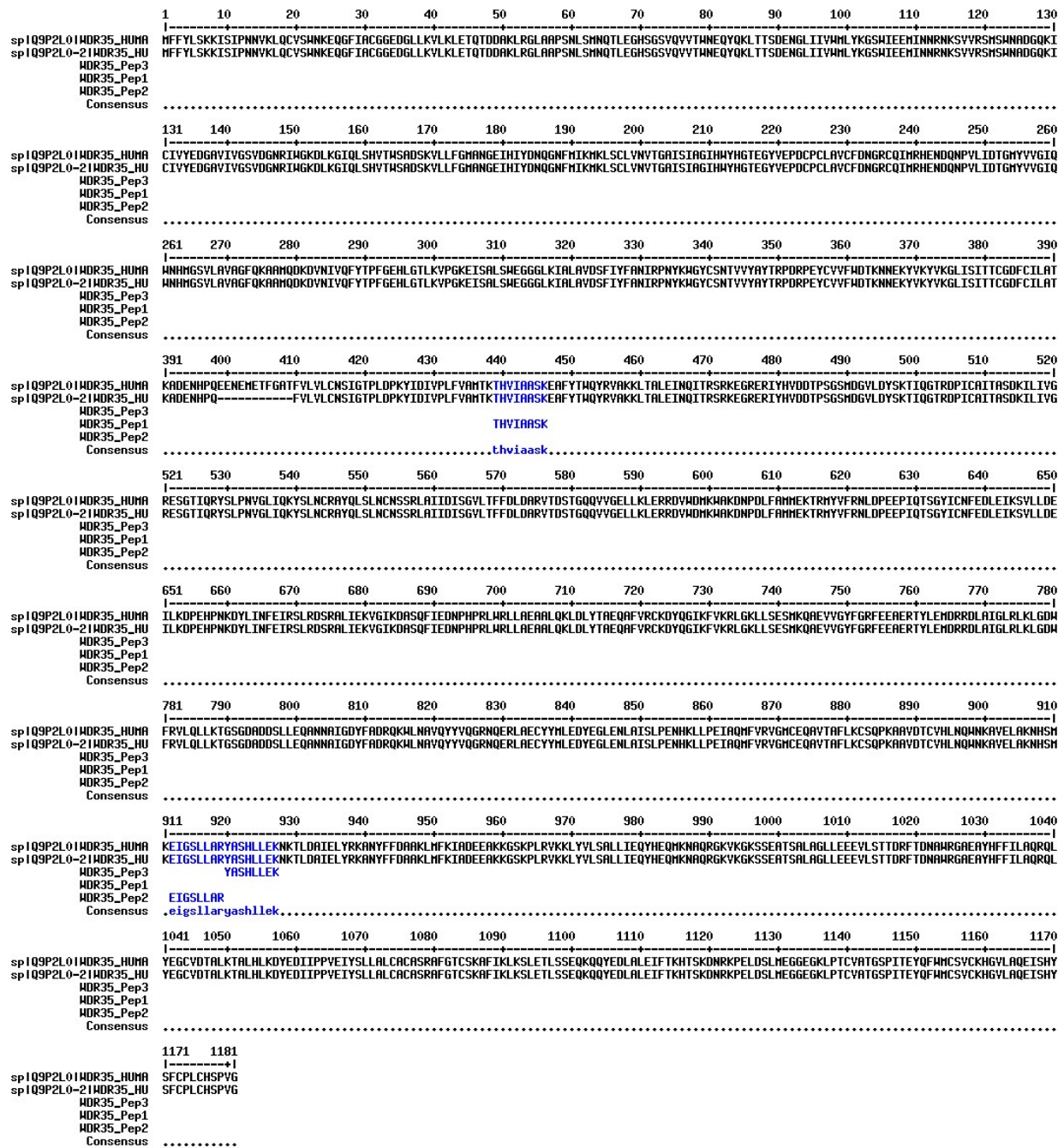


Figure 54: Alignment of WDR35 isoforms and chosen representative peptides

Table 19: Identified links of IFT-A from Flp-In (N)-SF-TULP3 cells after SEC (in at least 2of3)

Id	Protein1	Protein2	XLType	Biological Replicates	AbsPos1	AbsPos2	Mr
ALIKTHNSMAITYEAALK-K4-156	TT21B		monolink	YWI1931/YWI1936	768	n/a	2456.256
AYEYMGYIMEKEQAYTDAALINYEAWK-K27-156	TT21B		monolink	YWI1931/YWI1936	1259	n/a	3453.539
EKMADYLKHR-K2-155	TT21B		monolink	YWI1931/YWI1936	693	n/a	1557.828
EKMADYLKHR-K9-155	TT21B		monolink	YWI1931/YWI1936/YWI1953	700	n/a	1557.828
FFSMAEKRSR-K7-156	TT21B		monolink	YWI1931/YWI1936/YWI1953	1014	n/a	1527.762
MKLQLALQDWDQVETAQR-K2-155	TT21B		monolink	YWI1936/YWI1953	220	n/a	2428.226
XDSQELKTLINVCQER-K7-155	TT21B		monolink	YWI1936/YWI1953	7	n/a	2361.086
ALIKYFEEGLQDNDTFALLGK-HNRSCKKAYEYMGYIMEK-a3-b7	TT21B	TT21B	intra-protein xl	YWI1931/YWI1936	166	1232	4805.254
DGTLASKMGK-CGKLEDPVR-a7-b3	TT21B	TT21B	intra-protein xl	YWI1936/YWI1953	761	1001	2217.101
FFSMAEKRSR-GKEPYTK-a7-b2	TT21B	TT21B	intra-protein xl	YWI1936/YWI1953	1014	157	2331.161
QCALLQSDQNEAATMMADLMFRKQDYEQAVFHQLQLLER-DKMLYTCFR-a26-b2	TT21B	TT21B	intra-protein xl	YWI1936/YWI1953	965	705	6512.052
SCCKAYEYMGYIMEKEQAYTDAALINYEAWK-IRKDILDK-a4-b3	TT21B	TT21B	intra-protein xl	YWI1931/YWI1953	1232	1304	4954.263
TLHMAMSLPGXKR-DILDKAR-a12-b5	TT21B	TT21B	intra-protein xl	YWI1931/YWI1953	596	1309	2455.284
VKLCQSLELCLSDFKVR-HNRSCKKAYEYMGYIMEK-a17-b7	TT21B	TT21B	intra-protein xl	YWI1931/YWI1953	558	1232	4821.262
VYSKMEK-VYSKMEK-a4-b4	TT21B	TT21B	intra-protein xl	YWI1936/YWI1953	843	843	1904.955
DKMLYTCFR-KIEGXVR-a2-b1	TT21B	WDR19	inter-protein xl	YWI1931/YWI1936	705	1233	2331.156
DRKTEVDTSHR-KKIEGMVR-a3-b1	TT21B	WDR19	inter-protein xl	YWI1931/YWI1953	608	1232	2440.263
EQRKGAGEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	104	1171	2124.059
KPDNYMTLR-AKEADGR-a1-b2	TT21B	WDR19	inter-protein xl	YWI1931/YWI1936	982	905	2107.03
ALIKYFEEGLQDNDTFALLGKAOCLXR-KYTOAGNKLK-a3-b1	TT21B	IF140	inter-protein xl	YWI1936/YWI1953	166	1211	4520.283
DKMLYTCFR-KPEEADR-a2-b1	TT21B	IF140	inter-protein xl	YWI1931/YWI1953	705	733	2327.111
EKMADYLK-KPEEADR-a2-b1	TT21B	IF140	inter-protein xl	YWI1936/YWI1953	693	733	2091.043
IMENYCLMATKOK-YVEEKGVQXDR-a11-b5	TT21B	IF140	inter-protein xl	YWI1936/YWI1953	1134	1077	3199.44
LAFNYLAK-AEEVXVK-a7-b6	TT21B	IF140	inter-protein xl	YWI1931/YWI1936	1278	270	2153.194
VKYLSDIEAFAFNLOHCLHNPYSADAHLLAQVLSQEK-GENXNCVCYKVK-a2-b11	TT21B	IF140	inter-protein xl	YWI1931/YWI1953	502	334	6513.051
YDMAEDLKR-KEPEIKK-a9-b1	TT21B	IF140	inter-protein xl	YWI1936/YWI1953	1221	1261	2280.116
AKLEGFQYCK-ETKMLTK-a2-b3	TT21B	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	1020	748	2440.268
AOCLERQNSGALETVNIQIIVNFFSFLPAFVKK-GHKDYYCVAYAK-a33-b3	TT21B	IF122	inter-protein xl	YWI1936/YWI1953	217	53	5560.852
KXGEIADAIK-NKNGEEK-a1-b2	TT21B	IF122	inter-protein xl	YWI1931/YWI1936	575	161	2046.024
SCCKAYEYMGYIMEK-NIKEPK-a4-b3	TT21B	IF122	inter-protein xl	YWI1931/YWI1936	1232	762	2813.283
AIKFYREALVHCETDNK-DNPDLFAMMEKTR-a3-b11	TT21B	WDR35	inter-protein xl	YWI1936/YWI1953	905	614	3797.802
AYEYMGYIMEKEQAYTDAALINYEAWK-DNPDLFAMMEKTR-a11-b11	TT21B	WDR35	inter-protein xl	YWI1931/YWI1953	1243	614	5002.248
EALVHCETDNKXIEELARLYAQDDPSCLR-SMSWADGQKICIVYEDGAVIGSVGNGNR-a11-b10	TT21B	WDR35	inter-protein xl	YWI1936/YWI1953	919	129	6981.31
KDRDWGQNALYNXIECLNPDNETVGGVEFENLDGDLGNSTEK-KPELDSLMIEGGEK-a1-b1	TT21B	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1051	1123	6512.035
WYDKAEVLQHALAHEPYNELSAIMEDGR-LKLDWFR-a7-b2	TT21B	WDR35	inter-protein xl	YWI1931/YWI1953	810	776	4520.283
YDMAEDLKR-MFFYLSKK-a9-b7	TT21B	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1221	7	2469.221
MDDRTSAEESISVILGKK-K19-156	WDR19		monolink	YWI1931/YWI1953	209	n/a	2209.098
VATCGDNCKIQDLVDLK-K10-156	WDR19		monolink	YWI1931/YWI1936/YWI1953	299	n/a	2217.098
VATCGDNCKIQDLVDLK-K18-156	WDR19		monolink	YWI1931/YWI1953	307	n/a	2217.102
DDWTVCPHDFPALYSELKIMLINTESTCPMCSERLNAAQLK-KDKMLYTCFR-a19-b1	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	1305	703	6512.037
GITGDNKEHDEACLAGVAQXSIR-IGASTKSK-a7-b6	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	807	603	3415.665
GVIHAYDDKVVYVFKHDITQGAQK-KDKMLYTCFR-a10-b1	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	580	703	4520.287
HVLMWPMDKGVFIAVDDKVVYVFKH-ALYFEEGLQDNDTFALLGKAOCLER-a21-b3	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	580	166	6889.327
HMLKDDWTVCPHDFPALYSELKIMLINTESTCPMCSER-SAEIYLHVLAXKK-a23-b14	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	1305	407	6511.047
HMLKDDWTVCPHDFPALYSELKIMLINTESTCPMCSER-SAEIYLHVLAXKK-a4-b14	WDR19	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	1286	407	6511.033
HMLKDDWTVCPHDFPALYSELKIMLINTESTCPXCSR-MIKISDGSKQGHVLLK-a4-b3	WDR19	TT21B	inter-protein xl	YWI1931/YWI1953	1286	136	6465.028
KTLFLNLEPNADLEFQQDFGNVYVWYVGDGR-FYHAYGTLMEFGTQELAR-a1-b12	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	209	52	6561.015



KTLEFLINLPDPADLEFQDDFNINVCYWGDR-RIAKMNVNNAIDAEEFEK-a1-b4	WDR19	TT21B	inter-protein xl	YWI1936/YWI1953	209	1185	6511.033
YKEAVVAYENAK-KMGEIADA1K-a2-b1	WDR19	TT21B	inter-protein xl	YWI1931/YWI1953	912	575	2596.345
CNNEAFTLAQQHNKMEIYADIGSEDTTINEDYQSIALYFEGEKR-KISDCTQYL-R-a14-b1	WDR19	WDR19	intra-protein xl	YWI1936/YWI1953	992	1328	6576.038
HMLKDDWTVCPHCFDPALYSEKIMLINTESTCPMCSER-DAKYLFRLYMALK-a23-b3	WDR19	WDR19	intra-protein xl	YWI1931/YWI1936/YWI1953	1305	1090	6439.994
HMLKDDWTVCPHCFDPALYSEKIMLINTESTCPMCSER-DAKYLFRLYMALK-a4-b3	WDR19	WDR19	intra-protein xl	YWI1936/YWI1953	1286	1090	6439.997
AGLKNFAFAAMLMRPEYSK-AVSGDEKALDMFNWKK-a4-b16	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1207	187	4521.29
ALKHFLKPSSEEDNVAIEMAIETVGOAK-GENVILSPDEKFGKGENNVCYCK-a3-b16	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1041	323	6511.006
GITGNKHDEACLAGVAQMSIRMGDIR-KYQEAQLCLGQNSITEEMAEKMTVAK-a7-b23	WDR19	IF140	inter-protein xl	YWI1931/YWI1936	807	1173	6440.999
LNAAQLKK-KEPEIXK-a7-b1	WDR19	IF140	inter-protein xl	YWI1931/YWI1936	1327	1261	1912.077
NGDHMKGAR-KEPEIMK-a6-b1	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1171	1261	1995.969
NSAFSAAMLIMRPEYRSKIDAK-FGFEKGENXNVCYCKV-a18-b16	WDR19	IF140	inter-protein xl	YWI1931/YWI1953	1225	334	4955.261
NSAFSAAMLIMRPEYRSKIDAK-FGFEKGENXNVCYCKV-a18-b5	WDR19	IF140	inter-protein xl	YWI1931/YWI1953	1225	323	4955.256
SSCIYLDANTKTSQLDNGXRDMISLWSK-GKTTQVVSADSTIQMLFYXKR-a13-b21	WDR19	IF140	inter-protein xl	YWI1931/YWI1953	88	241	6511.033
DCGAILENXKOFSEAAQLYKGLYDK-GETNNDLFLADVFSYQGFHEAAKLYK-a10-b24	WDR19	IF122	inter-protein xl	YWI1936/YWI1953	854	711	6442.02
IDAKYK-NIKEPK-a4-b3	WDR19	IF122	inter-protein xl	YWI1931/YWI1953	1229	762	1601.913
RITCGWNAENLALGGEDKMITVSNQEGDIR-HPEFKDDIYXPAQWLAENDR-a20-b5	WDR19	IF122	inter-protein xl	YWI1931/YWI1953	160	861	6512.035
CNNEAFTLAQQHNKMEIYADIGSEDTTINEDYQSIALYFEGEKR-KANVFFDAK-a14-b1	WDR19	WDR35	inter-protein xl	YWI1936/YWI1953	992	939	6466.935
CNNEAFTLAQQHNKXEIYADIGSEDTTINEDYQSIALYFEGEKR-YASHLLEKKNK-a43-b8	WDR19	WDR35	inter-protein xl	YWI1931/YWI1936	1021	927	6511.047
DDWTVCPHCFDPALYSEKIMLINTESTCPXCSER-LGKLLSESMKQAEVYGFGR-a19-b3	WDR19	WDR35	inter-protein xl	YWI1931/YWI1953	1305	736	6527.026
HMLKDDWTVCPHCFDPALYSEKIMLINTESTCPXCSER-VLLETQTDADAKLR-a4-b12	WDR19	WDR35	inter-protein xl	YWI1936/YWI1953	1286	51	6454.022
HMLKDDWTVCPHCFDPALYSEKIMLINTESTCPXCSER-VLLETQTDADAKLR-a4-b3	WDR19	WDR35	inter-protein xl	YWI1931/YWI1936	1286	42	6454.006
IEGWVRRDPISEIATTPCFKFLPECELLCPGCK-NKSVRSXSNWADGOK-a24-b2	WDR19	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1257	115	6511.042
NSIPYCIATGRHMLKDDWTVCPHCFDPALYSELK-DNPDFAMIMEKTRXYVFR-a15-b11	WDR19	WDR35	inter-protein xl	YWI1931/YWI1953	1286	614	6511.006
RKASEIEDFR-K2-156	IFT43		monolink	YWI1931/YWI1953	76	n/a	1534.766
EDDVGVWDHFLTEVSSEVLTEWDPLOTEKEDPAGQAR-DKMLYTCFREIAER-a30-b2	IFT43	TT21B	inter-protein xl	YWI1931/YWI1936	198	705	6511.033
EDDVGVWDHFLTEVSSEVLTEWDPLOTEKEDPAGQARHT-RLQCLFSFGVKEK-a30-b11	IFT43	IF122	inter-protein xl	YWI1936/YWI1953	198	463	6383.979
VMTYRDLDNDLMKYSAGIQLDGEIDLK-DDIYMPYAQWLAENDRFEAQKAFHK-a13-b22	IFT43	IF122	inter-protein xl	YWI1931/YWI1936	141	883	6512.045
RAQOESAQENHLINGKNSSLLTGETSSAK-AVELAKNHSVK-a16-b6	IFT43	WDR35	inter-protein xl	YWI1936/YWI1953	40	906	4521.262
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	YWI1931/YWI1936/YWI1953	187	n/a	2065.017
AAVSGDEKALDMFNWKK-K17-155	IF140		monolink	YWI1931/YWI1953	188	n/a	2064.034
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	YWI1931/YWI1936/YWI1953	188	n/a	2065.007
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	YWI1931/YWI1936/YWI1953	179	n/a	2065.006
AAVSGDEKALDMFNWKK-K8-155	IF140		monolink	YWI1936/YWI1953	179	n/a	1951.93
ALSYKSDTHR-K7-156	IF140		monolink	YWI1931/YWI1936/YWI1953	926	n/a	1624.767
DFVLEDCDKATR-K10-156	IF140		monolink	YWI1936/YWI1953	770	n/a	1680.773
DKTLWRWAQYLESQGEMLDAALHYELAR-K2-156	IF140		monolink	YWI1931/YWI1953	955	n/a	3784.769
GENMNCYCKYKGLLAAGTDR-K11-155	IF140		monolink	YWI1936/YWI1953	334	n/a	2670.249
GENMNCYCKYKGLLAAGTDR-K13-155	IF140		monolink	YWI1931/YWI1953	336	n/a	2670.216
GVQMDRAVMILYHKAGHFSK-K13-156	IF140		monolink	YWI1931/YWI1936/YWI1953	1090	n/a	2330.174
GVQMDRAVMILYHKAGHFSK-K19-156	IF140		monolink	YWI1936/YWI1953	1096	n/a	2330.167
KYQEAQLCLGQNSITEEMAEK-K23-155	IF140		monolink	YWI1931/YWI1953	1173	n/a	2868.349
KYQEAQLCLGQNSITEEMAEK-K1-155	IF140		monolink	YWI1931/YWI1953	1151	n/a	2884.346
RTYTEDPKSIK-K12-156	IF140		monolink	YWI1936/YWI1953	1358	n/a	1621.825
SEAVWENMARXCVKQTR-K14-155	IF140		monolink	YWI1936/YWI1953	813	n/a	2266.073
VPDFLGSFGAEGKDR-K13-155	IF140		monolink	YWI1931/YWI1936	366	n/a	1698.845
VQGTPLLKHEYGK-K8-156	IF140		monolink	YWI1931/YWI1936/YWI1953	145	n/a	1624.877
XTVAKDSDLPESR-K5-155	IF140		monolink	YWI1936/YWI1953	1178	n/a	1834.858
AAVSGDEKALDMFNWKK-XEKLGDALTAALQAR-a8-b3	IF140	TT21B	inter-protein xl	YWI1936/YWI1953	179	846	3594.747

AEEVXKVLGK-EKMADIYLK-a8-b2	IF140	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	272	693	2581.358
DFVLEDCDKATRAMLHFSFVTIGDXDEAFK-NYLVDLAEILLKLLKWYDK-a10-b15	IF140	TT21B	inter-protein xl	YWI1931/YWI1936	770	803	6441.014
DHFLSVRIHCFQGNVQKAAQJANETGNLAASYHLAR-HNRSCCKAYEYMGYIMEK-a17-b7	IF140	TT21B	inter-protein xl	YWI1936/YWI1953	999	1232	6513.046
ENGLDQQLXNLLALLSPEDMIEAARYEEKGVQMDR-EALVHCETDNKXILELAR-a30-b11	IF140	TT21B	inter-protein xl	YWI1931/YWI1936	1077	919	6453.984
ESIKQCEILLEPDLSTIRIGDYGVEFLVEHYVVR-HNRSCCKAYEYMGYIMEK-a4-b7	IF140	TT21B	inter-protein xl	YWI1936/YWI1953	1358	1232	6512.052
QGSYHLATTK-GKEPYTK-a9-b2	IF140	TT21B	inter-protein xl	YWI1931/YWI1953	1210	157	2091.097
SEAVWENXARMCKVQTR-DKMLYITCFREIAER-a14-b2	IF140	TT21B	inter-protein xl	YWI1936/YWI1953	813	705	4192.994
SSSGSLLKMGHEGLLFFVSLMDGTVHYVDEK-YEEGLQDGNDFALLGKAQCLEMR-a8-b18	IF140	TT21B	inter-protein xl	YWI1931/YWI1953	196	184	6512.061
YEEKGVQMDR-YDXAEDLLKR-a5-b9	IF140	TT21B	inter-protein xl	YWI1936/YWI1953	1077	1221	2823.292
YEEKGVQDR-EKXADIYLK-a5-b2	IF140	TT21B	inter-protein xl	YWI1931/YWI1953	1077	693	2696.25
ALDLAGFYDACAQVEIDYQNYDKAHGALTEAYKCLAK-SEPSNMQFLMKXDDR-a35-b12	IF140	WDR19	inter-protein xl	YWI1936/YWI1953	1312	190	6565.956
AVELLAAKYQALQLCLGQNMMSITEMAEK-SKNWAK-a10-b2	IF140	WDR19	inter-protein xl	YWI1936/YWI1953	1151	880	4520.287
CSGGSTISLQKADNSPDKFCYDVEIMDTVTYDFDK-NAHDVLFXYAEIKSQK-a14-b14	IF140	WDR19	inter-protein xl	YWI1936/YWI1953	590	1137	6480.992
DKTLWRWAAQYLESQGMDDAALHYVELAR-RDCGAILENKXQF5EAAQYK-a2-b11	IF140	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	955	854	6382.968
FGFEKGNMNCVYCK-NGDHMKGARMILIR-a5-b6	IF140	WDR19	inter-protein xl	YWI1936/YWI1953	323	1171	3677.646
FGFEKGNMNCVYCK-IHVKNGDHXKGAR-a5-b10	IF140	WDR19	inter-protein xl	YWI1931/YWI1953	323	1171	3673.641
GENYLSPEKFGFEKGNMNCVYCK-NGDHMKGARXLR-a16-b6	IF140	WDR19	inter-protein xl	YWI1931/YWI1953	323	1171	4955.238
GKTTQWVSADSTIQMLFYXK-EAVVAYENAKWQSVIR-a2-b10	IF140	WDR19	inter-protein xl	YWI1931/YWI1953	222	922	4520.273
KPEEADREDEVEFGCHHPQIMVSR-SEPSNMQFLMKDDR-a1-b12	IF140	WDR19	inter-protein xl	YWI1931/YWI1953	733	190	4957.224
RETLFNEQETNKSHLFVDEGLK-VLKRDCGAILENKX-a13-b3	IF140	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	634	843	4520.297
SEAVWENMARXCVKQTR-CROGGWAGDSKASK-a14-b12	IF140	IF143	inter-protein xl	YWI1931/YWI1953	813	69	3854.786
DFVLEDCDKATRAMLHFSFVTIGDMDEAFK-GKTTQWVSADSTIQMLFYMEK-a10-b2	IF140	IF140	intra-protein xl	YWI1931/YWI1936	770	222	6340.933
DFVLEDCDKATRAMLHFSFVTIGDMDEAFK-GKTTQWVSADSTIQMLFYMEK-a10-b21	IF140	IF140	intra-protein xl	YWI1931/YWI1953	770	241	6513.023
DKTLWRWAAQYLESQGMDDAALHYVELAR-KYQEAALQCLGQNMMSITEMAEK-a2-b1	IF140	IF140	intra-protein xl	YWI1931/YWI1936/YWI1953	955	1151	6480.026
ETLSFQETNKSHLFVDEGLK-FGFEKGNMNCVYCK-a12-b5	IF140	IF140	intra-protein xl	YWI1931/YWI1953	634	323	4987.251
GENYLSPEKFGFEKGNMNCVYCK-AHGALTEAYKCLAK-a16-b10	IF140	IF140	intra-protein xl	YWI1931/YWI1953	323	1312	4957.252
GENYLSPEKFGFEKGNMNCVYCK-LIKSEAVWENMMAR-a11-b3	IF140	IF140	intra-protein xl	YWI1931/YWI1953	318	799	4987.248
GKTTQWVSADSTIQMLFYXK-RFGFEKGNMNCVYCK-a21-b16	IF140	IF140	intra-protein xl	YWI1931/YWI1953	221	334	4971.266
GKTTQWVSADSTIQMLFYXK-RFGFEKGNMNCVYCK-a2-b5	IF140	IF140	intra-protein xl	YWI1931/YWI1953	222	323	4955.274
IGDYVFLVEHYRKEEYQYAYR-FGFEKGNMNCVYCK-a15-b5	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1389	323	5014.254
KYQEAALQCLGQNMMSITEMAEKXTVAK-YQEAALQCLGQNMMSITEMAEKXTVAK-a23-b22	IF140	IF140	intra-protein xl	YWI1931/YWI1936	1173	1173	6529.035
KYQEAALQCLGQNMMSITEMAEKXTVAK-YQEAALQCLGQNMMSITEMAEKXTVAK-a23-b22	IF140	IF140	intra-protein xl	YWI1936/YWI1953	1218	187	2455.276
KYQEAALQCLGQNMMSITEMAEKXTVAK-YQEAALQCLGQNMMSITEMAEKXTVAK-a23-b22	IF140	IF140	intra-protein xl	YWI1936/YWI1953	1047	336	6511.062
LCKENGLDQLMNLALLSPEDMIEAARYEEK-GENMNCVYCKVGLLAAGTDR-a3-b13	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1178	770	3482.61
MTVAKDSSDLPESRR-DFVLEDCDKATR-a5-b10	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1178	770	3482.61
MTVAKDSSDLPESRR-MTVAKDSSDLPESR-a5-b5	IF140	IF140	intra-protein xl	YWI1936/YWI1953	1178	1178	3621.742
QKEIYMAANYLQSLDWRK-FGFEKGNMNCVYCK-a2-b16	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1244	334	4792.253
TYTEDPKESIK-KYTOAGNLIK-a7-b8	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1354	1218	2597.348
TYTEDPKESIK-REKAHCSQR-a7-b4	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1354	554	2721.258
XKDKTLWR-KPEEADR-a2-b1	IF140	IF140	intra-protein xl	YWI1931/YWI1953	953	733	2074.034
XTVAKSSDLPESR-KEEYQYAYR-a5-b1	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1178	1389	3004.388
XTVAKSSDLPESR-MCVKQTR-a5-b4	IF140	IF140	intra-protein xl	YWI1931/YWI1953	1178	813	2739.267
YEEKGVQMDR-LKAMRALK-a5-b2	IF140	IF140	intra-protein xl	YWI1931/YWI1936/YWI1953	1077	1220	2597.358
AEEVMKVK-SVSKHK-a6-b4	IF140	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	270	128	1754.958
FGFEKGNMNCVYCK-CKELVK-a5-b6	IF140	IF122	inter-protein xl	YWI1931/YWI1953	323	392	3099.411
GENYLSPEKFGFEKGNMNCVYCK-XGDKSLVQLHVETQRWDEAFALGEK-a11-b5	IF140	IF122	inter-protein xl	YWI1931/YWI1936	318	835	6441.019
GENYLSPEKFGFEKGNMNCVYCK-XGDKSLVQLHVETQRWDEAFALGEK-a16-b5	IF140	IF122	inter-protein xl	YWI1931/YWI1936	323	835	6441.013
VQGTPLLKHEYGK-CKELVK-a8-b6	IF140	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	145	392	2510.373
YQEAALQCLGQNMMSITEMAEKXTVAKDSSDLPESR-DSXTDVIVQHLETEQKVR-a27-b16	IF140	IF122	inter-protein xl	YWI1931/YWI1936	1178	382	6512.053

YEEKGVQMDRAVXLYHK-CLDMSASRKK-a5-b9	IF140	IF122	inter-protein xl	YW11931/YW11936	1077	527	3607.712
ADNSPDKICFYDVEKXTVYDFKGTGIDR-LLSESMKQAEVGYGRFEAE-a25-b7	IF140	WDR35	inter-protein xl	YW11936/YW11953	615	743	6440.984
AEVXKVLGK-LGKLLSEK-a8-b3	IF140	WDR35	inter-protein xl	YW11936/YW11953	272	736	2592.416
LSGKTGR-KEGR-a4-b1	IF140	WDR35	inter-protein xl	YW11931/YW11936	276	473	1628.888
DTVYCVAYADGK-K10-155	IF122		monolink	YW11931/YW11953	63	n/a	1643.79
ELAXEALGLDFETAKK-K16-156	IF122		monolink	YW11931/YW11936/YW11953	663	n/a	2066.021
ELAXEALGLDFETAKK-K17-155	IF122		monolink	YW11931/YW11936/YW11953	664	n/a	2065.049
ELAXEALGLDFETAKK-K17-156	IF122		monolink	YW11931/YW11936/YW11953	664	n/a	2066.003
LDSPGAAETVIXKGDLLK-K13-155	IF122		monolink	YW11931/YW11953	830	n/a	2142.073
XLIITQADWAR-K5-156	IF122		monolink	YW11931/YW11953	753	n/a	1503.764
CLDXSARKK-GKEPYTKK-a9-b7	IF122	TT21B	inter-protein xl	YW11931/YW11936/YW11953	527	162	2298.164
FASGSADKSVIWT5K-LKWYDK-a8-b2	IF122	TT21B	inter-protein xl	YW11931/YW11936	75	803	2685.423
FECNLLVCANHIILCOEKR-NVLCYDIAELLK-K19-b13	IF122	TT21B	inter-protein xl	YW11931/YW11936/YW11953	452	801	4521.286
HPFKDDIYMPYQWLAEADRFEAAQK-LFPVDDKCR-a5-b8	IF122	WDR19	inter-protein xl	YW11931/YW11936	861	475	4727.227
IMGDLKSLVQLHVETQRWDEAFALGEK-YKEAVVAYENAK-a5-b2	IF122	WDR19	inter-protein xl	YW11936/YW11953	835	912	4521.278
GHKDTVYCVAYADGK-KASEIEDFR-a3-b1	IF122	IF143	inter-protein xl	YW11931/YW11936	53	76	3171.491
DGKRFAASADK-KYTOAGNK-a3-b1	IF122	IF140	inter-protein xl	YW11931/YW11953	66	1211	2284.132
KLIFKAYQIACLVGTDWNR-GENXNCYCKVK-a4-b11	IF122	IF140	inter-protein xl	YW11931/YW11936	631	334	4227.919
NDILAVADWQGVSYQLSGK-YAMWRKVPDFLGSFGAEGK-a12-b6	IF122	IF140	inter-protein xl	YW11931/YW11936/YW11953	262	353	4520.301
RCKDDPGP-KEPEIMK-a3-b1	IF122	IF140	inter-protein xl	YW11936/YW11953	1236	1261	1954.955
CLDMSARKK-NKNGEEK-a9-b2	IF122	IF122	intra-protein xl	YW11931/YW11953	527	161	2150.049
CLDXSARKK-CKELVK-a9-b2	IF122	IF122	intra-protein xl	YW11931/YW11936	527	388	2124.061
CLDXSARKK-QIGKDR-a9-b4	IF122	IF122	intra-protein xl	YW11931/YW11936	527	275	2064.041
DTSPGSKVK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	YW11931/YW11936/YW11953	995	1021	2070.103
FLLHSLPKDTPSGISK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	YW11931/YW11936/YW11953	987	1021	2778.506
HKELVK-NGEEKVK-a4-b5	IF122	IF122	intra-protein xl	YW11936/YW11953	388	166	1957.092
QSKALGAYR-NIKPK-a3-b3	IF122	IF122	intra-protein xl	YW11931/YW11936	1007	762	1858.03
QSKALGAYR-RCKDDPGP-a3-b3	IF122	IF122	intra-protein xl	YW11931/YW11936	1007	1236	2074.033
FNDAAYYWMLSMQCLDAQDPAQKDTXIGKHFQR-NKSVRSXSNADGQK-a25-b2	IF122	WDR35	inter-protein xl	YW11936/YW11953	935	115	6541.031
IMGDLKSLVQLHVETQRWDEAFALGEK-CKDYQGIKFK-a5-b2	IF122	WDR35	inter-protein xl	YW11931/YW11936	835	723	4522.284
QSKALGAYR-HTSKDNR-a3-b4	IF122	WDR35	inter-protein xl	YW11931/YW11936	1007	1119	1987.014
YIQEIPSTLKSAYSSQSGSEAFEEPEEEDDSDPR-LYVLSALLIEQYHEQMIKNAQR-a10-b17	IF122	WDR35	inter-protein xl	YW11931/YW11936	219	986	6512.038
DNPDLFAMMEKTR-K11-155	WDR35		monolink	YW11936/YW11953	614	n/a	1721.821
DNPDLFAMMEKTR-K11-156	WDR35		monolink	YW11931/YW11936	614	n/a	1722.808
DNRKPELDSLMEGEGKLPCTCATGSPITVEYQFWMCSVCK-K40-155	WDR35		monolink	YW11931/YW11953	1159	n/a	4745.208
DVWDMKWKNDPDLFAMMEK-K6-156	WDR35		monolink	YW11931/YW11953	600	n/a	2625.201
DVWDMKWKNDPDLFAMMEK-K9-156	WDR35		monolink	YW11931/YW11953	603	n/a	2625.171
DVWDMKWKNDPDLFAMMEK-K20-156	WDR35		monolink	YW11931/YW11936	614	n/a	2641.177
KPELDSLXEGGEGK-K1-155	WDR35		monolink	YW11931/YW11953	1123	n/a	1659.818
RDVWDXKWK-K10-155	WDR35		monolink	YW11931/YW11936	603	n/a	1504.764
SXSNWADGQKICVYEDGAVVGSVDGNRIWVK-K10-156	WDR35		monolink	YW11931/YW11936	129	n/a	3795.805
SXSNWADGQKICVYEDGAVVGSVDGNRIWVK-K33-156	WDR35		monolink	YW11936/YW11953	152	n/a	3795.804
WAKDNDPDLFAMMEKTR-K14-156	WDR35		monolink	YW11936/YW11953	614	n/a	2108.013
AVELAKNHSMK-KXGEIADAIK-a6-b1	WDR35	TT21B	inter-protein xl	YW11936/YW11953	906	575	2455.28
CKDYQGIKFKV-KMGEIADAIK-a2-b1	WDR35	TT21B	inter-protein xl	YW11936/YW11953	723	575	2597.351
DNPDLFAMMEKTR-KGAGEK-a11-b1	WDR35	TT21B	inter-protein xl	YW11931/YW11953	614	104	2293.102
GKSSFAT5ALAGLLEEVLSITDR-HNRSCCKAYEXGYIMEK-a2-b7	WDR35	TT21B	inter-protein xl	YW11931/YW11953	996	1232	4956.257
KPELDSLMEGEGK-YDXAEDLLKR-a1-b9	WDR35	TT21B	inter-protein xl	YW11931/YW11936	1123	1221	2895.414

KPELDSLXEGGEGK-GKPEPTYTK-a1-b2	WDR35	TT21B	inter-protein xl	YWI1936/YWI1953	1123	157	2592.299
KPELDSLXEGGEGK-KPDNYTLR-a1-b1	WDR35	TT21B	inter-protein xl	YWI1931/YWI1936	1123	982	2882.37
LAECYMLEDEYGLLENLAISLPENHKLPEIAQMFVR-SCCKAYEYMGYIMEK-a26-b4	WDR35	TT21B	inter-protein xl	YWI1936/YWI1953	857	1232	6480.977
AAVDTCVHUNQWNAKAVELAKNHSMK-TSAAESMISVIGKK-a20-b14	WDR35	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	906	208	4521.277
LMFKIADEEAK-AKXGRR-a4-b2	WDR35	IF143	inter-protein xl	YWI1936/YWI1953	952	21	2165.138
AAMQDKDVNIQYTPFGEHLGLT-LIKSEAVWENXAR-a6-b3	WDR35	IF140	inter-protein xl	YWI1931/YWI1936	281	799	4521.267
COJMRHENDONPVLIDTGMVYVGIQWNHMGSVLAVAFQKAAAMODK-KYPDFLGSPGAEK-a40-b1	WDR35	IF140	inter-protein xl	YWI1936/YWI1953	275	353	6738.26
DNRRKPELDSLMEGGEGK-KEPEIK-a4-b1	WDR35	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1123	1261	2901.398
HTSKDNRRKPELDSLMEGGEGK-ALLKSGDTEKITFFASVSR-a8-b10	WDR35	IF140	inter-protein xl	YWI1936/YWI1953	1123	1233	4534.296
KPELDSLMEGGEGKLPCTCVATGSPITEYQFWXCSYCK-SEAVWENXARMCVKR-a1-b14	WDR35	IF140	inter-protein xl	YWI1931/YWI1936	1123	813	6469.983
SXSWNADGQKICVYEDGAVVGSVDGNR-GENMNCVCYCK-a10-b11	WDR35	IF140	inter-protein xl	YWI1931/YWI1953	129	334	4954.251
VGMCEQAVTFLKCSQPK-KYTOAGNKLK-a13-b1	WDR35	IF140	inter-protein xl	YWI1931/YWI1936	881	1211	3340.719
AVELAKNHSMKIEGSLAR-DFLGSDDPKETK-a11-b9	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	911	745	3496.831
DVWDMKWAK-HAYDKLR-a6-b5	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	600	1021	2217.102
IADDEAKK-SVSKHK-a7-b4	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	959	128	1724.919
KPELDSLMEGGEGKLPCTCVATGSPITEYQFWMCSYCK-KLDSPGYAAETLYKMGDLK-a1-b1	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	1123	817	6442.001
IMKLSCLVNVVTGAIAGIHWHYHGTEGYVEPDCPLAVCFDNGRCQJXR-VRIKCK-a2-b4	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	194	386	6511.064
NNKEYVK-CKELVK-a4-b2	WDR35	IF122	inter-protein xl	YWI1936/YWI1953	368	388	1806.96
SMSWNADGQKICVYEDGAVVGSVDGNR-KAFIRVQDLR-a10-b1	WDR35	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	129	664	4522.289
WGYCSNTVYAYTRPDRPEYCVFWDTKNNEK-NIKEPKAAVEXYISAGEHVK-a28-b3	WDR35	IF122	inter-protein xl	YWI1936/YWI1953	364	762	6383.976
WGYCSNTVYAYTRPDRPEYCVFWDTKNNEKYVK-ELAXEALLEGIDFETAKK-a28-b16	WDR35	IF122	inter-protein xl	YWI1931/YWI1936	364	663	6454.998

Table 20: Identified links of IFT-A from Flp-In (N)-SF-TULP3 cells after SEC (in at least 3of3)

Id	Protein1	Protein2	XLType	Biological Replicates	AbsPos1	AbsPos2	Mr
EKMADIYKLR-K9-155	TT21B		monolink	YWI1931/YWI1936/YWI1953	700	n/a	1557.828
FFSMAEKRSR-K7-156	TT21B		monolink	YWI1931/YWI1936/YWI1953	1014	n/a	1527.762
EQRKAGEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	104	1171	2124.059
AKLEPGFYCK-ETKMLTK-a2-b3	TT21B	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	1020	748	2440.268
KDRDWGQNALYNXIEICLNPNETVGGVEFELDGLGNSTEK-KPELDSLMGEGEK-a1-b1	TT21B	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1051	1123	6512.035
YDMAEDLLKR-XFYLSKK-a9-b7	TT21B	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1221	7	2469.221
VATCGDNKIQDLVDLK-K10-156	WDR19		monolink	YWI1931/YWI1936/YWI1953	299	n/a	2217.098
HMLKDDWTVCPHCDPFPALYSELKIMLNTSTPCMSER-SAEILYLHVLAXKK-a4-b14	WDR19	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	1286	407	6511.033
YKEAVVAYENAK-VKEALKWYK-a2-b2	WDR19	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	912	341	2685.427
HMLKDDWTVCPHCDPFPALYSELKIMLNTSTPCMSER-DAKYLFRLYMALK-a23-b3	WDR19	WDR19	intra-protein xl	YWI1931/YWI1936/YWI1953	1305	1090	6439.994
ALIKNSAFSAAMLIMRPEYSK-AVSGDEKALDMFNWKK-a4-b16	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1207	187	4521.29
ALKHLKCPSSDNNVAIEMAIETVGOAK-GENYILSPDEKFGFKGENMNCVCYCK-a3-b16	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1041	323	6511.006
NGDHMKGAR-KEPEIMK-a6-b1	WDR19	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1171	1261	1995.969
IEGMVRRPDISEIEAATPCPFCKFLPCELLCPGCK-NKSVRSXSWNADGQK-a24-b2	WDR19	WDR35	inter-protein xl	YWI1931/YWI1936/YWI1953	1257	115	6511.042
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	YWI1931/YWI1936/YWI1953	187	n/a	2065.017
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	YWI1931/YWI1936/YWI1953	188	n/a	2065.007
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	YWI1931/YWI1936/YWI1953	179	n/a	2065.006
ALSYEKSDTHR-K7-156	IF140		monolink	YWI1931/YWI1936/YWI1953	926	n/a	1624.767
GVQMDRAVMLYHKAGHFSK-K13-156	IF140		monolink	YWI1931/YWI1936/YWI1953	1090	n/a	2330.174
VQGTPLLKHGYGK-K8-156	IF140		monolink	YWI1931/YWI1936/YWI1953	145	n/a	1624.877
AEEVXKVLGSK-EKMADYK-a8-b2	IF140	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	272	693	2581.358
DKTLWRWVAQYLESQEMDAALHYELAR-RDCGAILNKKQFSEAAQLYEK-a2-b11	IF140	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	955	854	6382.968
RETLFNFQETNKSHUFVDEGLK-VLKRDCGAILNKK-a13-b3	IF140	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	634	843	4520.297
DKTLWRWVAQYLESQEMDAALHYELAR-KYQEAQLCLGQNMSITEEMAEK-a2-b1	IF140	IF140	intra-protein xl	YWI1931/YWI1936/YWI1953	955	1151	6480.026
YYEEKGVQMDR-LKAMRALLK-a5-b2	IF140	IF140	intra-protein xl	YWI1931/YWI1936/YWI1953	1077	1220	2597.358
AEEVMKYK-SVSKHK-a6-b4	IF140	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	270	128	1754.958
VQGTPLLKHGYGK-CKELVKK-a8-b6	IF140	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	145	392	2510.373
ELAXEALGLDFETAKK-K16-156	IF122		monolink	YWI1931/YWI1936/YWI1953	663	n/a	2066.021
CLDXASRKK-GPEYTKK-a9-b7	IF122	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	527	162	2298.164
FECNLLVVCANHIILCOEKR-NVLCYDLAELLK-a19-b13	IF122	TT21B	inter-protein xl	YWI1931/YWI1936/YWI1953	452	801	4521.286
NDILAVADWGGKVSFYQLSGK-VAMWRKVPDFLSPGAEGK-a12-b6	IF122	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	262	353	4520.301
DTPSGISKVK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	YWI1931/YWI1936/YWI1953	995	1021	2070.103
FLHSLPKDTPSGISK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	YWI1931/YWI1936/YWI1953	987	1021	2778.506
AAVDTCVHLNQWNAKAVELAKNHSMK-TSAAESMISVVLGKK-a20-b14	WDR35	WDR19	inter-protein xl	YWI1931/YWI1936/YWI1953	906	208	4521.277
DNRRPELDSLMGEGEK-KEPEIKK-a4-b1	WDR35	IF140	inter-protein xl	YWI1931/YWI1936/YWI1953	1123	1261	2901.398
SMSWNADGGKICVYEDGAVIVSGVDGMR-KAFIRVQDLR-a10-b1	WDR35	IF122	inter-protein xl	YWI1931/YWI1936/YWI1953	129	664	4522.289

**Table 21: Links of IFT-A (Flp-In (N)-SF-TULP3) identified after both enrichment methods**

Id	Protein1	Protein2	XLType	Method(s)	AbsPos1	AbsPos2	Mr
CLDMSASRKK-NKNGEEK-a9-b2	IF122	IF122	intra-protein xl	SEC/CutOff	527	161	2150.05
DTPSGISVKV-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	SEC/CutOff	995	1021	2070.111
FLLHSLPKDTPSGISK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	SEC/CutOff	987	1021	2778.506
LDSPGYAAETYLKXGDLK-QGGWAGDSVKASK-a13-b10	IF122	IFT43	inter-protein xl	SEC/CutOff	830	69	3414.633
ELAXEALGLDFETAKK-K16-156	IF122		monolink	SEC/CutOff	663	n/a	2066.003
ELAXEALGLDFETAKK-K17-155	IF122		monolink	SEC/CutOff	664	n/a	2065.019
ELAXEALGLDFETAKK-K17-156	IF122		monolink	SEC/CutOff	664	n/a	2066.003
LVETKDSIGDEDPFTAK-K17-155	IF122		monolink	SEC/CutOff	1153	n/a	2019.014
AEVVMKVK-SVSKHK-a6-b4	IF140	IF122	inter-protein xl	SEC/CutOff	270	128	1754.95
FGFEKGENXNCVYCK-ALLKSGDTEK-a5-b4	IF140	IF140	intra-protein xl	SEC/CutOff	323	1227	3256.458
AAVSGDEKALDXFNWK-CGKLEDVPR-a8-b3	IF140	TT21B	inter-protein xl	SEC/CutOff	179	1001	3007.419
NIIGFYTKGR-RCGKLEDVPR-a8-b4	IF140	TT21B	inter-protein xl	SEC/CutOff	1275	1001	2534.318
YEEKGVQMDR-YDXAEDLLKR-a5-b9	IF140	TT21B	inter-protein xl	SEC/CutOff	1077	1221	2823.303
FGFEKGENMNCVYCK-IHVKNGDHXKGAR-a5-b10	IF140	WDR19	inter-protein xl	SEC/CutOff	323	1171	3657.637
YEEKGVQMDR-KKIEGMVR-a5-b2	IF140	WDR19	inter-protein xl	SEC/CutOff	1077	1233	2514.274
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	SEC/CutOff	187	n/a	2065.035
AAVSGDEKALDMFNWKK-K17-155	IF140		monolink	SEC/CutOff	188	n/a	2064.036
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	SEC/CutOff	188	n/a	2065.029
AAVSGDEKALDMFNWKK-K8-155	IF140		monolink	SEC/CutOff	179	n/a	2064.036
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	SEC/CutOff	179	n/a	2065.019
ALSYEKS DTHR-K7-155	IF140		monolink	SEC/CutOff	926	n/a	1623.793
GENMNCVYCKVKGLLAAGTDR-K11-155	IF140		monolink	SEC/CutOff	334	n/a	2670.249
GVQMDRAVMMLYHKAGHFSK-K13-156	IF140		monolink	SEC/CutOff	1090	n/a	2330.176
KYQEALQLCLGQNM SITEEMA EKXTVAK-K1-155	IF140		monolink	SEC/CutOff	1151	n/a	3414.664
RIGASTKSK-HAYDKLR-a7-b5	TT21B	IF122	inter-protein xl	SEC/CutOff	603	1021	1986.085
IMENYCLMATKQK-YEEKGVQXDR-a11-b5	TT21B	IF140	inter-protein xl	SEC/CutOff	1134	1077	3199.441
FFSMAEKRSR-GKEPYTK-a7-b2	TT21B	TT21B	intra-protein xl	SEC/CutOff	1014	157	2331.16
AKLEPGFYCK-LGKLLSESMK-a2-b3	TT21B	WDR35	inter-protein xl	SEC/CutOff	1020	736	2582.343
YDMAEDLLKR-MFFYLSKK-a9-b7	TT21B	WDR35	inter-protein xl	SEC/CutOff	1221	7	2469.216
EALVHCETDNKIXLELAR-K11-155	TT21B		monolink	SEC/CutOff	919	n/a	2312.168
EALVHCETDNKIXLELAR-K11-156	TT21B		monolink	SEC/CutOff	919	n/a	2313.144
FFSMAEKRSR-K7-156	TT21B		monolink	SEC/CutOff	1014	n/a	1527.762
IAKMNWNAIDAEFEK-K3-156	TT21B		monolink	SEC/CutOff	1185	n/a	2063.972
XDSQELKTLINYYCQER-K7-155	TT21B		monolink	SEC/CutOff	7	n/a	2361.091
IDAKYK-NIKEPK-a4-b3	WDR19	IF122	inter-protein xl	SEC/CutOff	1229	762	1601.912
TSAESXISVVLGKK-KYTQAGNK-a14-b1	WDR19	IF140	inter-protein xl	SEC/CutOff	208	1211	2582.339
YKEAVVAYENAK-MKDCTLWR-a2-b2	WDR19	IF140	inter-protein xl	SEC/CutOff	912	953	2598.34
ETQSLDGAKXVAR-K9-155	WDR19		monolink	SEC/CutOff	955	n/a	1575.78
GVLWENWPM DKG VFIAYDDDK-K21-156	WDR19		monolink	SEC/CutOff	580	n/a	2653.251
SEINLPGNCVAMDWDKGDVLA VIAEK-K16-156	WDR19		monolink	SEC/CutOff	64	n/a	3114.472
VATCGDNCIKIQLVDL K-K10-156	WDR19		monolink	SEC/CutOff	299	n/a	2217.103
KPELDSLMEGGGK-LSGKTGR-a1-b4	WDR35	IF140	inter-protein xl	SEC/CutOff	1123	276	2344.189
RDVWDXKWAK-KPEEADR-a7-b1	WDR35	IF140	inter-protein xl	SEC/CutOff	600	733	2331.155
DNPDLFAMMEKTR-FFSXA EKR-a11-b7	WDR35	TT21B	inter-protein xl	SEC/CutOff	614	1014	2735.266
VLKLETQTD DAK-AKEADGRYK-a3-b2	WDR35	WDR19	inter-protein xl	SEC/CutOff	42	905	2534.312
DVWDMKWAKDNPDLFAXMEK-DVWDMKWAK-a6-b6	WDR35	WDR35	intra-protein xl	SEC/CutOff	600	600	3800.704
VGXCEQAVTAF LKCSQPK-K18-156	WDR35		monolink	SEC/CutOff	886	n/a	2225.033

Table 22: Identified links of IFT-A from HEK293T cells after SEC (in at least 2of3)

Id	Protein1	Protein2	XLType	Biological Replicates	AbsPos1	AbsPos2	deltaAA	Annotation	Mr
ALKYFEEGLQDNDTFALLGK-K21-155	TT21B		monolink	SEC_BR1/2	184	n/a	-	-	2483.241
EALVHCETDNKIXLELAR-K11-155	TT21B		monolink	SEC_BR1/2	919	n/a	-	-	2312.164
EALVHCETDNKIXLELAR-K11-156	TT21B		monolink	SEC_BR1/2/4	919	n/a	-	-	2313.12
EKMADIYLKHR-K9-155	TT21B		monolink	SEC_BR1/2/4	700	n/a	-	-	1557.825
FYHAYGTLXEGKTQEALR-K12-155	TT21B		monolink	SEC_BR1/3	52	n/a	-	-	2285.109
XIKISDGSQGHVVK-K15-155	TT21B		monolink	SEC_BR1/2	148	n/a	-	-	1811
XIKISDGSQGHVVK-K3-155	TT21B		monolink	SEC_BR2/4	136	n/a	-	-	1810.99
XIKISDGSQGHVVK-K9-155	TT21B		monolink	SEC_BR1/2/3	142	n/a	-	-	1810.989
YDMAEDLLKRCRLR-K9-155	TT21B		monolink	SEC_BR2/3	1221	n/a	-	-	1836.907
YDXAEDLLKRCRLR-K9-156	TT21B		monolink	SEC_BR2/4	1221	n/a	-	-	1853.898
DGTLASKXGK-IGASTKSK-a7-b6	TT21B	TT21B	intra-protein xl	SEC_BR1/3	761	603	158	-	1951.034
DKMLYITCFR-GKEPYTKK-a2-b2	TT21B	TT21B	intra-protein xl	SEC_BR1/2/3	705	157	548	-	2433.238
EKXADIYK-RIGASTKSK-a2-b7	TT21B	TT21B	intra-protein xl	SEC_BR1/2	693	603	90	-	2210.212
EPYTKKALK-HFNKAR-a5-b4	TT21B	TT21B	intra-protein xl	SEC_BR1/3/4	162	1048	886	-	1986.089
GKEPYTK-AQSQKK-a2-b5	TT21B	TT21B	intra-protein xl	SEC_BR2/3/4	157	574	417	-	1647.894
GKEPYTK-GKEPYTK-a2-b2	TT21B	TT21B	intra-protein xl	SEC_BR1/4	157	157	0	-	1780.914
GKEPYTK-KDILDK-a2-b1	TT21B	TT21B	intra-protein xl	SEC_BR1/4	157	1304	1147	-	1689.908
KPDNYXTLSR-KDILDK-a1-b1	TT21B	TT21B	intra-protein xl	SEC_BR1/4	982	1304	322	-	2108.07
TLHXAMSLPGMKR-AQSQKK-a12-b5	TT21B	TT21B	intra-protein xl	SEC_BR1/3	596	574	22	-	2314.182
DRKTEVDTSR-AKEADGR-a3-b2	TT21B	WDR19	inter-protein xl	SEC_BR2/4	608	905	297	-	2226.078
EQRKGAGEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	SEC_BR1/2/3/4	104	1171	1067	-	2124.061
IGASTKSK-SKNWAK-a6-b2	TT21B	WDR19	inter-protein xl	SEC_BR2/3	603	880	277	-	1660.904
MDIYLIKHR-NGDHMKGAR-a7-b6	TT21B	WDR19	inter-protein xl	SEC_BR2/4	700	1171	471	-	2284.102
VKEALKWYK-IDAKYKK-a2-b6	TT21B	WDR19	inter-protein xl	SEC_BR1/2	341	1231	890	-	2166.262
FFSMAEKNSR-SIKLIK-a7-b3	TT21B	IF140	inter-protein xl	SEC_BR1/2/3	1014	796	218	-	2210.213
FFSXAERNR-KYTQAGNK-a7-b1	TT21B	IF140	inter-protein xl	SEC_BR2/4	1014	1211	197	-	2434.224
DKMLYITCFR-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	SEC_BR1/2/3/4	705	1236	531	-	2271.03
EQRKGAGEK-NKNGEEK-a4-b2	TT21B	IF122	inter-protein xl	SEC_BR1/2/3/4	104	161	57	-	1956.982
GKEPYTK-LDKAER-a2-b3	TT21B	IF122	inter-protein xl	SEC_BR1/3	157	802	645	-	1689.908
KMGEIADAIK-CKDDPGP-a1-b2	TT21B	IF122	inter-protein xl	SEC_BR2/3	575	1236	661	-	1999.941
KPDNYXTLSR-LDKAER-a1-b3	TT21B	IF122	inter-protein xl	SEC_BR2/3	982	802	180	-	2108.064
SCCKAYEYMGYIXEK-KLDKAER-a4-b4	TT21B	IF122	inter-protein xl	SEC_BR2/4	1232	802	430	-	2944.332
DILDKAR-KEGRER-a5-b1	TT21B	WDR35	inter-protein xl	SEC_BR1/4	1309	473	836	-	1740.962
EKXADIYK-GSKPLR-a2-b3	TT21B	WDR35	inter-protein xl	SEC_BR1/3	693	963	270	-	1920.055
GKEPYTK-SRKEGR-a2-b3	TT21B	WDR35	inter-protein xl	SEC_BR1/4	157	473	316	-	1690.9
HFNKARK-KEGRER-a4-b1	TT21B	WDR35	inter-protein xl	SEC_BR1/2/3/4	1048	473	575	-	1810.992
IRKDILDK-HTSKDNR-a3-b4	TT21B	WDR35	inter-protein xl	SEC_BR3/4	1304	1119	185	-	1994.091
KTEVDTSR-GKVKGK-a1-b4	TT21B	WDR35	inter-protein xl	SEC_BR2/4	608	994	386	-	1825.009
YDXAEDLLKRMFFYLSKK-a9-b7	TT21B	WDR35	inter-protein xl	SEC_BR1/2/3	1221	7	1214	-	2469.219

DCGAILENXKQF5EAAQLYK-K10-155	WDR19		monolink	SEC_BR1/2/3	854	n/a	-	-	2615.247
ETQSLDGAKMVAR-K9-156	WDR19		monolink	SEC_BR2/3	955	n/a	-	-	1560.77
GLYDKAASVYIR-K6-155	WDR19		monolink	SEC_BR1/2/3/4	871	n/a	-	-	1672.882
IHLQYAKAKEADGR-K9-155	WDR19		monolink	SEC_BR2/3/4	905	n/a	-	-	1753.947
NGDHMKGARMILIR-K6-155	WDR19		monolink	SEC_BR1/3	1171	n/a	-	-	1652.874
NWAKVGDLLPHVSSPK-K16-155	WDR19		monolink	SEC_BR3/4	896	n/a	-	-	1902.043
RGVNOALKHPSR-K8-156	WDR19		monolink	SEC_BR1/2/4	836	n/a	-	-	1517.833
IHKNGDHMK-KTEVDITSHR-a4-b1	WDR19	TT21B	inter-protein xl	SEC_BR2/4	1165	608	557	-	2387.193
MKRIFLLEK-IGASTKSK-a2-b6	WDR19	TT21B	inter-protein xl	SEC_BR1/2	2	603	601	-	2192.239
NGDHMKGAR-XADYILKHR-a6-b7	WDR19	TT21B	inter-protein xl	SEC_BR2/3/4	1171	700	471	-	2284.1
AKADGR-KIEGMVR-a2-b1	WDR19	WDR19	intra-protein xl	SEC_BR3/4	905	1233	328	-	1714.896
RGVNOALKHPSR-LYXALKQYR-a8-b6	WDR19	WDR19	intra-protein xl	SEC_BR1/2	836	1100	264	-	2700.461
KISDCTQYLR-KEEYQIAYR-a1-b1	WDR19	IF140	inter-protein xl	SEC_BR1/2	1328	1389	61	-	2607.279
LYXALKQYR-VKLSGK-a6-b2	WDR19	IF140	inter-protein xl	SEC_BR2/3	1100	272	828	-	1969.094
ALKHFLK-CKDDPGP-a3-b2	WDR19	IF122	inter-protein xl	SEC_BR1/2/3	1041	1236	195	-	1780.903
ALKHFLK-NKNGEEK-a3-b2	WDR19	IF122	inter-protein xl	SEC_BR1/4	1041	161	880	-	1810.993
ETQSLDGAKXVAR-CKDDPGP-a9-b2	WDR19	IF122	inter-protein xl	SEC_BR1/2/3/4	955	1236	281	-	2346.061
IHKNGDHMK-CKDDPGP-a4-b2	WDR19	IF122	inter-protein xl	SEC_BR2/4	1165	1236	71	-	2102.969
IHKNGDHMKGAR-CKDDPGP-a10-b2	WDR19	IF122	inter-protein xl	SEC_BR1/4	1171	1236	65	-	2387.163
NGDHMKGAR-NGEEKYK-a6-b5	WDR19	IF122	inter-protein xl	SEC_BR3/4	1171	166	1005	-	1924.945
NGDHMKGAR-NKNGEEK-a6-b2	WDR19	IF122	inter-protein xl	SEC_BR2/3	1171	161	1010	-	1939.917
CROGGWAGDSVKASK-K15-155	IFT43		monolink	SEC_BR1/2	72	n/a	-	-	1760.865
AKMGRR-KEGRER-a2-b1	IFT43	WDR35	inter-protein xl	SEC_BR2/4	21	473	452	-	1628.898
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	SEC_BR1/2/3/4	187	n/a	-	-	2065.017
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	SEC_BR1/2/3	188	n/a	-	-	2065.022
AAVSGDEKALDMFNWKK-K8-155	IF140		monolink	SEC_BR1/4	179	n/a	-	-	2064.035
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	SEC_BR1/2/3/4	179	n/a	-	-	2065.017
AAVSGDEKALDXFNWKK-K16-155	IF140		monolink	SEC_BR1/2/3	187	n/a	-	-	2080.019
AAVSGDEKALDXFNWKK-K17-155	IF140		monolink	SEC_BR1/2/3/4	188	n/a	-	-	2080.019
AAVSGDEKALDXFNWKK-K8-155	IF140		monolink	SEC_BR2/4	179	n/a	-	-	2080.019
ALSYYEKSDTHR-K7-156	IF140		monolink	SEC_BR1/3/4	926	n/a	-	-	1624.771
AQAFKNAIRLCK-K12-156	IF140		monolink	SEC_BR2/3	1047	n/a	-	-	1574.853
AQAFKNAIRLCK-K5-156	IF140		monolink	SEC_BR1/2/3	1040	n/a	-	-	1574.852
DFVGLDCKATR-K10-156	IF140		monolink	SEC_BR1/4	770	n/a	-	-	1680.774
GENMNCVCYCKVGLLAAGTDR-K11-156	IF140		monolink	SEC_BR2/3	334	n/a	-	-	2671.246
GENMNCVCYCKVGLLAAGTDR-K13-156	IF140		monolink	SEC_BR1/2	336	n/a	-	-	2671.247
GENMNCVCYCKVGLLAAGTDR-K11-155	IF140		monolink	SEC_BR1/2/4	334	n/a	-	-	2686.221
GENMNCVCYCKVGLLAAGTDR-K13-155	IF140		monolink	SEC_BR1/2	336	n/a	-	-	2686.226
GVQMDRAVMLYHKAGHFSK-K13-156	IF140		monolink	SEC_BR1/2/3/4	1090	n/a	-	-	2330.173
GVQMDRAVXLYHKAGHFSK-K13-155	IF140		monolink	SEC_BR1/2/4	1090	n/a	-	-	2345.178



MLSEDLPSLEYVNMK-K15-156	IF140		monolink	SEC_BR1/4	951	n/a	-	-	2165.086
SEAVWENMARXCVKQR-K14-155	IF140		monolink	SEC_BR1/2	813	n/a	-	-	2266.061
SEAVWENMARXCVKQR-K14-156	IF140		monolink	SEC_BR2/4	813	n/a	-	-	2267.03
VPDFGSPGAEKDR-K13-156	IF140		monolink	SEC_BR1/3	366	n/a	-	-	1699.846
VQGTPLLKHEYK-K8-156	IF140		monolink	SEC_BR1/4	145	n/a	-	-	1624.878
XCVKQRLDVAK-K4-156	IF140		monolink	SEC_BR1/3/4	813	n/a	-	-	1619.851
YTOAGNKLKAMR-K7-156	IF140		monolink	SEC_BR1/4	1218	n/a	-	-	1535.799
EAKAHCSR-FFSXAEKR-a3-b7	IF140	TT21B	inter-protein xl	SEC_BR3/4	554	1014	460	-	2286.033
FGFEKGENMNCVCK-KGAGEK-a5-b1	IF140	TT21B	inter-protein xl	SEC_BR2/3	323	104	219	-	2768.189
GENMNCVCKVK-MADIYKHR-a11-b7	IF140	TT21B	inter-protein xl	SEC_BR1/2	334	700	366	-	2944.33
GENMNCVCKVK-TAEKLLK-a11-b4	IF140	TT21B	inter-protein xl	SEC_BR1/2	334	1106	772	-	2616.238
VKLSGKTGR-KIEGXVR-a6-b1	IF140	WDR19	inter-protein xl	SEC_BR3/4	276	1233	957	-	1930.091
LDVAKVCLGNXGHAR-AKMGR-a5-b2	IF140	IF143	inter-protein xl	SEC_BR1/2	821	21	800	-	2511.303
MTVAKDSSDLPEESR-EAKAHCSR-a5-b3	IF140	IF140	inter-protein xl	SEC_BR1/2	1178	554	624	-	2919.325
REAKAHCSR-KPEADR-a4-b1	IF140	IF140	intra-protein xl	SEC_BR1/3	554	733	179	-	2255.035
VKLSGKTGR-AEEVXVK-a6-b6	IF140	IF140	intra-protein xl	SEC_BR2/4	276	270	6	-	2031.124
XALVKRFIQAR-KYTQAGNK-a5-b1	IF140	IF140	intra-protein xl	SEC_BR1/2	1340	1211	129	-	2394.313
AEEVMKVK-SVSKHK-a6-b4	IF140	IF122	inter-protein xl	SEC_BR1/2/3/4	270	128	142	-	1754.952
AKSLDQETR-EKIHK-a2-b5	IF140	IF122	inter-protein xl	SEC_BR2/4	1320	432	888	-	2039.149
AVXLYKHAGHFSK-RCKDDPGP-a7-b3	IF140	IF122	inter-protein xl	SEC_BR1/2/3	1090	1236	146	-	2585.23
LSGKTGR-RCKDDPGP-a4-b3	IF140	IF122	inter-protein xl	SEC_BR2/4	276	1236	960	-	1954.99
NIIGFTKGR-RCKDDPGP-a8-b3	IF140	IF122	inter-protein xl	SEC_BR1/2/4	1275	1236	39	-	2249.131
REAKAHCSR-NKNGEEK-a4-b2	IF140	IF122	inter-protein xl	SEC_BR1/4	554	161	393	-	2229.021
XCVKQRLDVAK-RCKDDPGP-a4-b3	IF140	IF122	inter-protein xl	SEC_BR1/2	813	1236	423	-	2545.238
YVEEKVQMDR-LDKAER-a5-b3	IF140	IF122	inter-protein xl	SEC_BR1/3/4	1077	802	275	-	2285.082
CKRHLLNK-GSKPLR-a2-b3	IF140	WDR35	inter-protein xl	SEC_BR1/4	870	963	93	-	1977.076
GENMNCVCKVK-KEGRER-a11-b1	IF140	WDR35	inter-protein xl	SEC_BR1/2/4	334	473	139	-	2588.15
LSGKTGR-KEGRER-a4-b1	IF140	WDR35	inter-protein xl	SEC_BR1/2	276	473	197	-	1628.899
DTXLGKPYHFQR-K6-155	IF122		monolink	SEC_BR2/3/4	941	n/a	-	-	1712.821
DTXLGKPYHFQR-K6-156	IF122		monolink	SEC_BR1/2/3/4	941	n/a	-	-	1713.818
ELAXEALGLDFETAKK-K16-156	IF122		monolink	SEC_BR1/2/3/4	663	n/a	-	-	2066
ELAXEALGLDFETAKK-K17-156	IF122		monolink	SEC_BR1/2/3/4	664	n/a	-	-	2066
FEEAQAFHKAGR-K6-155	IF122		monolink	SEC_BR1/3	883	n/a	-	-	1672.887
LQCLFSGVKER-K10-156	IF122		monolink	SEC_BR1/3/4	463	n/a	-	-	1578.804
WDEAFALGEKHPEFK-K10-156	IF122		monolink	SEC_BR1/3/4	856	n/a	-	-	1958.947
XLITKQADWAR-K5-155	IF122		monolink	SEC_BR1/2/3	753	n/a	-	-	1502.805
XLITKQADWAR-K5-156	IF122		monolink	SEC_BR2/3/4	753	n/a	-	-	1503.763
CKELVKK-WYDKAEK-a2-b4	IF122	TT21B	inter-protein xl	SEC_BR1/2/3	388	807	419	-	1980.04
DTPSGISKVK-AQSQKK-a8-b5	IF122	TT21B	inter-protein xl	SEC_BR1/2/4	995	574	421	-	1857.03
ETKMLITK-VYSKMEK-a3-b4	IF122	TT21B	inter-protein xl	SEC_BR1/2	748	843	95	-	1984.058

HAYDKLR-VYSKMEK-a5-b4	IF122	TT21B	inter-protein xl	SEC_BR2/4	1021	843	178	-	1922.979
NKNGEEKVK-VKEQRK-a2-b2	IF122	TT21B	inter-protein xl	SEC_BR1/2/4	161	100	61	-	1969.099
CKDDPGP-DAKYLFK-a2-b3	IF122	WDR19	inter-protein xl	SEC_BR3/4	1236	1090	146	-	1836.859
QVSLFTKDGVR-NGDHKKGAR-a7-b6	IF122	WDR19	inter-protein xl	SEC_BR2/3	309	1171	862	-	2387.179
CKDDPGP-XCVKTQR-a2-b4	IF122	IF140	inter-protein xl	SEC_BR1/4	1236	813	423	-	1862.84
CLDXSASRKK-CLAKAK-a9-b4	IF122	IF140	inter-protein xl	SEC_BR1/2	527	1316	789	-	2038.021
CLDXSASRKK-KYTOAGNK-a9-b1	IF122	IF140	inter-protein xl	SEC_BR2/3/4	527	1211	684	-	2257.108
DGKRFASSADK-KYTOAGNK-a3-b1	IF122	IF140	inter-protein xl	SEC_BR1/3	66	1211	1145	-	2284.142
DGKRFASSADK-MALVKK-a3-b5	IF122	IF140	inter-protein xl	SEC_BR1/4	66	1340	1274	-	2092.102
AFHKAGRQR-CKDDPGP-a4-b2	IF122	IF122	intra-protein xl	SEC_BR1/2/3/4	887	1236	349	-	1994.979
AFHKAGRQR-VRIKCK-a4-b4	IF122	IF122	intra-protein xl	SEC_BR1/4	887	386	501	-	2010.16
CKDDPGP-NIKEPK-a2-b3	IF122	IF122	intra-protein xl	SEC_BR3/4	1236	762	474	-	1652.821
CLDMSASRKK-NKNGEEK-a9-b2	IF122	IF122	intra-protein xl	SEC_BR1/3	527	161	366	-	2150.05
CLDXSASRKK-QJGKDR-a9-b4	IF122	IF122	intra-protein xl	SEC_BR2/3/4	527	275	252	-	2064.034
MLITKQADWAR-SVSKHK-a5-b4	IF122	IF122	intra-protein xl	SEC_BR2/4	753	128	625	-	2154.178
QSKALGAYR-NIKEPK-a3-b3	IF122	IF122	intra-protein xl	SEC_BR2/3/4	1007	762	245	-	1858.03
QSKALGAYR-RCKDDPGP-a3-b3	IF122	IF122	intra-protein xl	SEC_BR2/4	1007	1236	229	-	2074.035
AFHKAGRQR-DVWDMKWAK-a4-b6	IF122	WDR35	inter-protein xl	SEC_BR1/2	887	600	287	-	2385.229
ANYFFDAAKLXFK-K9-155	WDR35		monolink	SEC_BR1/3/4	948	n/a	-	-	1735.875
ANYFFDAAKLXFK-K9-156	WDR35		monolink	SEC_BR1/4	948	n/a	-	-	1736.84
DNPDLFAMMEKTR-K11-156	WDR35		monolink	SEC_BR1/3/4	614	n/a	-	-	1722.81
HTSKDNRPDLXLGEGEGK-K21-156	WDR35		monolink	SEC_BR1/4	1136	n/a	-	-	2499.198
IYHVDDTPSGSMDGVLDSYKTIQTR-K20-155	WDR35		monolink	SEC_BR1/2	498	n/a	-	-	3009.422
KANYFFDAAKLMFK-K1-156	WDR35		monolink	SEC_BR1/3/4	939	n/a	-	-	1848.938
LMFKIADEEAKK-K11-155	WDR35		monolink	SEC_BR1/2	959	n/a	-	-	1576.86
RDVVDXKWAK-K10-155	WDR35		monolink	SEC_BR2/4	603	n/a	-	-	1504.765
RDVVDXKWAK-K7-155	WDR35		monolink	SEC_BR1/2/3/4	600	n/a	-	-	1504.753
THVIAASKEAFYTWQYR-K8-156	WDR35		monolink	SEC_BR3/4	446	n/a	-	-	2226.132
VGMCEQAVTAFKCSQPK-K13-156	WDR35		monolink	SEC_BR1/2	881	n/a	-	-	2209.04
VGMCEQAVTAFKCSQPK-K18-156	WDR35		monolink	SEC_BR2/4	886	n/a	-	-	2209.037
VTDSTGQVVGELLKLER-K15-156	WDR35		monolink	SEC_BR2/4	590	n/a	-	-	2127.146
WAKDNPDFAMMEKTR-K3-155	WDR35		monolink	SEC_BR1/2/4	603	n/a	-	-	2107.038
AVELAKNHSMK-KGAGEK-a6-b1	WDR35	TT21B	inter-protein xl	SEC_BR2/3	906	104	802	-	1953.053
GSKPLRYKK-LAFNYLKAK-a3-b7	WDR35	TT21B	inter-protein xl	SEC_BR2/4	963	1278	315	-	2216.325
KPELDSLXEGEGEK-AOSQKK-a1-b5	WDR35	TT21B	inter-protein xl	SEC_BR1/2/3	1123	574	549	-	2331.158
KPELDSLXEGEGEK-DYEKAIKIFYR-a1-b4	WDR35	TT21B	inter-protein xl	SEC_BR1/2	1123	902	221	-	2974.487
KPELDSLXEGEGEK-IGASTKSK-a1-b6	WDR35	TT21B	inter-protein xl	SEC_BR2/3/4	1123	603	520	-	2433.237
DNPDLFAMMEKTR-NGDHMKGAR-a11-b6	WDR35	WDR19	inter-protein xl	SEC_BR2/4	614	1171	557	-	2689.258
WAKDNPDFAXMEK-IDAKYK-a3-b4	WDR35	WDR19	inter-protein xl	SEC_BR1/2/4	603	1229	626	-	2585.235
CKDYOIIL-LSGKTGR-a2-b4	WDR35	IF140	inter-protein xl	SEC_BR1/2	723	276	447	-	1865.973

LETQTDDAKLR-KYTOAGNK-a9-b1	WDR35	IF140	inter-protein xi	SEC_BR2/4	51	1211	1160	-	2335.203
LGKLLSESMK-MKDKTLWR-a3-b2	WDR35	IF140	inter-protein xi	SEC_BR2/4	736	953	217	-	2319.277
RDVWDMKWAK-MALVKR-a7-b5	WDR35	IF140	inter-protein xi	SEC_BR1/3/4	600	1340	740	-	2188.151
DNPDLFAMMEKTR-CKDDPGP-a11-b2	WDR35	IF122	inter-protein xi	SEC_BR1/2/3/4	614	1236	622	-	2492.11
DVWDMKWAK-NKNGEEK-a6-b2	WDR35	IF122	inter-protein xi	SEC_BR1/2	600	161	439	-	2133.017
HTSKDNR-NGEEKV-k-a4-b5	WDR35	IF122	inter-protein xi	SEC_BR2/4	1119	166	953	-	1796.898
KANYFFDAAK-SVSKHKSSK-a1-b6	WDR35	IF122	inter-protein xi	SEC_BR2/3	939	130	809	-	2385.225
KISIPNNVK-NKNGEEK-a1-b2	WDR35	IF122	inter-protein xi	SEC_BR1/4	8	161	153	-	1967.065
LMFKIADDEAK-LDKAER-a4-b3	WDR35	IF122	inter-protein xi	SEC_BR1/2/4	952	802	150	-	2162.111
NNEKYVK-CKELVK-a4-b2	WDR35	IF122	inter-protein xi	SEC_BR2/4	368	388	20	-	1806.949
CKDYQGK-GKYKGGK-a2-b2	WDR35	WDR35	intra-protein xi	SEC_BR1/2	723	992	269	-	1763.963
DVWDMKWAK-GSKPLR-a6-b3	WDR35	WDR35	intra-protein xi	SEC_BR1/2/3/4	600	963	363	-	1972.011
DVWDMKWAK-KGSKPLR-a6-b4	WDR35	WDR35	intra-protein xi	SEC_BR1/2/3/4	600	963	363	-	2100.103
KPELDSLXEGGEGK-HTSKDNR-a1-b4	WDR35	WDR35	intra-protein xi	SEC_BR2/4	1123	1119	4	-	2499.198
VLKLETQDDAK-GSKPLR-a3-b3	WDR35	WDR35	intra-protein xi	SEC_BR2/3	42	963	921	-	2154.179

**Table 23: Crosslinks of IFT-A from HEK293T cells after SEC (in at least 3of3)**

Id	Protein1	Protein2	XLType	Biological Replicates	AbsPos1	AbsPos2	Mr
EALVHCETDNKIXLELAR-K11-156	TT21B		monolink	SEC_BR1/2/4	919	n/a	2313.12
EKMADIYLKHR-K9-155	TT21B		monolink	SEC_BR1/2/4	700	n/a	1557.825
XIKISDGSKQGHVVK-K9-155	TT21B		monolink	SEC_BR1/2/3/4	142	n/a	1810.989
EPYTKKALK-HFNKAR-a5-b4	TT21B	TT21B	intra-protein xl	SEC_BR1/3/4	162	1048	1986.089
GKEPYTK-AQSQKK-a2-b5	TT21B	TT21B	intra-protein xl	SEC_BR2/3/4	157	574	1647.894
EQKRGAGEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	SEC_BR1/2/3/4	104	1171	2124.061
FFSMAEKRSR-SIKLIK-a7-b3	TT21B	IF140	inter-protein xl	SEC_BR1/2/3	1014	796	2210.213
DKMLYITCFR-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	SEC_BR1/2/3/4	705	1236	2271.03
EQKRGAGEK-NKNGEEK-a4-b2	TT21B	IF122	inter-protein xl	SEC_BR1/2/3/4	104	161	1956.982
HFNKARK-KEGRER-a4-b1	TT21B	WDR35	inter-protein xl	SEC_BR1/2/3/4	1048	473	1810.992
YDMAEDLLKR-XFFYLSKK-a9-b7	TT21B	WDR35	inter-protein xl	SEC_BR1/2/3	1221	7	2469.218
DCGAILENXQFSEAAQLEYK-K10-155	WDR19		monolink	SEC_BR1/2/3	854	n/a	2615.247
GLYYDKAASVYIR-K6-155	WDR19		monolink	SEC_BR1/2/3/4	871	n/a	1672.882
IHLQYAKAKEADGR-K9-155	WDR19		monolink	SEC_BR2/3/4	905	n/a	1753.947
RGVNQALKHPSR-K8-156	WDR19		monolink	SEC_BR1/2/4	836	n/a	1517.833
NGDHMKGAR-XADIYLKHR-a6-b7	WDR19	TT21B	inter-protein xl	SEC_BR2/3/4	1171	700	2284.1
ALKHFLK-CKDDPGP-a3-b2	WDR19	IF122	inter-protein xl	SEC_BR1/2/3	1041	1236	1780.903
ETQSLDGAIXVAR-CKDDPGP-a9-b2	WDR19	IF122	inter-protein xl	SEC_BR1/2/3/4	955	1236	2346.061
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	SEC_BR1/2/3/4	187	n/a	2065.017
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	SEC_BR1/2/3	188	n/a	2065.022
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	SEC_BR1/2/3/4	179	n/a	2065.017
AAVSGDEKALDXFNWKK-K16-155	IF140		monolink	SEC_BR1/2/3	187	n/a	2080.019
AAVSGDEKALDXFNWKK-K17-155	IF140		monolink	SEC_BR1/2/3/4	188	n/a	2080.019
ALSYEKS DTHR-K7-156	IF140		monolink	SEC_BR1/3/4	926	n/a	1624.771
AQAFKNAILCK-K5-156	IF140		monolink	SEC_BR1/2/3	1040	n/a	1574.852
GENXNCVCYCKVGLAAGTDR-K11-155	IF140		monolink	SEC_BR1/2/4	334	n/a	2686.221
GVQMDRAVMLYHKAGHFSK-K13-156	IF140		monolink	SEC_BR1/2/3/4	1090	n/a	2330.173
GVQMDRAVXLYHKAGHFSK-K13-155	IF140		monolink	SEC_BR1/2/4	1090	n/a	2345.178
XCVTKQRLDVAK-K4-156	IF140		monolink	SEC_BR1/3/4	813	n/a	1619.851
AEEVMKVK-SVSKHK-a6-b4	IF140	IF122	inter-protein xl	SEC_BR1/2/3/4	270	128	1754.952
AVXLYHKAGHFSK-RCKDDPGP-a7-b3	IF140	IF122	inter-protein xl	SEC_BR1/2/3	1090	1236	2585.23
NIIGFYTKGR-RCKDDPGP-a8-b3	IF140	IF122	inter-protein xl	SEC_BR1/2/4	1275	1236	2249.131
YYEEKGVQMDR-LDKAER-a5-b3	IF140	IF122	inter-protein xl	SEC_BR1/3/4	1077	802	2285.082
GENXNCVCYCKVK-KEGRER-a11-b1	IF140	WDR35	inter-protein xl	SEC_BR1/2/4	334	473	2588.15
DTXLGKFYHFQR-K6-155	IF122		monolink	SEC_BR2/3/4	941	n/a	1712.821
DTXLGKFYHFQR-K6-156	IF122		monolink	SEC_BR1/2/3/4	941	n/a	1713.818
ELAXEALGLDFETAKK-K16-156	IF122		monolink	SEC_BR1/2/3/4	663	n/a	2066
ELAXEALGLDFETAKK-K17-156	IF122		monolink	SEC_BR1/2/3/4	664	n/a	2066
LQCLSFSGVKER-K10-156	IF122		monolink	SEC_BR1/3/4	463	n/a	1578.804
WDEAFALGEKHPFK-K10-156	IF122		monolink	SEC_BR1/3/4	856	n/a	1958.947
XLITKQADWAR-K5-155	IF122		monolink	SEC_BR1/2/3	753	n/a	1502.805
XLITKQADWAR-K5-156	IF122		monolink	SEC_BR2/3/4	753	n/a	1503.763
CKELVKK-WYDKAEK-a2-b4	IF122	TT21B	inter-protein xl	SEC_BR1/2/3	388	807	1980.04
DTPSGISKVK-AQSQKK-a8-b5	IF122	TT21B	inter-protein xl	SEC_BR1/2/4	995	574	1857.03
NKNGEEKVK-VKEQRK-a2-b2	IF122	TT21B	inter-protein xl	SEC_BR1/2/4	161	100	1969.099
CLDXSASRKK-KYTQAGNK-a9-b1	IF122	IF140	inter-protein xl	SEC_BR2/3/4	527	1211	2257.108
AFHKAGRQR-CKDDPGP-a4-b2	IF122	IF122	intra-protein xl	SEC_BR1/2/3/4	887	1236	1994.979
CLDXSASRKK-QIGKDR-a9-b4	IF122	IF122	intra-protein xl	SEC_BR2/3/4	527	275	2064.034
QSKALGAYR-NIKEPK-a3-b3	IF122	IF122	intra-protein xl	SEC_BR2/3/4	1007	762	1858.03
ANYFFDAAKLXFK-K9-155	WDR35		monolink	SEC_BR1/3/4	948	n/a	1735.875
DNPDLFAMMEKTR-K11-156	WDR35		monolink	SEC_BR1/3/4	614	n/a	1722.81
KANYFFDAAKLMFK-K1-156	WDR35		monolink	SEC_BR1/3/4	939	n/a	1848.938
RDVWDXKWAK-K7-155	WDR35		monolink	SEC_BR1/2/3/4	600	n/a	1504.753
WAKDNPDLFAMMEKTR-K3-155	WDR35		monolink	SEC_BR1/2/4	603	n/a	2107.038
KPELDSLXEGGEGK-AQSQKK-a1-b5	WDR35	TT21B	inter-protein xl	SEC_BR1/2/3	1123	574	2331.158
KPELDSLXEGGEGK-IGASTKSK-a1-b6	WDR35	TT21B	inter-protein xl	SEC_BR2/3/4	1123	603	2433.237
WAKDNPDLFAXMEK-IDAKYK-a3-b4	WDR35	WDR19	inter-protein xl	SEC_BR1/2/4	603	1229	2585.235
RDVWDMKWAK-MALVKR-a7-b5	WDR35	IF140	inter-protein xl	SEC_BR1/3/4	600	1340	2188.151
DNPDLFAMMEKTR-CKDDPGP-a11-b2	WDR35	IF122	inter-protein xl	SEC_BR1/2/3/4	614	1236	2492.11
LMFKIADEEAK-LDKAER-a4-b3	WDR35	IF122	inter-protein xl	SEC_BR1/2/4	952	802	2162.111
DVWDMKWAK-GSKPLR-a6-b3	WDR35	WDR35	intra-protein xl	SEC_BR1/2/3/4	600	963	1972.011
DVWDMKWAK-GSKPLR-a6-b4	WDR35	WDR35	intra-protein xl	SEC_BR1/2/3/4	600	963	2100.103

Table 24: Crosslinks of IFT-A (HEK293T) cells after spin column filtration (in at least 2of3)

Id	Protein1	Protein2	XLType	Biological Replicates	AbsPos1	AbsPos2	Mr
EALVHCETDNKIMLELAR-K11-156	TT21B		monolink	CutOff_BR1/2/3/4	919	n/a	2297.151
EALVHCETDNKIXLELAR-K11-155	TT21B		monolink	CutOff_BR2/4	919	n/a	2312.164
EHPALLGMATAYXILKQTPR-K17-156	TT21B		monolink	CutOff_BR3/4	1171	n/a	2525.327
IMGKALIKTHNSMAITYEAALK-K3-155	TT21B		monolink	CutOff_BR2/4	764	n/a	2771.436
SAELYLHAVLAXKK-K14-155	TT21B		monolink	CutOff_BR1/3	407	n/a	1857.034
SCCKAYEYMGYMEKEQAYTDAALNYEXAWK-K31-156	TT21B		monolink	CutOff_BR1/2/3/4	1259	n/a	3988.705
SCCKAYEYMGYMEKEQAYTDAALNYEMAWK-K15-156	TT21B		monolink	CutOff_BR1/4	1243	n/a	3988.702
XIKISDGSKQGHVLIK-K3-155	TT21B		monolink	CutOff_BR1/3	136	n/a	1811.006
XIKISDGSKQGHVLIK-K9-155	TT21B		monolink	CutOff_BR1/2	142	n/a	1810.994
CGKLEDVPR-IRKDIIDK-a3-b3	TT21B	TT21B	intra-protein xl	CutOff_BR1/2/4	1001	1304	2210.221
CQVLLAKVSK-EPYTKK-a7-b5	TT21B	TT21B	intra-protein xl	CutOff_BR2/3	839	162	2210.215
EFEAIKKNK-VKEALK-a6-b2	TT21B	TT21B	intra-protein xl	CutOff_BR2/3	64	341	1802.015
EKKADIYLLK-RIGASTKSK-a2-b7	TT21B	TT21B	intra-protein xl	CutOff_BR2/3/4	693	603	2210.206
EPYTKKALK-HFNKAR-a5-b4	TT21B	TT21B	intra-protein xl	CutOff_BR1/4	162	1048	1986.087
GKEPYTK-AOSQKK-a2-b5	TT21B	TT21B	intra-protein xl	CutOff_BR1/3/4	157	574	1647.891
HDKAREVIDR-AIKFYR-a3-b3	TT21B	TT21B	intra-protein xl	CutOff_BR1/4	126	905	2236.158
IXENYCLMATKQK-IRKDIIDK-a11-b3	TT21B	TT21B	intra-protein xl	CutOff_BR1/2/3/4	1134	1304	2782.463
KTEVDTSHR-KTEVDTSHR-a1-b1	TT21B	TT21B	intra-protein xl	CutOff_BR1/2	608	608	2281.137
LEDVPRFFSXAERK-DGLASKXGK-a13-b7	TT21B	TT21B	intra-protein xl	CutOff_BR1/2	1014	761	2900.442
IMDSQELKTLINYYCQER-GKEPYTKK-a7-b2	TT21B	TT21B	intra-protein xl	CutOff_BR2/2	7	157	3277.626
IMGEIADAIKTLHMVXSLPGMK-MADIYLLKHR-a9-b7	TT21B	TT21B	intra-protein xl	CutOff_BR2/4	584	700	3543.793
XEKLGDALQQAR-XIKISDGSK-a3-b3	TT21B	TT21B	intra-protein xl	CutOff_BR2/3	846	136	2791.42
ELKPLQVQGHVQLRIMENYCLXATKQK-NGDHXXGAR-a3-b6	TT21B	WDR19	inter-protein xl	CutOff_BR2/3	1112	1171	4397.148
EQRKGAEEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3/4	104	1171	2124.064
HDKAREVIDR-NGDHXXGAR-a3-b6	TT21B	WDR19	inter-protein xl	CutOff_BR1/2	126	1171	2440.173
KMGEIADAIK-MKRIFSLLEK-a1-b2	TT21B	WDR19	inter-protein xl	CutOff_BR2/4	575	2	2476.397
RCGKLEDVPR-SKIDAK-a4-b2	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3/4	1001	1225	2027.089
SCCKAYEYMGYMEK-DAKYLFR-a4-b3	TT21B	WDR19	inter-protein xl	CutOff_BR2/4	1232	1090	2997.345
XEKLGDALQQARELOAR-IDAKYK-a3-b6	TT21B	WDR19	inter-protein xl	CutOff_BR2/4	846	1231	3259.74
YDXAEDLLKRLR-KIPVLGK-a9-b1	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3	1221	131	2589.392
YDMAEDLLKRAKXGRR-a9-b2	TT21B	IFT43	inter-protein xl	CutOff_BR2/4	1221	21	2124.062
DILDKARASLRP-LSGKTGR-a5-b4	TT21B	IF140	inter-protein xl	CutOff_BR1/3	1309	276	2209.24
EQAYTDAALNYEXAWKYSNR-MALVKR-a16-b5	TT21B	IF140	inter-protein xl	CutOff_BR3/4	1259	1340	3293.56
FFSMAEKNSR-SIKLIK-a7-b3	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	1014	796	2210.206
FFSMAEKNSR-KYTOAGNK-a7-b1	TT21B	IF140	inter-protein xl	CutOff_BR2/3/4	1014	1211	2434.223
KTEVDTSHR-VKLSGKTGR-a1-b2	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	608	272	2154.183
KTEVDTSHR-VKLSGKTGR-a1-b6	TT21B	IF140	inter-protein xl	CutOff_BR3/4	608	276	2154.17
LAFNYLKAK-AEEVXVK-a7-b6	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	1278	270	2153.192
VKEALKWYK-KYTOAGNK-a2-b1	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	341	1211	2210.218

DKMLYITCFR-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	CutOff_BR1/2/3/4	705	1236	2271.04
EKMADYILKHR-KIAIYR-a2-b1	TT21B	IF122	inter-protein xl	CutOff_BR1/4	693	393	2303.274
ISDGGKOGHVLK-DVLIKR-a6-b5	TT21B	IF122	inter-protein xl	CutOff_BR1/3	142	1184	2148.238
KPDNYXTLR-LDKAER-a1-b3	TT21B	IF122	inter-protein xl	CutOff_BR2/3/4	982	802	2108.067
LKWYDKAEK-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	CutOff_BR1/2/4	803	1236	2105.019
RIGASTSK-HAYDKLR-a7-b5	TT21B	IF122	inter-protein xl	CutOff_BR1/3	603	1021	1986.087
SCCKAYEYMGYIMEK-FHEAAKLYKR-a4-b6	TT21B	IF122	inter-protein xl	CutOff_BR1/2/4	1232	711	3331.563
SCCKAYEYMGYIXEK-KLDKAER-a4-b4	TT21B	IF122	inter-protein xl	CutOff_BR1/2/3/4	1232	802	2944.338
VKEALKWYK-SVSKHK-a2-b4	TT21B	IF122	inter-protein xl	CutOff_BR1/2/4	341	128	1986.136
CQVLLAKVYSKMEK-GSKPLRVKK-a11-b8	TT21B	WDR35	inter-protein xl	CutOff_BR1/2	843	968	2845.654
DGTLASKXGK-XFFYLSKK-a7-b7	TT21B	WDR35	inter-protein xl	CutOff_BR1/4	761	7	2239.126
EKKADIYLK-GSKPLR-a2-b3	TT21B	WDR35	inter-protein xl	CutOff_BR1/2/3/4	693	963	1920.053
HFNKARK-KEGRER-a4-b1	TT21B	WDR35	inter-protein xl	CutOff_BR2/3	1048	473	1810.99
IAKXNWNDAIEEFK-WAKDNPDILFAMXEK-a3-b3	TT21B	WDR35	inter-protein xl	CutOff_BR2/3/4	1185	603	3772.72
KPDNYMTLSR-ALIEKVGIK-a1-b5	TT21B	WDR35	inter-protein xl	CutOff_BR1/3	982	680	2331.281
SCCKAYEYMGYIXEK-CKDYQGIK-a4-b2	TT21B	WDR35	inter-protein xl	CutOff_BR2/3/4	1232	723	3096.365
AGLKNFAFSAAMLXRPYR-K4-155	WDR19		monolink	CutOff_BR2/4	1207	n/a	2430.223
AGLKNFAFSAAMLXRPYRSK-K22-156	WDR19		monolink	CutOff_BR2/3/4	1225	n/a	2646.306
AGLKNFAFSAAMLXRPYRSK-K4-156	WDR19		monolink	CutOff_BR2/3/4	1207	n/a	2646.307
DAKYLFRLYXALK-K13-155	WDR19		monolink	CutOff_BR1/4	1100	n/a	1801.969
DTIQGAKVILAGSTK-K7-156	WDR19		monolink	CutOff_BR1/4	595	n/a	1656.927
IHVKNGDHXXGAR-K10-155	WDR19		monolink	CutOff_BR2/4	1171	n/a	1632.847
KISDCTQYLRTTEEL-K1-156	WDR19		monolink	CutOff_BR2/4	1328	n/a	2039.974
RGVNOALKHPSR-K8-156	WDR19		monolink	CutOff_BR1/2/3/4	836	n/a	1517.832
SEPSNMQFFLMKMDDR-K12-156	WDR19		monolink	CutOff_BR3/4	190	n/a	2130.96
SSCYLWDANTNKTSQDNGMR-K13-155	WDR19		monolink	CutOff_BR1/4	88	n/a	2728.272
VGDLLPHVSSPKIHLQYAK-K12-155	WDR19		monolink	CutOff_BR3/4	896	n/a	2256.265
VGDLLPHVSSPKIHLQYAK-K19-155	WDR19		monolink	CutOff_BR2/3/4	903	n/a	2256.262
AGLKNFAFSAAMLXRPYRSK-YDXAEDLLKRCR-a4-b9	WDR19	TT21B	inter-protein xl	CutOff_BR1/3/4	1207	1221	4310.122
ETOSLDGAKXXVAR-HFNKAR-a9-b4	WDR19	TT21B	inter-protein xl	CutOff_BR2/3	955	1048	2330.178
IHLQYAKAK-DILDKAR-a7-b5	WDR19	TT21B	inter-protein xl	CutOff_BR2/3	903	1309	2038.165
IHVKNGDHMK-FFSAEKR-a4-b7	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1165	1014	2330.175
IHVKNGDHMK-KTEVDTSR-a4-b1	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/4	1165	608	2387.199
IHVKNGDHMK-CGKLEDVPR-a4-b3	WDR19	TT21B	inter-protein xl	CutOff_BR1/2	1165	1001	2404.215
KIEGMVR-LKWYDK-a1-b2	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1233	803	1821
KISDCTQYLRTTEEL-YDMAEDLLK-a1-b9	WDR19	TT21B	inter-protein xl	CutOff_BR2/4	1328	1221	3274.548
SEPSNMQFFLMKMDDR-KXGEIADAIK-a12-b1	WDR19	TT21B	inter-protein xl	CutOff_BR1/2	190	575	3203.479
SEPSNMQFFLMKXDDR-MADYILKHR-a12-b7	WDR19	TT21B	inter-protein xl	CutOff_BR1/2	190	700	3274.555
SEPSNMQFFLXKMDDR-KMGEIADAIK-a12-b1	WDR19	TT21B	inter-protein xl	CutOff_BR1/3	190	575	3203.481
IFDRHGQKR-NGDHMKGAR-a8-b6	WDR19	WDR19	intra-protein xl	CutOff_BR2/4	47	1171	2278.153

LYMALKQYR-KIEGMVR-a6-b1	WDR19	WDR19	WDR19	intra-protein xl	CutOff_BR3/4	1100	1233	2154.177
TSAAESISVVLGKK-IHVKNGDHMK-a14-b4	WDR19	WDR19	WDR19	intra-protein xl	CutOff_BR1/2/4	208	1165	2851.516
VATCGDNCKIQDLVDLK-KIEGMVR-a10-b1	WDR19	WDR19	WDR19	intra-protein xl	CutOff_BR1/3/4	299	1233	3030.525
NGDHXKGARMILR-AKXGRR-a6-b2	WDR19	IFT43	IFT43	inter-protein xl	CutOff_BR1/3	1171	21	2385.222
HPVSVKK-KEPEIKK-a6-b1	WDR19	IF140	IF140	inter-protein xl	CutOff_BR2/4	518	1261	1820.999
KISDCTQYLRTTEEL-REKAHCSCR-a1-b4	WDR19	IF140	IF140	inter-protein xl	CutOff_BR1/2	1328	554	3295.529
LYMALKQYR-MALVKR-a6-b5	WDR19	IF140	IF140	inter-protein xl	CutOff_BR2/4	1100	1340	2039.146
SEPSNMQFLMKXDDR-QGSYHLATKK-a12-b9	WDR19	IF140	IF140	inter-protein xl	CutOff_BR1/4	190	1210	3260.503
YKEAVAYENAK-ALLKSGDTEK-a2-b4	WDR19	IF140	IF140	inter-protein xl	CutOff_BR2/3	912	1227	2582.328
DTGPELRLPMLAQNLLIKR-RDVLIKR-a18-b6	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/4	660	1184	3249.765
ETQSLDGAKXVAR-CKDDPGP-a9-b2	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/2/3/4	955	1236	2346.062
GLYYDKAASVYIRSK-QIGKDR-a6-b4	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/4	871	275	2586.365
IDAKYK-NIKEPK-a4-b3	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/4	1229	762	1601.913
IHVKNGDHXKGAR-NKNGEEK-a10-b2	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/3	1171	161	2433.238
LFPVDDKCR-CKDDPGP-a8-b2	WDR19	IF122	IF122	inter-protein xl	CutOff_BR1/2/4	475	1236	2144.985
QTQVRSEPSNMQFLMKXDDR-KLDKAEK-a17-b1	WDR19	IF122	IF122	inter-protein xl	CutOff_BR2/4	190	799	3599.721
NSSLTLTGETSSAKLPRCR-KIEGMVR-a14-b1	IFT43	WDR19	WDR19	inter-protein xl	CutOff_BR1/4	54	1233	3046.613
NSSLTLTGETSSAKLPRCR-EAKAHCSKR-a14-b3	IFT43	IF140	IF140	inter-protein xl	CutOff_BR1/3	54	554	3332.581
NSSLTLTGETSSAKLPR-LXFKIADEEAKK-a14-b11	IFT43	WDR35	WDR35	inter-protein xl	CutOff_BR2/3	54	959	3336.752
QGWAGDSVKASK-IADEEAKK-a10-b7	IFT43	WDR35	WDR35	inter-protein xl	CutOff_BR1/4	69	959	2330.177
AAVSGDEKALDMFNWKK-K16-156	IF140			monolink	CutOff_BR1/2/3/4	187	n/a	2065.018
AAVSGDEKALDMFNWKK-K17-155	IF140			monolink	CutOff_BR1/2/3	188	n/a	2064.038
AAVSGDEKALDMFNWKK-K17-156	IF140			monolink	CutOff_BR1/2/3/4	188	n/a	2065.019
AAVSGDEKALDMFNWKK-K8-155	IF140			monolink	CutOff_BR1/2/3/4	179	n/a	2064.041
ADNSPDKICFYDVEMDTTFVDFK-K8-155	IF140			monolink	CutOff_BR2/4	598	n/a	3097.374
ADNSPDKICFYDVEMDTTFVDFK-K25-156	IF140			monolink	CutOff_BR1/2/3/4	615	n/a	3114.381
ALSYEKSDTHR-K7-155	IF140			monolink	CutOff_BR1/3	926	n/a	1623.787
AMRALLKSGDTEK-K7-156	IF140			monolink	CutOff_BR1/2/3/4	1227	n/a	1574.839
AAQAFKNAIRLCK-K12-156	IF140			monolink	CutOff_BR3/4	1047	n/a	1574.855
AAQAFKNAIRLCK-K5-156	IF140			monolink	CutOff_BR1/3	1040	n/a	1574.852
AVXLYHKAGHFSK-K7-156	IF140			monolink	CutOff_BR1/2/3/4	1090	n/a	1659.842
GENMNCVCYCKVKGLAAGTDR-K13-156	IF140			monolink	CutOff_BR1/2/3/4	336	n/a	2671.247
GENXNCVCYCKVKGLAAGTDR-K11-156	IF140			monolink	CutOff_BR1/2/3	334	n/a	2687.195
GVQMDRAVMLYHKAGHFSK-K13-155	IF140			monolink	CutOff_BR1/2/3/4	1090	n/a	2329.194
GVQMDRAVMLYHKAGHFSK-K13-156	IF140			monolink	CutOff_BR1/2/3/4	1090	n/a	2330.176
GVQMDRAVMLYHKAGHFSK-K19-155	IF140			monolink	CutOff_BR1/2/4	1096	n/a	2329.194
GVQMDRAVMLYHKAGHFSK-K19-156	IF140			monolink	CutOff_BR1/2/3/4	1096	n/a	2330.175
GVQMDRAVXLYHKAGHFSK-K19-155	IF140			monolink	CutOff_BR1/3	1096	n/a	2345.175
ICFYDVEMDTTFVDFKFTGQIDR-K17-155	IF140			monolink	CutOff_BR1/3/4	615	n/a	2953.369
ICFYDVEMDTTFVDFKFTGQIDRR-K17-156	IF140			monolink	CutOff_BR1/2/3/4	615	n/a	3126.446

MLSEDLPSLELVNKM-K17-156	IF140			monolink	CutOff_BR1/4	953	n/a	2165.082
RPLRDFVGLGEDCDKATR-K14-156	IF140			monolink	CutOff_BR2/4	770	n/a	2203.102
AAVSGDEKALDMFNWK-XEKLGDALTAQQAR-a8-b3	IF140	TT21B		inter-protein xl	CutOff_BR1/4	179	846	3578.766
AAVSGDEKALDXFNWK-AKLEPGFYCK-a8-b2	IF140	TT21B		inter-protein xl	CutOff_BR2/3	179	1020	3274.545
AAVSGDEKALDXFNWK-FFSXAERK-a8-b7	IF140	TT21B		inter-protein xl	CutOff_BR1/2/4	179	1014	2965.397
AAEVMKVK-VKEQRK-a6-b2	IF140	TT21B		inter-protein xl	CutOff_BR1/4	270	100	1857.036
CSGGSTISILPSKADNSPSK-VYSXKEK-a14-b4	IF140	TT21B		inter-protein xl	CutOff_BR1/2	590	843	3274.547
DFVGLGEDCDKATR-IGASTKSKDR-a10-b8	IF140	TT21B		inter-protein xl	CutOff_BR3/4	770	605	2724.312
EAKAHCSCR-FFSXAERK-a3-b7	IF140	TT21B		inter-protein xl	CutOff_BR1/2	554	1014	2286.045
EAKAHCSCR-GKEPYTKK-a3-b7	IF140	TT21B		inter-protein xl	CutOff_BR1/4	554	162	2205.064
FGFEKGNXCVCYCK-EALKWYK-a5-b4	IF140	TT21B		inter-protein xl	CutOff_BR2/3	323	345	3132.411
ICFYDVMEDTVTFDFKFTGQIDR-ELKQTVQGHVQLR-a17-b3	IF140	TT21B		inter-protein xl	CutOff_BR1/4	615	1112	4568.277
VAMWRKVPDFLGSFGAEGK-KGAGEK-a6-b1	IF140	TT21B		inter-protein xl	CutOff_BR1/2	353	104	2770.442
XTVAKDSSDLPEESR-EALKWYK-a5-b4	IF140	TT21B		inter-protein xl	CutOff_BR1/4	1178	345	2754.327
XTVAKDSSDLPEESRR-CGKLEDVPR-a5-b3	IF140	TT21B		inter-protein xl	CutOff_BR2/4	1178	1001	3046.444
YYEEKGVQMDR-EQRKGAGEK-a5-b4	IF140	TT21B		inter-protein xl	CutOff_BR1/2	1077	104	2556.223
YYEEKGVQXDR-MGKALIK-a5-b3	IF140	TT21B		inter-protein xl	CutOff_BR3/4	1077	764	2330.179
YYEEKGVQXDR-YDMAEDLLKR-a5-b9	IF140	TT21B		inter-protein xl	CutOff_BR1/3	1077	1221	2823.282
AAVSGDEKALDXFNWK-DAKYLFR-a16-b3	IF140	WDR19		inter-protein xl	CutOff_BR2/4	187	1090	2974.484
GENMNCVCYCK-IHVKINGDHMKGAR-a11-b4	IF140	WDR19		inter-protein xl	CutOff_BR2/3	334	1165	3260.501
KEEYQTAYRLEEMR-KIEGXVR-a1-b1	IF140	WDR19		inter-protein xl	CutOff_BR2/3	1389	1233	2977.466
KPEEADREDEVEPGCHHIPQXVR-GITGDNKEHDEACLAGVAQMSIR-a1-b7	IF140	WDR19		inter-protein xl	CutOff_BR2/3	733	807	5469.483
XTVAKDSSDLPEESR-NGDHMKGARXLIR-a5-b6	IF140	WDR19		inter-protein xl	CutOff_BR2/3	1178	1171	3331.567
AKSPDQETRLAQLQSR-NSSLTLGETSSAKLPR-a2-b14	IF140	IF143		inter-protein xl	CutOff_BR1/4	1320	54	3839.06
DFVGLGEDCDKATR-KASEEIEDFR-a10-b1	IF140	IF143		inter-protein xl	CutOff_BR1/4	770	76	2885.333
LDVAKVCLGNXGHAR-AKMGRR-a5-b2	IF140	IF143		inter-protein xl	CutOff_BR1/2	821	21	2511.309
XTVAKDSSDLPEESRR-KASEEIEDFR-a5-b1	IF140	IF143		inter-protein xl	CutOff_BR3/4	1178	76	3196.519
AAEVMKVKLSGK-AEEVMKVK-a6-b6	IF140	IF140		intra-protein xl	CutOff_BR1/2/4	270	270	2388.297
AAEVMKVKLSGK-AEEVMKVK-a8-b6	IF140	IF140		intra-protein xl	CutOff_BR2/3/4	272	270	2388.295
FGFEKGNMNCVCYCKV-LSGKTGR-a16-b4	IF140	IF140		intra-protein xl	CutOff_BR3/4	334	276	3124.44
GENMNCVCYCKV-LSGKTGR-a11-b4	IF140	IF140		intra-protein xl	CutOff_BR2/3/4	334	276	2516.151
KEEYQTAYR-MALVKR-a1-b5	IF140	IF140		intra-protein xl	CutOff_BR2/4	1389	1340	2041.085
LDVAKVCLGNXGHARGAR-EAKAHCSCR-a5-b3	IF140	IF140		intra-protein xl	CutOff_BR1/2/3/4	821	554	3195.526
SIKLIKSEAVWENXAR-LSGKTGR-a6-b4	IF140	IF140		intra-protein xl	CutOff_BR3/4	799	276	2745.465
XALVKRFFIAR-KYTOAGNK-a5-b1	IF140	IF140		intra-protein xl	CutOff_BR1/2/3/4	1340	1211	2394.316
YYEEKGVQMDR-LKAMRALLK-a5-b2	IF140	IF140		intra-protein xl	CutOff_BR1/3	1077	1220	2597.394
AAEVMKVK-SVSKHK-a6-b4	IF140	IF122		inter-protein xl	CutOff_BR2/3	270	128	1754.947
ALLKSGDTEK-AFKAGR-a4-b4	IF140	IF122		inter-protein xl	CutOff_BR1/3/4	1227	887	1984.063
AVXLYHKAGHFSK-RCKDDPGP-a7-b3	IF140	IF122		inter-protein xl	CutOff_BR1/2/3/4	1090	1236	2585.246
LKSEAVWENMAR-AFKAGR-a3-b4	IF140	IF122		inter-protein xl	CutOff_BR1/2	799	887	2469.282



NIIGFYTKGR-RCKDDPGP-a8-b3	IF140	IF122	inter-protein xl	CutOff_BR2/3	1275	1236	2249.132
AEEVXVK-NAQRGKVK-a6-b6	IF140	WDR35	inter-protein xl	CutOff_BR2/3	270	992	1986.085
ALLKSGDTEKITFFASVSR-GSKPLR-a4-b3	IF140	WDR35	inter-protein xl	CutOff_BR2/4	1227	963	2863.583
DFVGLDCCDKATR-DVWDMKWAK-a10-b6	IF140	WDR35	inter-protein xl	CutOff_BR1/4	770	600	2840.289
FGFEKGNMNCVCYCK-NNEKYVYVK-a5-b4	IF140	WDR35	inter-protein xl	CutOff_BR1/2	323	368	3463.588
FGFEKGNXNCVCYCKV-KSKPLR-a5-b3	IF140	WDR35	inter-protein xl	CutOff_BR3/4	323	963	3079.466
GENXNCVCYCKVK-NKSVVR-a11-b2	IF140	WDR35	inter-protein xl	CutOff_BR1/2/3/4	334	115	2516.15
GRVQGTPLLKHEYGK-GSKPLR-a10-b3	IF140	WDR35	inter-protein xl	CutOff_BR1/4	145	963	2476.394
REAKHCSCR-CKDYQGIK-a4-b2	IF140	WDR35	inter-protein xl	CutOff_BR3/4	554	723	2422.118
AKPFHDEELVPLCYR-K2-156	IF122		monolink	CutOff_BR3/4	1045	n/a	2116.021
DRALNFDPCISYFTKGEYILLGGSDK-K27-155	IF122		monolink	CutOff_BR1/2/3/4	302	n/a	3293.579
DTVYVAYAKDGR-K13-156	IF122		monolink	CutOff_BR1/4	66	n/a	1800.877
ELAXEALGLDFETAKK-K16-156	IF122		monolink	CutOff_BR1/2/3/4	663	n/a	2066.007
ELAXEALGLDFETAKK-K17-156	IF122		monolink	CutOff_BR1/2/3/4	664	n/a	2065.999
ETKXLTQADWAR-K3-155	IF122		monolink	CutOff_BR1/4	748	n/a	1860.99
LYRSGHENLALEMYDLCXFEYAK-K3-156	IF122		monolink	CutOff_BR1/4	714	n/a	3253.538
XLITKQADWAR-K5-156	IF122		monolink	CutOff_BR1/2/3/4	753	n/a	1503.78
ETKMLITK-VYSKMEK-a3-b4	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	748	843	1984.063
GHKDTVYCVAYAK-TLHMAYSLPGMKR-a3-b12	IF122	TT21B	inter-protein xl	CutOff_BR1/2/3/4	53	596	3136.505
KAFIRVQDLR-GKEPYTK-a1-b2	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	664	157	2204.247
NKNGEEKVK-VKEQRK-a2-b2	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	161	100	1969.1
CLDMSASRKK-KEPEIMK-a9-b1	IF122	IF140	inter-protein xl	CutOff_BR1/2	527	1261	2206.105
HAYDKLR-KEPEIMK-a5-b1	IF122	IF140	inter-protein xl	CutOff_BR1/4	1021	1261	1912.995
KLDSFGYAAETLYLKMGLK-YYEEKGVQMDR-a14-b5	IF122	IF140	inter-protein xl	CutOff_BR2/4	830	1077	3653.772
CLDXSARKK-QIGKDR-a9-b4	IF122	IF122	intra-protein xl	CutOff_BR1/2/4	527	275	2064.038
DTPSGISKVK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	CutOff_BR1/3/4	995	1021	2070.113
LVETKDSIGDEDPFTAK-MLITKQADWAR-a5-b5	IF122	IF122	intra-protein xl	CutOff_BR2/3	1141	753	3333.681
MLITKQADWAR-SVSKHK-a5-b4	IF122	IF122	intra-protein xl	CutOff_BR1/2/3/4	753	128	2154.176
QSKALGAYR-NIKEPK-a3-b3	IF122	IF122	intra-protein xl	CutOff_BR1/3	1007	762	1858.034
AFHKAGRQR-DVWDMKWAK-a4-b6	IF122	WDR35	inter-protein xl	CutOff_BR2/3/4	887	600	2385.223
ALNFDPCISYFTKGEYILLGGSDK-KPELDSLMEGGEGK-a14-b1	IF122	WDR35	inter-protein xl	CutOff_BR2/4	291	1123	4494.124
HAYDKLR-NKSVVR-a5-b2	IF122	WDR35	inter-protein xl	CutOff_BR2/4	1021	115	1740.977
ANYFFDAKLXFK-K13-156	WDR35		monolink	CutOff_BR1/2/3/4	952	n/a	1736.861
DNPDLFAMMEKTRMYVFR-K11-156	WDR35		monolink	CutOff_BR1/2/4	614	n/a	2419.159
DPEHPNKDYLINEIRSLR-K7-156	WDR35		monolink	CutOff_BR3/4	660	n/a	2511.301
IYHVDDTPSGMSMDGVLDYSKTIQGTGR-K20-155	WDR35		monolink	CutOff_BR2/3/4	498	n/a	3009.419
KANYFFDAAKLMFK-K1-156	WDR35		monolink	CutOff_BR3/4	939	n/a	1848.957
KPELDSLXEGGEGK-K1-156	WDR35		monolink	CutOff_BR2/4	1123	n/a	1660.783
LKSLETLSSEQK-K12-156	WDR35		monolink	CutOff_BR1/3/4	1102	n/a	1517.832
LKSLETLSSEQK-K2-156	WDR35		monolink	CutOff_BR1/2/3/4	1092	n/a	1517.831

QQVEDLALIFTKHTSK-K13-156	WDR35		monolink	CutOff_BR3/4	1115	n/a	2206.105
RDVWDXKWAK-K10-155	WDR35		monolink	CutOff_BR1/2/3/4	603	n/a	1504.763
RDVWDXKWAK-K7-155	WDR35		monolink	CutOff_BR3/4	600	n/a	1504.763
SMSWNADGQICIVYEDGAVIVSGVDGMR-K10-155	WDR35		monolink	CutOff_BR1/4	129	n/a	3294.569
TLDAIELYRKANYFFDAAK-K10-156	WDR35		monolink	CutOff_BR1/2/3	939	n/a	2404.215
TLDAIELYRKANYFFDAAK-K19-156	WDR35		monolink	CutOff_BR1/2	948	n/a	2404.216
VGIKDASQFIEDNPHPR-K4-155	WDR35		monolink	CutOff_BR2/4	684	n/a	2077.052
AVELAKNHSXK-DGTLASKMGK-ag-b7	WDR35	TT21B	inter-protein xl	CutOff_BR1/3	906	761	2387.205
DNPDLFAMMEKTR-KGAGEK-a11-b1	WDR35	TT21B	inter-protein xl	CutOff_BR3/4	614	104	2293.098
DNPDLFAMMEKTR-LKWYDK-a11-b2	WDR35	TT21B	inter-protein xl	CutOff_BR2/4	614	803	2556.222
DVWDMKWAK-AQSQKK-ag-b5	WDR35	TT21B	inter-protein xl	CutOff_BR3/4	600	574	2004.013
DYQGIKVKR-KPDNYMTLSR-ag-b1	WDR35	TT21B	inter-protein xl	CutOff_BR1/3	729	982	2614.375
HTSKDNRKPELDSXEGGEGK-HLAAEICAEIAKHSHAQR-a8-b12	WDR35	TT21B	inter-protein xl	CutOff_BR1/4	1123	892	4484.175
KGSKPLR-WYDKAEK-a4-b4	WDR35	TT21B	inter-protein xl	CutOff_BR2/4	963	807	1860.997
KPELDSXEGGEGK-AQSQKK-a1-b5	WDR35	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1123	574	2331.153
KPELDSXEGGEGK-IGASTKSK-a1-b6	WDR35	TT21B	inter-protein xl	CutOff_BR1/2/4	1123	603	2433.239
VLKLETTQDDAK-FFSMAEKR-a3-b7	WDR35	TT21B	inter-protein xl	CutOff_BR1/2	42	1014	2512.299
WAKDNPDLFAMMEK-FFSMAEKRNSR-a3-b7	WDR35	TT21B	inter-protein xl	CutOff_BR2/3/4	603	1014	3204.525
HTSKDNR-KIEGVXR-a4-b1	WDR35	WDR19	inter-protein xl	CutOff_BR1/2	1119	1233	1841.931
KANYFFDAAK-LYXALKQYR-a1-b6	WDR35	WDR19	inter-protein xl	CutOff_BR2/3/4	939	1100	2512.3
KPELDSXEGGEGK-KKIEGMVR-a1-b2	WDR35	WDR19	inter-protein xl	CutOff_BR2/3	1123	1233	2586.365
KPELDSXEGGEGK-NGDPHXKGARMLIR-a1-b6	WDR35	WDR19	inter-protein xl	CutOff_BR2/3	1123	1171	3156.516
LGKLLSEMK-DAKYLFR-a3-b3	WDR35	WDR19	inter-protein xl	CutOff_BR3/4	736	1090	2154.197
LGKLLSEMKQAEVYGFGR-NGDHMKGAR-a3-b6	WDR35	WDR19	inter-protein xl	CutOff_BR2/4	736	1171	3333.679
LGKLLSEXK-SKIDAKYK-a3-b6	WDR35	WDR19	inter-protein xl	CutOff_BR1/3/4	736	1229	2210.206
AFGTCSKAFIK-SIKLIK-a7-b3	WDR35	IF140	inter-protein xl	CutOff_BR2/4	1086	796	2067.182
DNPDLFAMMEKTR-KYTAQGNK-a11-b1	WDR35	IF140	inter-protein xl	CutOff_BR2/3/4	614	1211	2613.243
DPEHPNKDYLINEIR-GENXVCYCKVK-a7-b11	WDR35	IF140	inter-protein xl	CutOff_BR1/4	660	334	3813.709
LMFKIADEEAK-CKRHDLNKK-a4-b2	WDR35	IF140	inter-protein xl	CutOff_BR2/3	952	870	2614.374
RDVWDMKWAK-KPEEADR-a7-b1	WDR35	IF140	inter-protein xl	CutOff_BR1/3	600	733	2315.127
VGXCEQAVTFLKCSQPK-MCVKTQRIDVAK-a13-b4	WDR35	IF140	inter-protein xl	CutOff_BR1/2	881	813	3654.797
VKLETTQDDAKLR-VKGLLAAGTDRGR-a3-b2	WDR35	IF140	inter-protein xl	CutOff_BR1/3	42	336	3079.735
AFGTCSKAFIKL-XLITQADWAR-a11-b5	WDR35	IF122	inter-protein xl	CutOff_BR3/4	1090	753	2955.572
DNPDLFAMMEKTR-CKDDPGP-a11-b2	WDR35	IF122	inter-protein xl	CutOff_BR1/2/3/4	614	1236	2492.114
KANYFFDAAK-RDVLIKR-a1-b6	WDR35	IF122	inter-protein xl	CutOff_BR2/4	939	1184	2210.215
KISIPNNVK-IKKELVK-a1-b4	WDR35	IF122	inter-protein xl	CutOff_BR1/4	8	388	2166.268
KPELDSXEGGEGK-GHKDTVYCVAYAK-a1-b3	WDR35	IF122	inter-protein xl	CutOff_BR1/2/4	1123	53	3137.513
KPELDSXEGGEGK-NKNGEEKVK-a1-b2	WDR35	IF122	inter-protein xl	CutOff_BR2/4	1123	161	2671.313
RLGKLLSESMK-KLDKAER-a4-b1	WDR35	IF122	inter-protein xl	CutOff_BR2/4	736	799	2257.269
DVWDMKWAK-KGSKPLR-a6-b4	WDR35	WDR35	intra-protein xl	CutOff_BR2/3/4	600	963	2100.109
KPELDSXEGGEGK-KISIPNNVK-a1-b1	WDR35	WDR35	intra-protein xl	CutOff_BR1/2	1123	8	2654.378
KPELDSXEGGEGK-NKSVVR-a1-b2	WDR35	WDR35	intra-protein xl	CutOff_BR2/4	1123	115	2344.18
LKSLETLSSEQK-HTSKDNR-a2-b4	WDR35	WDR35	intra-protein xl	CutOff_BR3/4	1092	1119	2356.22
VKLETTQDDAK-GSKPLR-a3-b3	WDR35	WDR35	intra-protein xl	CutOff_BR3/4	42	963	2154.177

Table 25: Crosslinks of IFT-A (HEK293T) cells after spin column filtration (in at least 3of3)

Id	Protein1	Protein2	XlType	Biological Replicates	AbsPos1	AbsPos2	Mr
EALVHCETDNKIMLELAR-K11-156	TT21B		monolink	CutOff_BR1/2/3/4	919	n/a	2297.151
SCCKAYEYMGYMEKEQAYTDAALNVEYAWK-K31-156	TT21B		monolink	CutOff_BR1/2/3/4	1259	n/a	3988.705
CGKLEDVPR-IRKDILDK-a3-b3	TT21B	TT21B	intra-protein xl	CutOff_BR1/2/4	1001	1304	2210.221
EKXADLYL-RIGASTKSK-a2-b7	TT21B	TT21B	intra-protein xl	CutOff_BR2/3/4	693	603	2210.206
GKEPYTK-AQSQKK-a2-b5	TT21B	TT21B	intra-protein xl	CutOff_BR1/3/4	157	574	1647.891
IXENYCLIMATKQK-IRKDILDK-a11-b3	TT21B	TT21B	intra-protein xl	CutOff_BR1/2/3/4	1134	1304	2782.463
EQRKAGEK-NGDHMKGAR-a4-b6	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3/4	104	1171	2124.064
RCGKLEDVPR-SKIDAK-a4-b2	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3/4	1001	1225	2027.089
YDXAEDLLKRCLR-KIPVLGK-a9-b1	TT21B	WDR19	inter-protein xl	CutOff_BR1/2/3/4	1221	131	2589.392
FFSMAEKRNSR-SIKLIK-a7-b3	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	1014	796	2210.206
FFSXAERNSR-KYQAGNK-a7-b1	TT21B	IF140	inter-protein xl	CutOff_BR2/3/4	1014	1211	2434.223
KTEVDTSR-VKLSGKTGR-a1-b2	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	608	272	2154.183
LAFNYLKAK-AEEVXKVK-a7-b6	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	1278	270	2153.192
VKEALKWYK-KYQAGNK-a2-b1	TT21B	IF140	inter-protein xl	CutOff_BR1/2/3/4	341	1211	2210.218
DKMLYITCFR-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	CutOff_BR1/2/3/4	705	1236	2271.04
KPDNYTLR-LDKAER-a1-b3	TT21B	IF122	inter-protein xl	CutOff_BR2/3/4	982	802	2108.067
LKWYDKAEK-CKDDPGP-a2-b2	TT21B	IF122	inter-protein xl	CutOff_BR1/2/4	803	1236	2105.019
SCCKAYEYMGYMEK-FHEAAKLYKR-a4-b6	TT21B	IF122	inter-protein xl	CutOff_BR1/2/4	1232	711	3331.563
SCCKAYEYMGYMEK-KLDKAER-a4-b4	TT21B	IF122	inter-protein xl	CutOff_BR1/2/3/4	1232	802	2944.338
EKXADLYL-GSKPLR-a2-b3	TT21B	WDR35	inter-protein xl	CutOff_BR1/2/3/4	693	963	1920.053
IAXKNWNAIDAEFEK-WAKNDPLFAMXEK-a3-b3	TT21B	WDR35	inter-protein xl	CutOff_BR2/3/4	1185	603	3772.72
SCCKAYEYMGYMEK-CKDYQGIK-a4-b2	TT21B	WDR35	inter-protein xl	CutOff_BR2/3/4	1232	723	3096.365
AGLKNSAFSFAAMLXRPYRSK-K22-156	WDR19		monolink	CutOff_BR2/3/4	1225	n/a	2646.306
AGLKNSAFSFAAMLXRPYRSK-K4-156	WDR19		monolink	CutOff_BR2/3/4	1207	n/a	2646.307
RGVNLKHPSR-K8-156	WDR19		monolink	CutOff_BR1/2/3/4	836	n/a	1517.832
VGDLLPHVSSPKHLQYAK-K19-155	WDR19		monolink	CutOff_BR2/3/4	903	n/a	2256.262
AGLKNSAFSFAAMLXRPYRSK-YDXAEDLLKRCLR-a4-b9	WDR19	TT21B	inter-protein xl	CutOff_BR1/3/4	1207	1221	4310.122
IHVKNGDHMK-FFSMAEKR-a4-b7	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1165	1014	2330.175
IHVKNGDHMK-KTEVDTSR-a4-b1	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/4	1165	608	2387.199
KIEGMVR-LKWYDK-a1-b2	WDR19	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1233	803	1821
TSAAESISVVLGKK-IHVKNGDHMK-a14-b4	WDR19	WDR19	intra-protein xl	CutOff_BR1/2/4	208	1165	2851.516
VATCGDNCKIQDLVDLIK-KIEGMVR-a10-b1	WDR19	WDR19	intra-protein xl	CutOff_BR1/3/4	299	1233	3030.525
ETQSLDGAKXVAR-CKDDPGP-a9-b2	WDR19	IF122	inter-protein xl	CutOff_BR1/2/3/4	955	1236	2346.062
LFPVAVDKCR-CKDDPGP-a8-b2	WDR19	IF122	inter-protein xl	CutOff_BR1/2/4	475	1236	2144.985
AAVSGDEKALDMFNWKK-K16-156	IF140		monolink	CutOff_BR1/2/3/4	187	n/a	2065.018
AAVSGDEKALDMFNWKK-K17-156	IF140		monolink	CutOff_BR1/2/3/4	188	n/a	2065.019
AAVSGDEKALDMFNWKK-K8-155	IF140		monolink	CutOff_BR1/2/3/4	179	n/a	2064.041
AAVSGDEKALDMFNWKK-K8-156	IF140		monolink	CutOff_BR1/2/3/4	179	n/a	2065.018
ADNSPDSKICFYDVEXTVTVFDFK-K25-156	IF140		monolink	CutOff_BR1/2/3/4	615	n/a	3114.381
AMRALLKSGTEK-K7-156	IF140		monolink	CutOff_BR1/2/3/4	1227	n/a	1574.839
GENMNCVCYKVKGLLAAGTDR-K13-156	IF140		monolink	CutOff_BR1/2/3/4	336	n/a	2671.247

GVQMDRAVMVLYHKAGHFSK-K13-155	IF140		monolink	CutOff_BR1/2/3	1090	n/a	2329.194
GVQMDRAVMVLYHKAGHFSK-K13-156	IF140		monolink	CutOff_BR1/2/3/4	1090	n/a	2330.176
GVQMDRAVMVLYHKAGHFSK-K19-155	IF140		monolink	CutOff_BR1/2/4	1096	n/a	2329.194
GVQMDRAVMVLYHKAGHFSK-K19-156	IF140		monolink	CutOff_BR1/2/3/4	1096	n/a	2330.175
ICFYDVEMDTVTFDKTQQIDR-K17-155	IF140		monolink	CutOff_BR1/3/4	615	n/a	2953.369
ICFYDVEMDTVTFDKTQQIDR-K17-156	IF140		monolink	CutOff_BR1/2/3/4	615	n/a	3126.446
AAVSGDEKALDXFNWK-FFSXAERK-a8-b7	IF140	TT21B	inter-protein xl	CutOff_BR1/2/4	179	1014	2965.397
AEEMVKVLSGK-AEEMVKV-k-a6-b6	IF140	IF140	intra-protein xl	CutOff_BR1/2/4	270	270	2388.297
AEEMVKVLSGK-AEEMVKV-k-a8-b6	IF140	IF140	intra-protein xl	CutOff_BR2/3/4	272	270	2388.295
GENMNCVCYKVK-LSGKTGR-a11-b4	IF140	IF140	intra-protein xl	CutOff_BR2/3/4	334	276	2516.151
LDVAKVCLGNXGHARGAR-EAKAHCSER-a5-b3	IF140	IF140	intra-protein xl	CutOff_BR1/2/3/4	821	554	3195.526
XALVKRFIQR-KYTOAGNK-a5-b1	IF140	IF140	intra-protein xl	CutOff_BR1/2/3/4	1340	1211	2394.316
ALLKSGDTEK-AFHAGR-a4-b4	IF140	IF122	inter-protein xl	CutOff_BR1/3/4	1227	887	1984.063
AVXLYHKAGHFSK-RCKDDPGP-a7-b3	IF140	IF122	inter-protein xl	CutOff_BR1/2/3/4	1090	1236	2585.246
GENXNCVCYKVK-NKSVVR-a11-b2	IF140	WDR35	inter-protein xl	CutOff_BR1/2/3/4	334	115	2516.15
DRALNFDPCISYFTKGEYLLGGSDK-K27-155	IF122		monolink	CutOff_BR1/2/3/4	302	n/a	3293.579
ELAXEALGLDFETAKK-K16-156	IF122		monolink	CutOff_BR1/2/3/4	663	n/a	2066.007
ELAXEALGLDFETAKK-K17-156	IF122		monolink	CutOff_BR1/2/3/4	664	n/a	2065.999
XLITKQADWAR-K5-156	IF122		monolink	CutOff_BR1/2/3/4	753	n/a	1503.78
ETKMLTK-VYSKMEK-a3-b4	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	748	843	1984.063
GHKDVTVCVAYAK-TLHXAMSLPGMKR-a3-b12	IF122	TT21B	inter-protein xl	CutOff_BR1/2/3/4	53	596	3136.514
KAFIRVQDLR-GKEPYTK-a1-b2	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	664	157	2204.247
NKNGEEKVK-VKEQRK-a2-b2	IF122	TT21B	inter-protein xl	CutOff_BR1/2/4	161	100	1969.1
CLDXSARRK-QJGKDR-a9-b4	IF122	IF122	intra-protein xl	CutOff_BR1/2/4	527	275	2064.038
DTPSGISKVK-HAYDKLR-a8-b5	IF122	IF122	intra-protein xl	CutOff_BR1/3/4	995	1021	2070.113
MLITKQADWAR-SVSKHK-a5-b4	IF122	IF122	intra-protein xl	CutOff_BR1/2/3/4	753	128	2154.176
AFHKAGRQR-DVWDMKWAK-a4-b6	IF122	WDR35	inter-protein xl	CutOff_BR2/3/4	887	600	2385.223
ANYFFDAAKLXFK-K13-156	WDR35		monolink	CutOff_BR1/2/3/4	952	n/a	1736.861
DNPDLFAMMEKTRMYVFR-K11-156	WDR35		monolink	CutOff_BR1/2/4	614	n/a	2419.159
IYHVDDTPSGMDGVLDSYKTIQGR-K20-155	WDR35		monolink	CutOff_BR2/3/4	498	n/a	3009.419
LKSLETLSSEQK-K12-156	WDR35		monolink	CutOff_BR1/3/4	1102	n/a	1517.832
LKSLETLSSEQK-K2-156	WDR35		monolink	CutOff_BR1/2/3/4	1092	n/a	1517.831
RDVWDXKWAK-K10-155	WDR35		monolink	CutOff_BR1/2/3/4	603	n/a	1504.763
TLDAIELYRKANYFFDAAK-K10-156	WDR35		monolink	CutOff_BR1/2/3	939	n/a	2404.215
KPELDSLXEGGEGK-AQSQKK-a1-b5	WDR35	TT21B	inter-protein xl	CutOff_BR1/2/3/4	1123	574	2331.153
KPELDSLXEGGEGK-IGASTKSK-a1-b6	WDR35	TT21B	inter-protein xl	CutOff_BR1/2/4	1123	603	2433.239
WAKDNPDLFAMMEK-FFSMAEKRNSR-a3-b7	WDR35	TT21B	inter-protein xl	CutOff_BR2/3/4	603	1014	3204.525
KANYFFDAAK-LYXALKQYR-a1-b6	WDR35	WDR19	inter-protein xl	CutOff_BR2/3/4	939	1100	2512.3
LGKLSSESXK-SKIDAKYK-a3-b6	WDR35	WDR19	inter-protein xl	CutOff_BR1/3/4	736	1229	2210.206
DNPDLFAMMEKTR-KYTOAGNK-a11-b1	WDR35	IF140	inter-protein xl	CutOff_BR2/3/4	614	1211	2613.243
DNPDLFAMMEKTR-CKDDPGP-a11-b2	WDR35	IF122	inter-protein xl	CutOff_BR1/2/3/4	614	1236	2492.114
KPELDSLMEGGEGK-GHKDVTVCVAYAK-a1-b3	WDR35	IF122	inter-protein xl	CutOff_BR1/2/4	1123	53	3137.513
DVWDMKWAK-KGSKPLR-a6-b4	WDR35	WDR35	intra-protein xl	CutOff_BR2/3/4	600	963	2100.109

---

## 7.2 Figure Index

Figure 1: Structure of a cilium.....	16
Figure 2: Ciliogenesis.....	17
Figure 3: Structure of the human eye with photoreceptors.....	19
Figure 4: Intraflagellar transport.....	20
Figure 5: Workflow of a bottom-up mass spectrometry approach.....	23
Figure 6: Identification of cross-linked peptides generating an isotopic signature.....	26
Figure 7: Stable isotope labelling of amino acids in cell culture (SILAC).....	53
Figure 8: Generation of stable cell lines using the Flp-In system.....	55
Figure 9: Gene editing using the CRISPR/Cas9 system.....	56
Figure 10: Protein sequence of (N)/(C)-SF-TAP tag.....	60
Figure 11: Tandem Affinity Purification (SF-TAP).....	61
Figure 12: Workflow for absolute quantification.....	66
Figure 13: Generating equimolar standard mixture performing EtEP.....	67
Figure 14: Workflow of chemical crosslinking of purified IFT-A.....	69
Figure 15: LC gradient profile.....	72
Figure 16: Selected Reaction Monitoring – SRM.....	73
Figure 17: Parallel Reaction Monitoring – PRM.....	74
Figure 18: Data analysis using Skyline.....	76
Figure 19: Experimental workflow.....	77
Figure 20: Validation of generated Flp-In monoclonal lines via PCR.....	79
Figure 21: Validation of bait protein expression of Flp-In monoclonal cell lines.....	80
Figure 22: Proteotypic peptides of IFT-A.....	82
Figure 23: Release of Equalizer Peptide after efficient tryptic proteolysis.....	84
Figure 24: Quantification of synthetic standard peptides using ‘heavy’ labelled EP.....	86
Figure 25: Quantification of synthetic peptides with adjusted amounts of spiked EP.....	87
Figure 26: Determination of the CE optimum for transitions using Skyline.....	91

---

Figure 27: Chosen baits for the purification of IFT-A .....	94
Figure 28: Validation of an entire purification of IFT-A using silver staining .....	95
Figure 29: Stoichiometry of purified IFT-A using different baits performing SRM.....	99
Figure 30: PRM-based stoichiometry of purified IFT-A using three different baits .....	102
Figure 31: Reproducibility of absolute quantification using different baits.....	103
Figure 32: Validation of generated CRISPPR clone mixtures via PCR.....	105
Figure 33: Generated mutations in IFT43 using CRISPR/Cas9 system.....	106
Figure 34: Generated mutations in WDR19 using CRISPR/Cas9 system .....	107
Figure 35: Effect on IFT-A composition of all identified single clones .....	108
Figure 36: Effect of selected single clones on complex composition of purified IFT-A .....	109
Figure 37: IFT-A complex composition within human fibroblasts.....	111
Figure 38: Complex composition of IFT-A during ciliary assembly and disassembly .....	113
Figure 39: Confluence-related complex composition of IFT-A .....	114
Figure 40: Number of identified crosslinks with or without DTT .....	116
Figure 41: Optimization of DSS concentration for chemical crosslinking .....	117
Figure 42: SEC or CutOff spin column filtration for the purification of cross-linked peptides ....	119
Figure 43: SEC and spin column filtration of cross-linked peptide mixture from HEK293Ts ....	121
Figure 44: Identified links within IFT-A from Flp-In (N)-SF TULP3 after SEC.....	124
Figure 45: Identified links of IFT-A (Flp-In (N)-SF TULP3) after SEC or spin column filtration..	127
Figure 46: Identified links of IFT-A from HEK293T cells (in at least 2 of 4 replicates).....	131
Figure 47: Identified links of IFT-A from HEK293T cells (in at least 3 of 4 replicates).....	133
Figure 48: Scheme of known isoforms of IFT-A components .....	157
Figure 49: Alignment of IFT122 isoforms and chosen representative peptides.....	158
Figure 50: Alignment of IFT140 isoforms and chosen representative peptides.....	159
Figure 51: Alignment of IFT43 isoforms and chosen representative peptides.....	160
Figure 52: Alignment of TTC21B isoforms and chosen representative peptides.....	161
Figure 53: Alignment of WDR19 isoforms and chosen representative peptides .....	162

---

Figure 54: Alignment of WDR35 isoforms and chosen representative peptides .....	163
---	-----

### 7.3 Table Index

Table 1: Oligonucleotides for gene editing using CRISPR/Cas9 system.....	42
Table 2: Sequencing and PCR Primers .....	42
Table 3: Vectors .....	43
Table 4: Constructs .....	43
Table 5: Primary Antibodies .....	44
Table 6: Secondary Antibodies.....	44
Table 7: Software .....	44
Table 8: Databases .....	45
Table 9: PCR program.....	49
Table 10: Calculation for an equimolar standard mixture .....	88
Table 11: Calculation for a standard mixture with adjusted amount of synthetic peptides .....	89
Table 12: Transition list for SRM with optimised CE values .....	92
Table 13: Parameter settings for PRM.....	93
Table 14: One-step affinity purification of generated Flp-In monoclonal lines .....	95
Table 15: Absolute quantification of purified IFT-A using Flp-In LCA5 performing SRM.....	98
Table 16: Absolute quantification of IFT-A using Flp-In (N)-SF-TULP3 performing PRM .....	101
Table 17: Absolute quantification of IFT-A using Flp-In (N)-SF-LCA5 performing SRM .....	155
Table 18: Absolute quantification of IFT-A using Flp-in (N)-SF-TULP3 performing PRM .....	156
Table 19: Identified links of IFT-A from Flp-In (N)-SF-TULP3 cells after SEC (in at least 2of3).....	164
Table 20: Identified links of IFT-A from Flp-In (N)-SF-TULP3 cells after SEC (in at least 3of3).....	169
Table 21: Links of IFT-A (Flp-In (N)-SF-TULP3) identified after both enrichment methods .....	170
Table 22: Identified links of IFT-A from HEK293T cells after SEC (in at least 2of3) .....	171
Table 23: Crosslinks of IFT-A from HEK293T cells after SEC (in at least 3of3) .....	176
Table 24: Crosslinks of IFT-A (HEK293T) cells after spin column filtration (in at least 2of3) ...	177

Table 25: Crosslinks of IFT-A (HEK293T) cells after spin column filtration (in at least 3of3)....183



---

## 7.4 Acknowledgements / Danksagung

Nun möchte ich mich an dieser Stelle noch bei all denen bedanken, die mich stets auf diesem Weg begleitet und somit zum Gelingen dieser Arbeit auf unterschiedlichste Weise beigetragen haben.

An erster Stelle möchte ich mich bei Dr. Karsten Boldt für die Betreuung meiner Arbeit bedanken. Die zahlreichen Besprechungen durch die ich wertvolle Ideen und Denkanstöße erhalten habe und die Ermutigung zum selbstständigen Arbeiten haben den wissenschaftlichen Wert dieser Arbeit vorangetrieben und mich dabei unterstützt Probleme eigenständig anzugehen und selbstständig nach Lösungsansätzen zu suchen. Für diese wertvolle Kombination aus Ermutigung und Unterstützung möchte ich mich herzlichst bedanken.

Bei Prof. Dr. Marius Ueffing möchte ich mich ganz herzlich bedanken, da er mir die Anfertigung dieser Arbeit in seiner Abteilung ermöglicht und dadurch den Grundstein zum Gelingen dieser Arbeit gelegt hat. Zusätzlich bedanke ich mich bei Ihm als Prüfer und Gutachter meiner Arbeit. Seine konstruktive Kritik im Rahmen von Präsentationen und Besprechungen waren sehr wertvoll für mich und meine Arbeit.

Für die Bereiterklärung als Gutachter dieser Arbeit von Seiten der mathematisch-naturwissenschaftlichen Fakultät möchte ich Prof. Dr. Robert Feil besonders danken.

Ebenfalls bedanken möchte ich mich bei Prof. Dr. Dirk Schwarzer und Prof. Dr. Stefan Stevanović, die als Prüfer dieser Arbeit fungiert haben.

Für die intensive Unterstützung von Dr. Andrea Meixner und Dr. Tina Beyer in molekularbiologischen Fragestellungen möchte ich mich ganz besonders bedanken. Ihr hattet beide immer ein offenes Ohr für mich und wart sehr engagiert und motiviert wenn es um eine Problemlösung ging. Vielen Dank euch beiden dafür.

Dr. Johannes Gloeckner gilt mein besonderer Dank für die Hilfestellungen und Diskussionen seitens der Bioinformatik. Durch seine nicht endende Geduld habe ich sehr viel über Kommandozeilenbefehle und über die Auswertung von *Crosslinks* gelernt.

Nicola und Franzi, euch möchte ich ganz besonders dafür danken, dass die Zusammenarbeit mit euch beiden bei der „Pflege und Instandsetzung“ unserer MassSpecs so unkompliziert abgelaufen ist. Dadurch habe ich viel von euch und vor allem auch mit euch gemeinsam über die Massenspektrometer und ihre „Wehwehchen“ gelernt.

Dr. Giambattista-Guaitoli möchte ich für die vielen Gespräche und Diskussionen bezüglich der *Size Exclusion Chromatography* danken.

Vielen lieben Dank auch an alle Kollegen des MPC in Tübingen, die mich während der Zeit als Doktorandin immer unterstützt haben. Danke für die tolle Arbeitsatmosphäre und eure unermüdliche Hilfsbereitschaft.

Zuletzt, aber für mich an wichtigster Stelle, möchte ich mich von ganzem Herzen bei meinen Freunden und meiner Familie bedanken. Danke dass ihr während der ganzen Zeit immer an meiner Seite gestanden habt, mir vor Präsentationen immer Händchen gehalten und in schwierigen Situationen immer ein offenes Ohr und tröstende Worte parat hattet. Meinem Freund Benjamin möchte ich dafür danken, dass er mich in schwierigen Momenten geduldig ertragen hat und mich jederzeit liebevoll unterstützt. Meinen Eltern möchte ich meinen besonderen Dank aussprechen. Durch ihre unermüdliche Unterstützung während meines Studiums und ihr stetiges Vertrauen in mich haben Sie mir diesen Weg ermöglicht.

Danke

## 7.5 Curriculum vitae

### Personal Information

Name	Wissinger, Yasmin
Date of birth	12.06.1988
Place of birth	Herrenberg
Nationality	German

### Education

Since October 2012	PhD student at Eberhard-Karls-University Tübingen, Germany Institute for ophthalmic research, Medical Proteome Center (Prof. Dr. rer. nat. Marius Ueffing)
October 2010 – September 2012	Studies in Analytical and Bioanalytical Chemistry at Hochschule Aalen, Germany Qualification: Master of Science Final grade: 1.3
October 2007 – July 2010	Studies in Applied Chemistry at Hochschule Reutlingen, Germany Qualification: Bachelor of Science Final grade: 1.6
September 1998 – July 2007	Carlo-Schmid-Gymnasium Tübingen, Germany Qualification: Abitur Final grade: 1.9
February 1996 – September 1998	Primary school Kusterdingen, Germany
September 1994 – February 1996	Primary school Pliezhausen, Germany

### Research Experience

Since October 2012	PhD thesis at the Institute for ophthalmic research at the Medical Proteome Center of Prof. Dr. rer. nat. Marius Ueffing at the Universitätsklinikum Tübingen, Germany Title: Stoichiometric and Structural Investigations of ciliopathy-related protein complex IFT-A
--------------------	---

March 2012 – August 2012	Master thesis at the Institute of ophthalmic research at the Medical Proteome Center of Prof. Dr. rer. nat. Marius Ueffing at the Universitätsklinikum Tübingen, Germany  Title: "Etablierung von HILIC-basierten Vortrennungsprotokollen für die phosphoproteomische Analytik"
February 2010 – August 2010	Bachelor thesis at the department of biochemistry at the Naturwissenschaftliches und Medizinisches Institut (NMI) Tübingen in Reutlingen, Germany  Title: "Epitope Protected Labelling: An approach to improve the sensitivity of antibody arrays"
August 2006	Internship at the Institute of Hirnforschung at the Universitätsklinikum Tübingen, Germany
November 2003	Internship at the department of Medical Genetics of Prof. Dr. med Olaf Rieß at the Universitätsklinikum Tübingen, Germany
<b>Time abroad</b>	
January 2014	Internship in the group of Karl Mechtler at the Institute of Molecular Pathology (IMP) in Vienna, Austria
September 2013	
<b>Publications</b>	
<p>Boldt, K., J. van Reeuwijk, Q. Lu, K. Koutroumpas, T. M. Nguyen, Y. Texier, S. E. van Beersum, N. Horn, J. R. Willer, D. A. Mans, G. Dougherty, I. J. Lamers, K. L. Coene, H. H. Arts, M. J. Betts, T. Beyer, E. Bolat, C. J. Gloeckner, K. Haidari, L. Hetterschijt, D. Iaconis, D. Jenkins, F. Klose, B. Knapp, B. Latour, S. J. Letteboer, C. L. Marcelis, D. Mitic, M. Morleo, M. M. Oud, M. Riemersma, S. Rix, P. A. Terhal, G. Toedt, T. J. van Dam, E. de Vrieze, <u>Y. Wissinger</u>, K. M. Wu, G. Apic, P. L. Beales, O. E. Blacque, T. J. Gibson, M. A. Huynen, N. Katsanis, H. Kremer, H. Omran, E. van Wijk, U. Wolfrum, F. Kepes, E. E. Davis, B. Franco, R. H. Giles, M. Ueffing, R. B. Russell, R. Roepman and U. K. R. D. Group (2016). "An organelle-specific protein landscape identifies novel diseases and molecular mechanisms." <u>Nat Commun</u> <b>7</b>: 11491</p>	

---

## Meetings and Conferences

June 2016	Thermo Fisher Scientific LC-MS Usermeeting, Dietzenbach, Germany
August 2015	APRS Symposium 2015, Klosterneuburg, Austria Poster presentation; Travel grant from Deutsche Gesellschaft für Proteomforschung (DGPF)
July 2015	GBM Annual Meeting 2015, Reutlingen, Germany Presentation
April 2015	Syscilia Annual Meeting, Naples, Italy
March 2015	Proteomic Forum 2015, Berlin, Germany Poster presentation; Travel grant from Deutsche Gesellschaft für Proteomforschung (DGPF)
October 2014	ERASysAPP Workshop, Larnaka, Cyprus
August 2014	Late Summer Practical Proteomics Seminar, Vienna, Austria
April 2014	Syscilia Annual Meeting, Heidelberg, Germany Presentation
September 2013	Late Summer Practical Proteomics Seminar, Vienna, Austria
April 2013	Syscilia Annual Meeting, Mallorca, Spain