

Structural Characterization of the Protein-Protein Interaction between the Tumor Suppressor p53 and S100 β

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften

Martin-Luther-Universität Halle-Wittenberg



Vorgelegt von An Jung Wei

Geboren am 01.11.1995 in Taoyuan, Taiwan

Gutachter:

1. Prof. Dr. Andrea Sinz
2. Prof. Dr. Stefan Hüttelmaier
3. PD Dr. Jürgen Schiller

Datum der öffentlichen Verteidigung: 13.02.2024

Table of Contents

Table of Contents.....	II
Abbreviations	VI
List of Figures.....	X
List of Tables.....	XII
List of Supplementary Figures	XIII
List of Supplementary Tables	XIV
Summary	1
1. Introduction.....	3
1.1. Intrinsically disordered proteins	3
1.2. The tumor suppressor p53 – The “guardian of the genome”	4
1.3. S100 proteins	8
1.4. Protein-protein interaction of p53 and S100 β	11
1.5. Mass Spectrometry.....	13
1.5.1. Ionization methods	14
1.5.1.1. Electrospray ionization.....	15
1.5.1.2. Nano-electrospray ionization	16
1.5.2. Mass analyzers.....	17
1.5.2.1. Orbitrap mass analyzer.....	17
1.5.2.2. Quadrupole mass analyzer	18
1.5.2.3. Quadrupole time-of-flight mass analyzer.....	19
1.5.3. Ion mobility spectrometry.....	20
1.5.4. Fragmentation	21
1.5.5. Mass spectrometers used for this work.....	22
1.5.5.1. Q-TOF mass spectrometer	22
1.5.5.2. Orbitrap mass spectrometer	23
1.5.5.3. Triple quadrupole mass spectrometer	24
1.5.5.4. Ion mobility Q-TOF mass spectrometer	25
1.5.6. Native mass spectrometry	27
1.5.7. Cross-linking mass spectrometry.....	28
1.5.7.1. The cross-linker disuccinimidyl dibutyric urea (DSBU).....	29
1.5.7.2. The cross-linker 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and sulfo-N-hydroxysulfosuccinimide (EDC/sulfo-NHS).....	30
1.5.7.3. The cross-linker sulfosuccinimidyl 4,4'-azipentanoate (sulfo-SDA)	31

1.5.8.	COMPetitive Pairing StatisticS.....	33
1.6.	Surface plasmon resonance	36
2.	Materials and Methods	38
2.1.	Materials.....	38
2.1.1.	Chemicals	38
2.1.2.	Plasmids	40
2.1.3.	Bacterial strains	40
2.1.4.	Media	40
2.1.5.	Protein.....	41
2.1.6.	Buffers and solutions	41
2.1.7.	Cross-linking reagents.....	42
2.1.8.	Enzymes	42
2.1.9.	Lab consumables	43
2.1.10.	Instruments	43
2.1.10.1.	Laboratory equipment.....	43
2.1.10.2.	Chromatography	44
2.1.10.3.	Mass spectrometers.....	45
2.1.11.	Software	45
2.2.	Methods.....	46
2.2.1.	Protein expression and purification.....	46
2.2.1.1.	Protein expression and purification of full-length p53	46
2.2.1.2.	Protein expression and purification of S100 β	48
2.2.2.	Buffer exchange.....	49
2.2.2.1.	Manual buffer exchange – Amicon filtration units	49
2.2.2.2.	Online buffer exchange using a self-packed column	49
2.2.3.	SDS-PAGE of proteins.....	49
2.2.4.	ESI mass spectrometry of intact cross-linked p53-S100 β complexes.....	50
2.2.5.	Cross-linking mass spectrometry.....	51
2.2.5.1.	Cross-linking reactions	52
2.2.6.	Enzymatic digestion	52
2.2.6.1.	In-gel proteolysis for protein identification.....	53
2.2.6.2.	In-solution proteolysis for cross-linked protein samples	53
2.2.7.	Mass spectrometry	54
2.2.7.1.	Electrospray ionization mass spectrometry	54
2.2.7.2.	Nano-HPLC-MS/MS-Orbitrap mass spectrometry	54

2.2.7.3.	Nano-HPLC-MS/MS-trapped ion mobility mass spectrometry.....	55
2.2.7.4.	Data analysis with MeroX.....	55
2.2.8.	Data analysis.....	56
2.2.8.1.	Data analysis with MeroX.....	56
2.2.9.	Surface plasmon resonance	58
2.2.9.1.	Buffer and sample preparation.....	58
2.2.9.2.	Protein immobilization	58
2.2.9.3.	Analyte injection and surface regeneration	58
3.	Results	59
3.1.	Expression and purification of p53 and S100 β	59
3.1.1.	Expression and purification of three p53 variants.....	59
3.1.2.	Expression and purification of ¹⁴ N and ¹⁵ N S100 β	60
3.2.	Native mass spectrometry of p53 and S100 β	61
3.2.1.	Tetrameric wild-type p53.....	61
3.2.2.	Dimeric L344A p53.....	63
3.2.3.	Monomeric L344P p53	64
3.2.4.	¹⁴ N S100 β	65
3.2.5.	¹⁵ N S100 β	66
3.3.	Covalent modification and mass spectrometry of p53-S100 β protein-protein interaction ..	67
3.3.1.	Tetrameric wild-type p53 and S100 β	67
3.3.2.	Dimeric L344A p53 and S100 β	68
3.3.3.	Monomeric L344P p53 and S100 β	70
3.4.	Cross-linking mass spectrometry of the interaction between p53-S100 β	72
3.4.1.	Cross-linking with DSBU	72
3.4.2.	Cross-linking with sulfo-SDA.....	74
3.4.3.	Cross-linking with EDC/sulfo-NHS.....	76
3.5.	COMPetitive PAirng StatisticS (COMPASS)	80
3.5.1.	Cross-linking of ¹⁴ N and ¹⁵ N S100 β	80
3.5.2.	Enrichment of S100 β dimer with SEC	80
3.5.3.	Data analyses with MeroX and COMPASS.....	82
3.6.	Surface Plasmon Resonance	86
4.	Discussion	90
4.1.	Expression and purification of full-length p53	90
4.2.	Expression and purification of S100 β	91
4.3.	Protein-protein interaction of p53 and S100 β	91

4.3.1.	Covalently modified ESI-MS of p53-S100 β protein complexes	92
4.3.2.	Cross-linking mass spectrometry	93
4.3.3.	Surface plasmon resonance	96
4.4.	COMPetitive PAiring Statistics	96
5.	Conclusions and Outlook	97
	References	i
	Appendix	xi
	Acknowledgement	xxv
	Publikationen und Tagungsbeiträge	xxvii
	Lebenslauf	xxix
	Selbständigkeitserklärung	xxx

Abbreviations

2xYT	Double yeast tryptone medium
AA	Ammonium acetate
ABC	Ammonium bicarbonate
APS	Ammonium persulfate
ACN	Acetonitrile
ATR	Attenuated Total Region
AUC	Analytical ultracentrifugation
BSA	Bovine serum albumin
CBP	CREB-binding protein
CCS	Collisional cross section
CID	Collision-induced dissociation
CMD	Carboxymethyl dextran
COMPASS	COMPetitive PAiring StatisticS
CREB	cAMP response element-binding protein
CRM	Charged residue model
C-trap	Curved linear trap
DBD	DNA-binding domain
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSBU	Disuccinimidyl dibutyric urea
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
ECD	Electron capture dissociation
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EGTA	Ethylene glycol-bis(aminoethyl ether)-N,N,N',N'-tetraacetic acid
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FA	Formic acid
FDR	False discovery rate
HCD	Higher-energy collision-induced dissociation
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethane-1-sulfonic acid
HLT	Histidine-lipoyl-tobacco etch virus
HPLC	High-performance liquid chromatography
IAA	Iodoacetamide
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IMAC	Immobilized metal ion affinity chromatography
IMS	Ion mobility spectrometry
ITC	Isothermal titration calorimetry
K_D	Dissociation constant
kDa	Kilodalton
LB	Lysogeny broth
LC	Liquid chromatography
LCMS	Liquid chromatography mass spectrometry
LIT	Linear ion trap
LTQ	Linear triple quadrupole
m/z	Mass-to-charge ratio
MDM2	Murine double minute 2
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry

MW	Molecular weight
NCE	Normalized collision energy
NHS	<i>N</i> -Hydroxysuccinimide
NMR	Nuclear magnetic resonance
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDB	Protein data bank
PRR	Proline-rich region
PTM	Post-translational modification
Q	Quadrupole
Q-TOF	Quadrupole time-of-flight
Reg.	Regulatory domain
RISE	Reporter ion scan event
SDA	Succinimidyl 4,4'-azipentanoate
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
SPR	Surface plasmon resonance
TAD	Transactivation domain
TCEP	Tris(2-carboxylethyl)phosphine
TD	Tetramerization domain
TEMED	<i>N,N,N',N'</i> -Tetramethylethane-1,2-diamine
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
Th	Thomson
TIC	Total ion current
TIMS	Trapped ion mobility spectrometry
TOF	time-of-flight

TQD	Triple quadrupole
Tris-HCl	Tris(hydroxymethyl)aminoethane hydrochloride
UV	Ultraviolet
UVPD	Ultraviolet photon dissociation
v/v	Volume per volume
YT	Yeast tryptone
XL	Cross-linking
XL-MS	Cross-linking mass spectrometry
z	Charge

Proteinogenic amino acids (three-letter code and one-letter code)

Alanine	Ala	A	Arginine	Arg	R
Asparagine	Asn	N	Aspartic acid	Asp	D
Cysteine	Cys	C	Glutamine	Gln	Q
Glutamic acid	Glu	E	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

List of Figures

Figure 1. Domain structure and amino acid sequence of p53.	6
Figure 2. Tetramerization domain (TD) of p53.	7
Figure 3. Protein binding partners of the p53's C-terminal regulatory domain (Reg).	8
Figure 4. Structure of S100 Protein.	9
Figure 5. Conformational change of homodimeric S100 β upon binding of calcium ions.	9
Figure 6. S100 protein involvement in tumorigenic processes.	10
Figure 7. Interaction of tumor suppressor protein p53 with S100 protein family.	11
Figure 8. Interaction between S100 β and p53 peptide.	12
Figure 9. Proposed p53-S100 β binding mechanism based on peptide studies.	13
Figure 10. Criteria for selecting ionization method.	15
Figure 11. The ESI process in positive ionization mode.	16
Figure 12. Schematic representation of an orbitrap mass analyzer and a C-trap in a commercial instrument (Thermo Fisher Scientific).	18
Figure 13. Schematic representation of a linear quadrupole.	19
Figure 14. Schematic representation of a quadrupole time-of-flight mass spectrometer.	19
Figure 15. (a) Schematic of the ion-mobility spectrometry (IMS) and (b) Calculation of ion-mobility with the Mason-Schamp equation ^[126]	20
Figure 16. (a) Nomenclature of MS/MS peptide fragments as proposed by Roepstorff, Biemann, Fohlman, and Johnson ^[144] . (b) Proposed structures of peptide fragment ions.	22
Figure 17. Schematic layout of the High-Mass Q-TOF II mass spectrometer (Micromass/MS Vision).	23
Figure 18. Schematic layout of the Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific).	24
Figure 19. Schematic representation of a triple quadrupole mass spectrometer (TQD).	25
Figure 20. Schematic layout of the timsTOF Pro mass spectrometer (Bruker Daltonik).	26
Figure 21. Illustration of a Parallel Accumulation Serial Fragmentation (PASEF) method in timsTOF Pro mass spectrometer.	27
Figure 22. Cross-linking chemistry of DSBU MS-cleavability by collisional activation.	29
Figure 23. Cross-linking chemistry of EDC/sulfo-NHS cross-linkers.	30
Figure 24. Cross-linking chemistry of diazirine cross-linkers.	32
Figure 25. Dissociation of a sulfo-SDA cross-linked product via the diazo pathway.	33
Figure 26. Schematic representation of intra/inter-protein cross-links.	34
Figure 27. Schematic illustration of quantitation with the COMPASS approach.	35
Figure 28. Crystal Structure (PDB: 3CZT) and amino acid sequence of S100 β	36
Figure 29. Illustration of SPR experimental setup.	37

Figure 30. Protocol for ESI-MS of intact cross-linked protein complexes.....	51
Figure 31. Cross-linking Mass Spectrometry experimental workflow.	51
Figure 32. Protocol of SEC enrichment and cross-link quantitation with COMPASS.....	56
Figure 33. Summary of the expression and purification of three variants of p53.....	60
Figure 34. Summary of the expression and purification of the ¹⁴ N and the isotope-labelled ¹⁵ N S100β.	61
Figure 35. Native mass spectrum of the wild-type tetrameric p53.....	62
Figure 36. Native mass spectrum of the dimeric variant L344A p53.....	63
Figure 37. Native mass spectrum of the monomeric variant L344P p53.....	64
Figure 38. Native mass spectrum of ¹⁴ N S100β.....	65
Figure 39. Native mass spectrum of ¹⁵ N S100β.....	66
Figure 40. ESI mass spectra of intact, DSBU cross-linked S100β and wild-type tetrameric p53 (a) with the addition of calcium (b) without calcium.....	67
Figure 41. CID-MS/MS of the complex of wild-type tetrameric p53 and S100β.....	68
Figure 42. ESI mass spectra of DSBU intact, cross-linked S100β and the dimeric L344A p53 variant (a) with the addition of calcium (b) without calcium.....	69
Figure 43. CID-MS/MS of the complex of the dimeric L344A p53 variant and S100β.....	70
Figure 44. ESI mass spectra of DSBU intact, cross-linked S100β and the monomeric L344P p53 variant (a) with the addition of calcium (b) without calcium.....	71
Figure 45. CID-MS/MS of the complex of the monomeric L344P p53 variant and S100β.....	71
Figure 46. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100β and wild-type p53 with DSBU.....	72
Figure 47. Fragment ion spectrum of the cross-linked product of the L344P monomeric variant with S100β ⁰ {GSHMSELEK ⁹ – ³⁷⁴ KGQSTR ³⁸⁰	74
Figure 48. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100β and wild-type p53 with sulfo-SDA.....	75
Figure 49. Fragment ion spectrum of the sulfo-SDA cross-linked product of the wild-type tetrameric p53 with S100β ³⁴⁰ EMFRELNEALEIK ³⁵² – ⁴³ ELSHFLEEIK ⁵²	76
Figure 50. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100β and wild-type p53 with EDC/sulfo-NHS.....	77
Figure 51. Fragment ion spectrum of the EDC/sulfo-NHS cross-linked product of the monomeric L344P p53 with S100β ⁵⁸ DPGPDEAPR ⁶⁶ – ⁰ {GSHMSELEK ⁹	79
Figure 52. Cross-linking mapping of wild-type p53 and commercial S100β without the artificial residues.....	80
Figure 53. SDS-PAGE (15%) of the cross-links of ¹⁴ N and ¹⁵ N S100β with DSBU and sulfo-SDA.....	80
Figure 54. Enrichment with size exclusion chromatography of cross-linked S100β dimer with DSBU.....	81
Figure 55. ESI-mass spectra of the dimeric and monomeric fractions from SEC enrichment of the DSBU cross-linked samples.....	82
Figure 56. Enrichment with SEC of cross-linked S100β dimer with sulfo-SDA. SDS-PAGE (15%) gel run was performed to check the dimeric and monomeric fractions.....	82
Figure 57. Fragment ion mass spectrum of the cross-linked product between ¹⁴ N and ¹⁵ N S100β ⁹ KAMVALIDVFHQYSGR ²⁴ – ⁵¹ ikeqevvdvme ⁶²	83
Figure 58. COMPASS cross-link analysis of the K9 to K59 cross-link.....	84
Figure 59. Summary of identified inter-protein crosslinks of the calcium-bound state of ¹⁴ N and ¹⁵ N S100β with DSBU and sulfo-SDA mapped with xiNET.....	85
Figure 60. Ligand preparation – pH scouting with 10 mM sodium acetate.....	87

Figure 61. Ligand immobilization of S100 β	87
Figure 62. Analyte injection of p53.	89
Figure 63. Schematic representation of the stoichiometry of p53-S100 β complexes upon investigation with covalently modified ESI-MS.	93
Figure 64. Summary of XL-MS of S100 β and different p53 variants. (a) Tetrameric wild-type (b) dimeric L344A variant (c) monomeric L344P variant with red line showing cross-links with DSBU, cyan line showing cross-links with EDC/sulfo-NHS, and the blue line showing cross-links with sulfo-SDA...95	

List of Tables

Table 1. Reported dissociation constants between p53 peptides and S100 β	13
Table 2. Composition of stacking and resolving gel for SDS-PAGE.	50
Table 3. Summary of the identified p53-S100 β cross-links with DSBU.....	73
Table 4. Summary of the identified p53-S100 β cross-links with sulfo-SDA.	75
Table 5. Summary of all identified cross-links with EDC/sulfo-NHS.	78
Table 6. Summary of DSBU cross-links with wild-type p53 and commercial S100 β	79
Table 7. Summary of inter-protein cross-links percentage of ¹⁴ N and ¹⁵ N S100 β	86

List of Supplementary Figures

Supplementary Figure 1. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100 β with DSBU (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.	xiii
Supplementary Figure 2. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100 β with sulfo-SDA (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.	xiv
Supplementary Figure 3. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100 β with EDC/sulfo-NHS (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.	xv
Supplementary Figure 4. SDS-PAGE (4-20% gradient) for cross-linking wild-type p53 and the commercial S100 β without the artificial residues (Gly-His-Ser).....	xvi
Supplementary Figure 5. Native mass spectrum of wild-type p53 and S100 β with no covalent fixation (no cross-linker added).	xvi
Supplementary Figure 6. ESI mass spectra showing the complex formation of S100 β and wild-type tetrameric p53. Peak assignments of calcium-added and -depleted samples are presented.	xvii
Supplementary Figure 7. ESI mass spectra showing the complex formation of S100 β and L344A dimeric p53. Peak assignments of calcium-added and -depleted samples are presented.....	xviii
Supplementary Figure 8. ESI mass spectra showing the complex formation of S100 β and L344P monomeric p53. Peak assignments of calcium-added and -depleted samples are presented.....	xix
Supplementary Figure 9. Instrumental setup of the online buffer exchange (OBE) at load (top) and inject (bottom) position.	xxi
Supplementary Figure 10. Full-scan mass spectrum showing signals of the SDA cross-linked peptide between Nterm and E49 of ^{14}N and ^{15}N S100 β with a retention time 44.0-45.8 min, scan number 14446-15046.	xxii
Supplementary Figure 11. Full-scan mass spectrum showing signals of the DSBU cross-linked peptide between K9 and K59 of ^{14}N and ^{15}N S100 β with a retention time 49.6-51.2 min, scan number 16975-17544.	xxiii
Supplementary Figure 12. Full-scan mass spectrum showing signals of the DSBU cross-linked peptide between N-terminus and K52 of ^{14}N and ^{15}N S100 β with a retention time 46.1-47.6 min, scan number 15759-16299.	xxiv

List of Supplementary Tables

Supplementary Table 1. Theoretical and experimental masses of p53 and complexes..... xx

Summary

The tumor suppressor protein p53, which is also referred to as the “guardian of the genome”, plays a crucial role in gene regulation and cancer development. Being a well-known intrinsically disordered protein (IDP), roughly 40% of p53’s sequence lacks a fixed three-dimensional structure and is natively unfolded. Despite being studied extensively for the past decades, the structural characterization of p53 remains controversial due to different proposed models. There are numerous protein binding partners of p53 with many of them binding to the regulatory domain located at the C-terminus of p53, one example being S100 β . The S100 protein family is a biomarker for malignant melanoma and plays an important role in tumorigenesis. S100 β , in particular, is reported to possess both stabilizing and inhibiting roles when bound to p53. So far, only peptide studies of p53 to S100 β have been reported. In order to shed light on the molecular details of the p53-S100 β interaction, a detailed structural elucidation of this interaction with full-length p53 is highly important.

Point mutations in the tetramerization region (TD) of p53 are known to alter the oligomeric state of p53. To tackle the ambiguity and to understand the controversial binding mechanism of p53 to S100 β , three variants of the full-length p53 were expressed in *E. coli* and purified. This allowed a structural characterization of S100 β with three different oligomeric states of p53: Tetrameric (wild-type), dimeric (L344A) and monomeric (L344P) species.

Integrating native mass spectrometry and cross-linking mass spectrometry with surface plasmon resonance provided concrete insights into the p53-S100 β protein-protein interactions. The stoichiometry of the p53-S100 β complexes as well as protein subunits composition were characterized by native mass spectrometry. Cross-linking mass spectrometry using different cross-linkers allowed targeting different functional groups in the side chains of amino acids in the p53-S100 β complexes and yielded site-specific information on the protein interfaces. Finally, thermodynamic parameters, such as the binding affinities between p53 and S100 β were characterized with surface plasmon resonance

Characterizing different types of protein cross-links are challenging. For a homodimeric protein like S100 β , differentiating the inter-protein cross-links (connecting two different monomers) from intra-protein cross-links (within one monomer) is crucial to obtain meaningful structural information. An integrative approach that has been recently developed in our group, termed COMPetitive PAiring StatisticS (COMPASS), allow distinguishing the two different types of protein cross-links and is able to give quantitative information on the cross-links observed. Due to the conformational changes induced

in S100 β by the addition of calcium, cross-link identification and quantitation of the calcium-loaded and -depleted states were compared for the p53-S100 β complexes.

In this study, all three p53 oligomeric states (tetrameric, dimeric, and monomeric p53) showed an identical binding behavior towards the S100 β homodimer. Also, stoichiometries of the p53-S100 β complexes, site-specific interactions, and binding affinities were found to be independent of p53's oligomeric state. This contradicts the current knowledge from p53 peptide studies. Speculation regarding the molecular mechanism of p53-S100 β interactions and the regulatory role of S100 β exerted on the p53 oligomeric state requires further *in vivo* studies.

For the cross-links quantitation of the S100 β homodimer, the differentiation of two different types of cross-links was achieved. The integrative COMPASS approach, previously applied only on IDPs, was applied also to the globular protein. With this approach, the calcium-bound and -depleted states of the S100 β homodimer were compared and the quantitation of the different cross-links confirmed the calcium-dependent conformational change. This opens a new opportunity to study the structural biology of globular proteins.

1. Introduction

1.1. Intrinsically disordered proteins

Up until the past few decades, it was believed that functional proteins possess a defined three-dimensional structure^[1]. The significance of a protein's tertiary structure for its functions was evident from multiple Nobel Prize awards: In 1958, Frederick Sanger was awarded with the Nobel Prize in Chemistry for his study of the protein structure of insulin^[2]. Later in 1962, John Cowdery Kendrew and Max Ferdinand Perutz were awarded with the same prize for their first reports of three-dimensional protein structures; the first tertiary structure of myoglobin at a resolution of 6 Å and the oxy-hemoglobin structure at 5.5 Å using X-ray diffraction^[3]. It was also believed that a protein's structure is determined by its amino acid composition. Christian B. Anfinsen who was awarded the Nobel Prize in Chemistry in 1972 proposed the "Anfinsen's dogma," also known as the thermodynamic hypothesis. This dogma states that the native structure of a protein is determined strictly by its amino acid sequence^[4]. The paradigm stating that the amino acid sequence leads to structure and creates function has been generally accepted^[5]. Proteins are polypeptide structures that consist of a combination of amino acid residues. Primary structure details the composition of amino acid residues, while secondary structure describes the chains that are stabilized by hydrogen bonds that can, for example, create an alpha-helical or a beta-sheet conformation. Tertiary and quaternary structures illustrate the three-dimensional shapes and the more complex interactions of proteins^[6].

The Anfinsen dogma came into questions as the concept of protein misfolding was brought to attention^[7]. The discovery of multiple native structures for a protein has also refuted the claim that a protein's three-dimensional structure is strictly determined by its amino acid sequence^[8]. In 1969, Cyrus Levinthal argued that due to a large number of the degrees of freedom in an unfolded polypeptide chain^[9], there could be many possible folded states yielding many possible conformations. This is termed as the "Levinthal's paradox"^[10]. Researchers have also come up with doubts of the idea that all functional proteins must possess a folded three-dimensional structure. In 1978, functional disorder of the photosynthetic protein subunits of *Rhodospseudomonas viridis* was observed with X-ray crystallography^[11] where missing electron density was attributed to the variation in protein conformation. In the same year, nuclear magnetic resonance (NMR) experiments also indicated a highly charged, functional, but disordered histone H5 tail^[12]. As more examples pointed out the evidence that protein folding does not directly influence protein function, more experiments were performed, and more theories were proposed to explain the newly observed phenomenon.

Since the year of 2000, the traditional "structure-function paradigm" is no longer the golden standard^{[5][13]}, as an increased amount of experimental evidence has pointed out that a rather large

amount of protein domains, and even the entire proteins, lack an ordered structure even under physiological conditions^[14]. It can certainly not be denied that there is a correlation between the amino acid sequence of protein and its folded structure, however, it has been more widely accepted recently that unfolded or partly folded proteins can also possess key biological functions^[1]. According to the current knowledge, between 3-17% of eukaryotic proteins are fully disordered and 30-50% of eukaryotic proteins contain at least one intrinsically disordered region (IDR)^[15]. Lacking a stable and well-defined three-dimensional structure is beneficial for exerting the different functions of a protein^[16] as they can exhibit different binding conformations for various interactions partners. IDRs are also important loci for alternative splicing^[17] and are hotspots for post-translational modifications (PTMs)^[18].

Understanding the molecular basis of intrinsically disordered proteins (IDPs) is therefore essential in order to fully understand the correlation between disorder and its biological implications. Structural investigations of IDPs have enabled a deeper understanding of biological functions and allowed the development of more complex bioinformatics algorithms for predicting protein disorder and its relations to protein functions^[19]. One of the pioneer algorithms to predict protein disorder, Predictor Of Natural Disordered Regions VLXT (PONDR VLXT) allowed prediction of protein termini. More sophisticated methods have been developed and more advanced machine learning techniques allow disorder prediction, such as DisEMBL^[20], AlphaFold^[21], and Disopred2^[22]. The most important techniques to structurally characterize IDPs include NMR, fluorescence spectroscopy, and increasingly also mass spectrometry (MS). The challenges for characterizing IDPs are that they are usually not amenable to X-ray crystallography as one of the classical high-resolution techniques for protein structure analysis. For X-ray crystallography, the difficulty of the formation of the crystal lattices was observed due to the less optimal solubility and the weaker stability of IDPs^{[23][24]}. NMR is currently the main technique to characterize IDPs and IDRs, but experiments might be challenging due to the lack of dispersion of protein signals and a substantial signal-overlap. In terms of protein amount, NMR requires milligrams of proteins, which is not always achievable with IDPs^[25]. Structural characterization with MS has in recent years become a popular technique with IDPs^[26]. MS measures the mass-to-charge ratio (m/z) of a protein and with the requirement of low sample amounts and the ability to combine a broad range of applications. MS can play an important role for investigating IDPs.

1.2. The tumor suppressor p53 – The “guardian of the genome”

The discovery of p53 took place in 1979 as a cellular protein bound to the simian virus (SV40) large T antigen was observed. Despite having a molecular mass of 43.7 kDa, the origin of the name p53 came from the observed molecular weight of 53 kDa in the SDS-PAGE analysis^[27]. At first, p53 was considered an oncogene due to its low concentration in healthy cells and its overexpression in cancer cells. It was

only until the late 1980s, when Vogelstein discovered that deletions, insertions, and point mutations in the TP53 gene were signatures of colorectal carcinoma^[28]. Through multiple studies with human cancer genetics, mouse models, and cell biology, it was realized that p53 is in fact a tumor suppressor protein instead of an oncogene^[29].

Given the name the “guardian of the genome”, p53 plays a crucial role in maintaining genomic integrity^[30]. Loss of p53’s activity due to inactivation or mutation is the cause of more than 50% of human cancers^[31]. Murine Double Minute 2 (MDM2) is an E3 ubiquitin ligase and plays an important role as a negative regulator of p53. It is actively involved in p53 signaling pathways and is closely associated with tumor progression. DNA damage signaling and oncogenic signaling are the two pathways that activate p53 during tumor progression. Upon DNA damage, a signaling pathway a cascade of kinases is started comprising ATM, ATR, Chk1, and Chk2 that phosphorylates the serine and threonine residues (S20, S33, S46, T81) in p53 and blocks the binding of MDM2^{[32][33]}. Within the oncogenic signaling route, p53 activity is stabilized by the induced expression of p14/16^{ARF}. Binding of p14/16^{ARF} to MDM2 inhibits the ligase activity and promotes the degradation of MDM2^[34]. Many PTMs are observed in p53. Aside from phosphorylation and ubiquitination, other PTMs like acetylation, methylation, and SUMOylation are important in p53 regulation^[35].

There are five domains that make up the tumor suppressor protein p53 (Figure 1), the *N*-terminal transactivation domain (TAD), the proline-rich region (PRR), the DNA-binding domain (DBD), the tetramerization domain (TD), and the *C*-terminal regulatory domain (Reg). Each domain has a different role, and every domain is crucial for the functionality of p53 protein as a tumor suppressor. The *N*-terminal TAD, PRR, and Reg, accounting for 40% of the whole protein, are intrinsically disordered. Phosphorylation and ubiquitination in p53’s TAD and acetylation in Reg are important PTMs that regulate p53’s activity. An interesting observation points out the higher frequency of PTMs located in the IDRs of p53, which underlines the need for a more detailed structural to clarify the relevance of disorder for the various biological functions of p53^[36].

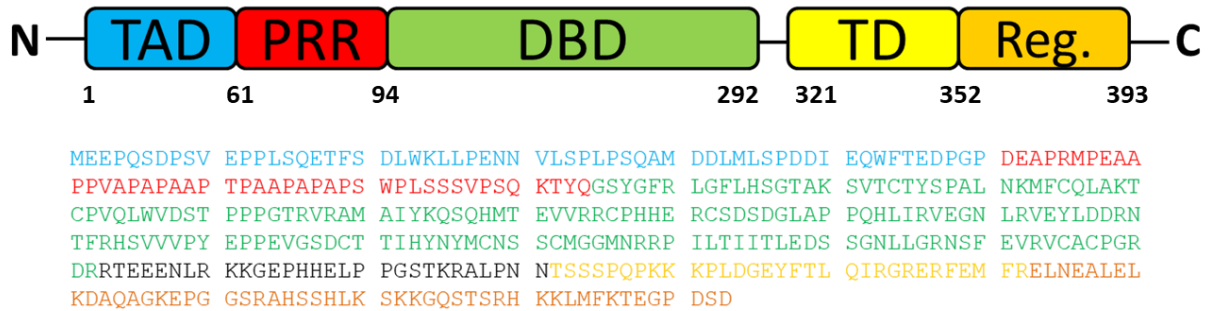


Figure 1. Domain structure and amino acid sequence of p53. The tumor suppressor p53 is a multi-domain protein (upper panel). The amino acid residues are color-coded according to the respective domains (lower panel). The five domains of p53 include the *N*-terminal transactivation domain (TAD, residues 1-61), the proline-rich region (PRR, residues 61-94), the DNA-binding domain (DBD, residues 94-292), the tetramerization domain (TD, residues 321-352), and the *C*-terminal regulatory domain (Reg, residues 352-393).

The highly acidic *N*-terminal TAD is responsible for proteasomal degradation and regulation. It can be further divided into two subdomains (TAD1 and TAD2). As briefly mentioned in the earlier paragraph, MDM2 is a negative regulator of p53. In the absence of cellular stress, most serine and threonine residues of the TAD remain unphosphorylated. Tight binding of MDM2 to Thr-18 in particular suppresses p53's activity and enhances nuclear export and proteasomal degradation^[37]. However, in response to cellular stress, multiple serine and threonine residues in the TAD are phosphorylated by protein kinases^[38]. Phosphorylation significantly decreases the binding affinity of MDM2 and increases the binding affinities with positive cofactors, such as CBP/p300, to stimulate cell cycle arrest and apoptosis^{[39][40][41]}.

The PRR contains PXXP repeats (X=any residue)^[42]. The region is not required for transactivation but is crucial for inducing apoptosis. As experimental results with a p53 mutant lacking the PRR have shown, p53-induced apoptosis could not be achieved without the PRR^[43].

DBD is a zinc-binding domain that spans from residue 94 to 292. The DBD binds to specific DNA sequence and it was discovered that two copies of 5'-PuPuPuC(A/T)(T/A)GPyPyPy-3' (Pu, purines and Py, pyrimidines) separated by 0 to 13 base pairs is one of the consensus binding sites^[44]. Due to a higher rate of mutation in this region, the DBD is known as mutation hotspot^[45]. Mutations classified as contact (R248 and R273) and structural (R175, G245, R249, and R282) mutants can disrupt DNA binding or alter the structure of the DNA-binding surface^[46]. Zinc-binding can be disrupted by mutations such as C176F, H179R, and C242F^[47]. The DBD was found to possess a compact structural domain unlike the *N*- and the *C*-termini. Digestion of p53 with subtilisin resulted in peptide fragments from only the *N*- and the *C*-terminus but showed the protease-resistance nature of the DBD^[48].

The TD is connected to the DBD via a disordered linked and plays a important role in p53 oligomerization. It is responsible for maintaining p53 as an active and functional homotetramer. With analytical ultracentrifugation (AUC) it was determined that the dissociation constant for dimerization of p53 is 1 nM and the dissociation constant for tetramer formation is 100-1000 nM^[49]. Protein-protein interactions, site-specific DNA binding, and PTMs are important factors that can either stabilize the tetrameric p53 or inhibit the oligomerization of p53, thereby hindering its transcriptional activity. Although most p53 mutations occur in the DBD, mutations in the TD are also common. Mutations in the TD account for more than 20% of germline mutations in Li-Fraumeni syndrome (LFS), a disorder that predisposes carriers to cancer development^[50]. The inhibition of p53 tetramer formation is a likely outcome of TD mutations, resulting in dimeric or monomeric forms of p53. The binding mechanisms can be directly affected with a single-point mutations and missense mutants can lead to different oligomeric states of p53 and effectively decrease the transcriptional activity of p53 (Figure 2). For studying the effect of different p53 oligomeric states on protein-protein interactions, the monomeric L344P variant and the dimeric L344A variant were used in this study to compare the different binding mechanisms with the tetrameric wild-type p53.

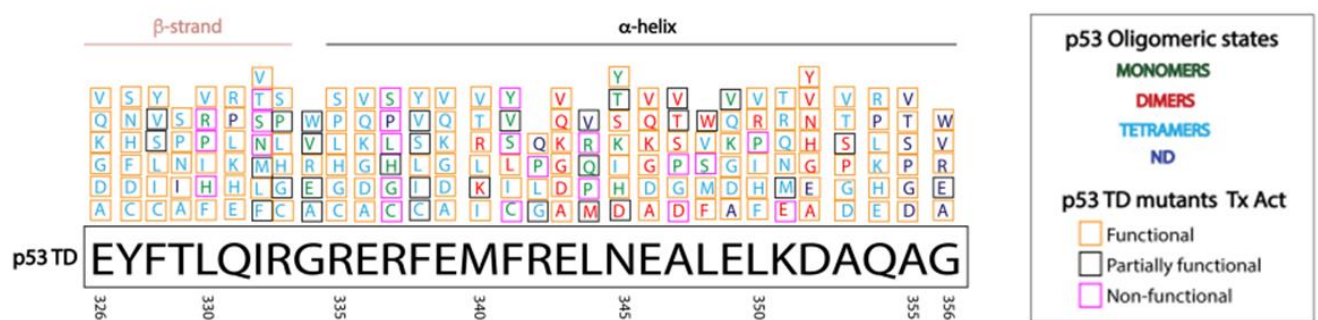


Figure 2. Tetramerization domain (TD) of p53. The oligomeric states and the functionality of p53 upon introduction of single-point mutations are indicated by the color of the amino acid (green: monomer, red: dimer, light blue: tetramer, dark blue: no data reported) and the color of the square (yellow: fully functional, black: partially functional, pink: non-functional). Figure adapted from Gencel-Augusto and Lozano. Copyright to Cold Spring Harbor Laboratory Press^[50].

The Reg at the C-terminus of p53, as suggested by its name, is critical for the regulation of p53 function^[51]. Deletion and introduction of point mutations in this region abolish p53-mediated transcription in the context of DNA and chromatin binding^[52]. In addition to regulating the DNA binding by the DBD, some nonspecific DNA binding also takes place in the Reg of p53. In fact, binding of Reg to nonspecific DNA sequences mediates the binding of specific DNA sequences to the DBD. It also plays a critical role for the association of p53 with cofactors and protein binding partners^[53]. Various protein binding partners interact with p53 specifically in the Reg (Figure 3). S100 β , sirtuin, CREB-binding

protein (CBP), and cyclin A2 are notable examples of proteins that bind specifically to this region from residue 365 to 392. In its native unbound state, Reg remains intrinsically disordered. The calculation of the solvent accessible surface area (SASA) indicates that upon binding to S100 β and sirtuin, a change in the structural conformation was observed. A disorder-to-order transition was shown upon binding of p53's Reg to S100 β and sirtuin. The Reg adopts an alpha-helical structure when bound to S100 β and a beta-sheet conformation when bound to sirtuin. When bound to CBP and cyclin A2, the p53 Reg. remains disordered^[54].

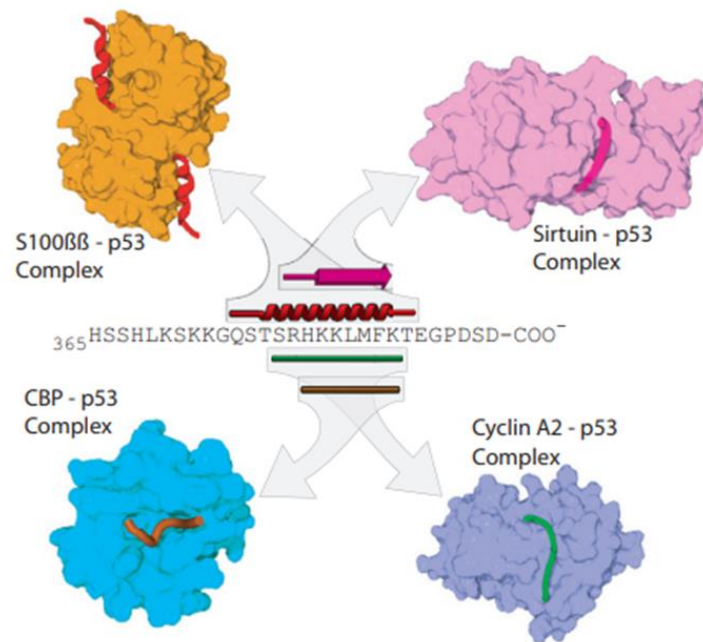


Figure 3. Protein binding partners of the p53's C-terminal regulatory domain (Reg). The amino acid sequence of the binding region and the secondary structures formed upon interaction with different protein binding partners are indicated accordingly. P53 adapts an alpha-helical structure (red helix) upon binding to S100 β and a beta-sheet structure (pink) upon binding to sirtuin. Upon interacting with CBP and cyclin A2, no disorder-to-order conformational change was observed (green lines indicated the binding region remains disordered). Figure adapted from Oldfield et al^[55].

1.3. S100 proteins

The discovery of the S100 protein family took place in 1965 by Moore^[56]. An S100 protein was discovered in a subcellular fraction that was thought to contain nervous system-specific proteins from bovine brain. The name "S100" originates from the observation that the isolated subcellular fraction containing the S100 protein was soluble in 100% saturated ammonium sulfate at neutral pH^[57]. Representing the largest subgroup within the EF-hand superfamily, the S100 protein family comprises more than 20 members with different degrees of homology^[58]. The protein family shares a similar MW (9-13 kDa) and structural conformation, as all of them contain symmetrical dimers with each S100 subunit containing four alpha-helices (Figure 4). Each of the monomer contains two calcium-binding domains, including a C-terminal canonical EF-hand motif (H_{III}, H_{IV}) composed of 12 amino acid residues

and an *N*-terminal “pseudo” canonical EF-hand motif (H_I, H_{II})^[59]. From the observed crystal structures it becomes apparent that the *N*-terminal EF-hand is in charge of dimer formation^{[60][61][62]}. Upon calcium binding, the C-terminal EF-hand of S100 undergoes a conformational rearrangement, where H_{III} (Helix III) is reoriented and the hydrophobic pocket is exposed (Figure 5)^[63].

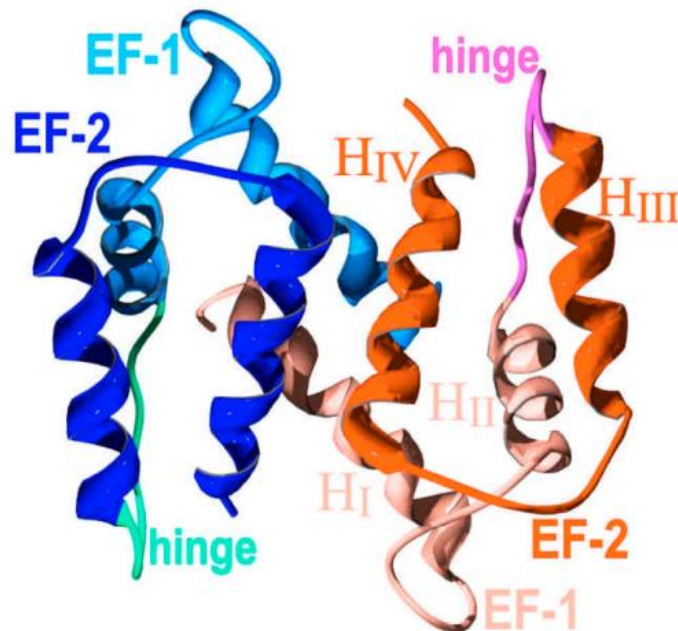


Figure 4. Structure of S100 Protein. Every member of the S100 protein family is comprised of symmetrical dimers. Each of the S100 monomer contains two calcium-binding domains, including a C-terminal canonical EF-hand motif (H_{III} and H_{IV}) composed of 12 amino acid residues and an *N*-terminal “pseudo” canonical EF-hand motif (H_I and H_{II}). Figure adapted from Heizmann et al^[64].

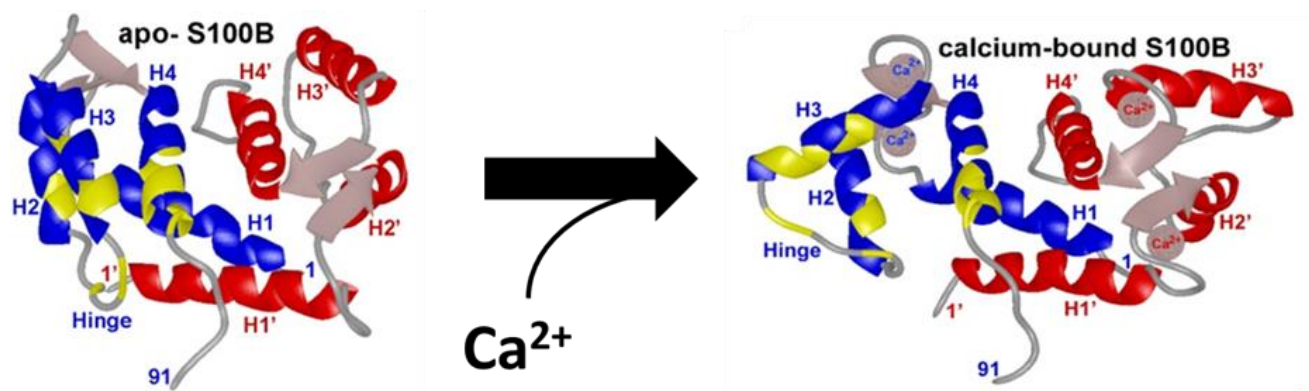


Figure 5. Conformational change of homodimeric S100β upon binding of calcium ions. Calcium binding of S100β re-orient helix 3 and exposes a hydrophobic pocket. Figure adapted from Hartman et al^[65]. © 2013 Kira G. Hartman, Paul T. Wilder, Kristen Varney, Alexander D. Jr. MacKerell, Andrew Coop, Danna Zimmer, Rena Lapidus and David J. Weber. Adapted from Hartman et al; originally published under CC BY 3.0 license. Available from: 10.5772/55176.

Despite having similar structural conformations, members of the S100 protein family show different binding behaviors towards the same protein^[66]. S100A1, S100A2, S100A4, and S100 β all bind to p53 and affect its biological functions but with different binding affinities^[67]. Binding to phosphoglucomutase has been observed for both S100A1 and S100 β and interestingly, the former stimulates phosphoglucomutase activity while the latter inhibits its activity^[68]. S100 proteins regulate both intra- and extracellular activities^[59]. Within cells, S100 proteins regulate calcium homeostasis, energy metabolism, inflammation, and proliferation. On the extracellular level, S100 proteins interact with cell surface receptors, including G-protein coupled receptors (GPCRs), receptor For advanced glycation end products (RAGE), and N-glycans^[69]. S100 proteins are found in serum and other biological fluids and an elevated S100 protein concentration serves as biomarker for various pathological conditions^[70]. The association between S100 proteins and cancer has seen a growing interest due to several observations indicating an involvement of S100 proteins in tumorigenesis. Members of the S100 protein family interact with target proteins and play an active role during different steps of cancer progression such as cell cycle progression, cell proliferation, cell apoptosis, and cell differentiation (Figure 6)^[63]. Elucidation of protein interactions involving S100 proteins provides further structural information and biological insights for potential therapeutic targets.

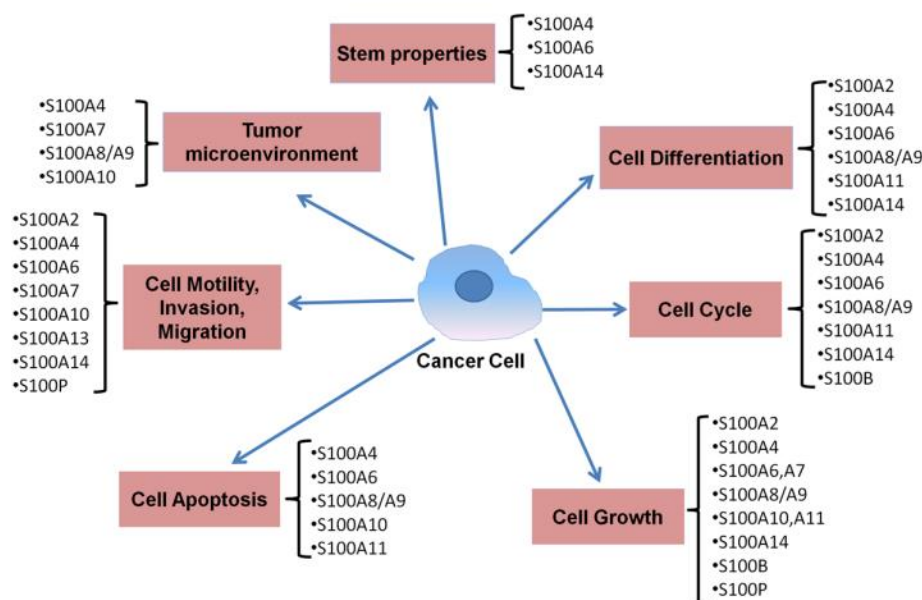


Figure 6. S100 protein involvement in tumorigenic processes. Members of the S100 protein family have effects on tumorigenic processes by affecting different stages of cell cycle regulation. Figure adapted from Chen et al^[63].

1.4. Protein-protein interaction of p53 and S100 β

Several members of the S100 protein family bind to p53 and thereby exert several regulatory functions (Figure 7). One member of the protein family, S100 β is of particular interest. Being one of the most extensively studied proteins in the family, S100 β is a well-known biomarker for the prognosis of malignant melanoma^[71]. Regarding its biological functions, S100 β plays a dual regulatory role on p53. In normal cells, S100 β promotes p53-mediated cell growth and apoptosis via a calcium-dependent, cell-contact signaling pathway^[72]. In cancer cells, S100 β inhibits p53 phosphorylation, disrupts the formation of the p53 tetramer, thus deactivating the role of p53 as a functional tumor suppressor^[73]. The dual regulatory role is reflected by structural studies as the structural characterization of the protein-protein interactions between p53 and S100 β has shown contradicting results. Reported experimental results have shown disagreements on the stoichiometry of the protein complexes between p53 and S100 β ^{[74][75]}. It was proposed that S100 β binds solely to tetrameric p53, thereby exerting a stabilizing effect of S100 β and protects p53 from degradation^[76]. On the other hand, it was also observed that S100 β binds solely to monomeric p53, thus inhibiting the effect on p53-mediated transcription^[77].

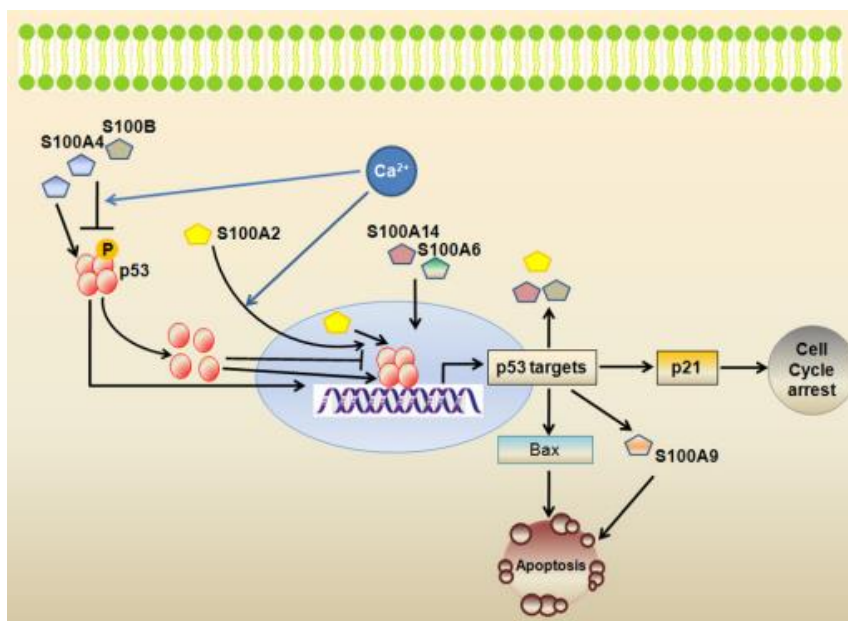


Figure 7. Interaction of tumor suppressor protein p53 with S100 protein family. Several S100 proteins are involved in the regulation of cell growth, cell-cycle progression, and apoptosis by interacting various intracellular signal-regulation pathways. Figure adapted from Chen et al^[63].

As previously mentioned in Chapter 1.3., the interaction between p53 and S100 β is calcium-dependent. Upon addition of calcium, helix 3 (H3) of S100 β is re-oriented and the hydrophobic binding pocket is

exposed, thereby allowing p53 binding (Figure 8). The dual binding pathway of S100 β to p53 (Figure 9), as shown with reported AUC and fluorescence anisotropy experiments, illustrates the contradicting nature of the p53-S100 β interaction. Attempts to characterize the binding affinities of S100 β to the C-terminus of p53 have been made. In Table 1, different dissociation constants were reported using different regions of the p53 peptides. No conclusion can be drawn regarding the binding affinity of S100 β -p53 protein-protein interaction, as there are drastic differences among the dissociation constants. Further studies are necessary to elucidate the true binding affinities between the two proteins.

So far, structural data of the p53-S100 β interaction have exclusively been based on p53 peptide studies and a structural characterization of the interaction between full-length p53 and S100 β is still lacking. A detailed study using full-length, wild-type p53, which will reflect the dynamics of the whole proteins, is urgently needed to increase our understanding about the molecular basis of the p53-S100 β interaction.

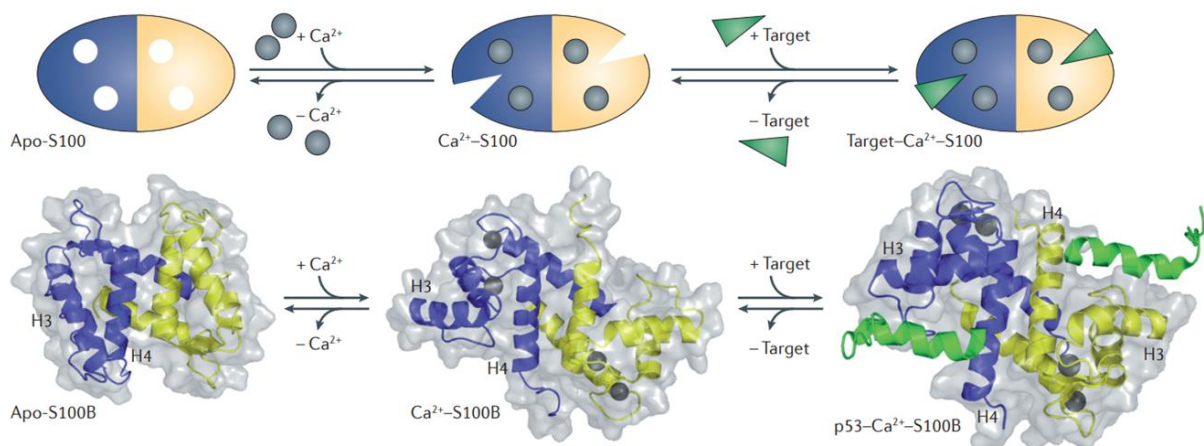


Figure 8. Interaction between S100 β and p53 peptide. The peptide is derived from the Reg of p53. Upon addition of calcium ion, the hydrophobic binding pocket of S100 β is exposed and binding with the p53 peptide takes place. The p53 peptide adopts an alpha-helical structure in the complex with S100 β . Figure adapted from Chen et al^[63].

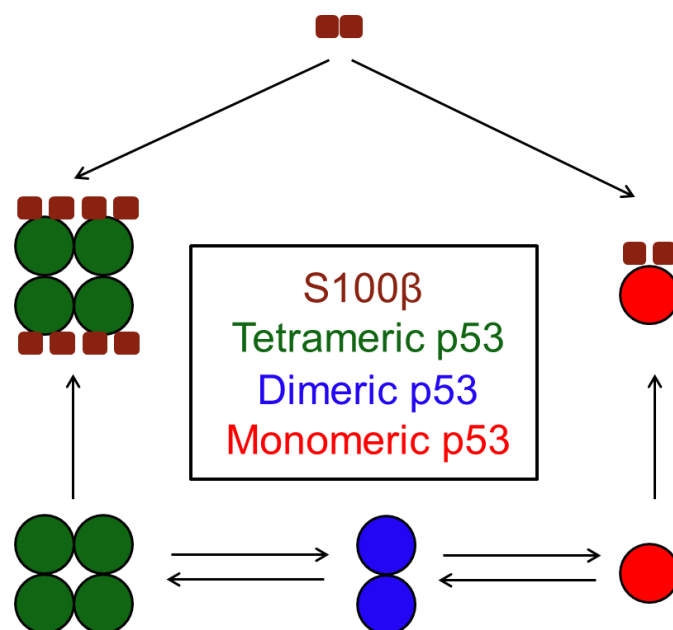


Figure 9. Proposed p53-S100 β binding mechanism based on peptide studies. Two pathways were observed for the protein-protein interaction of S100 β and p53. One is the pathway of S100 β bound to monomeric p53. It inhibits the oligomerization of p53 and suppresses the protein activity (red). The other pathway is S100 β bound to tetrameric p53. S100 β stabilizes the tetrameric state of p53 and its tumor suppressor activity.

Table 1. Reported dissociation constants between p53 peptides and S100 β .

Interaction	Dissociation constant (K_D) (μM)	Method
p53 (aa. 293-393)/S100 β	$0.25 \pm 0.05^{[78]}$	Fluorescence Anisotropy
p53 (aa. 325-339)/S100 β	172 ± 4	
p53 (aa. 325-355)/S100 β	112 ± 7	
p53 (aa. 340-351)/S100 β	302 ± 7	
p53 (aa. 367-393)/S100 β	102 ± 3	
p53 (aa. 305-322)/S100 β	570 ± 50	Analytical Ultracentrifugation
p53 (aa. 325-339)/S100 β	180 ± 20	
p53 (aa. 325-355)/S100 β	77 ± 11	
p53 (aa. 340-351)/S100 β	260 ± 20	
p53 (aa. 367-393)/S100 β	82 ± 6	
p53-K382acetylated (aa. 372-389)/Sir2Tm	$4.3 \pm 0.5^{[79]}$	Isothermal Calorimetry

1.5. Mass Spectrometry

MS measures the mass-to-charge ratio (m/z) of ions. The three main components of a mass spectrometer are the ion source, the mass analyzer, and the detector. Protein samples are first ionized into gas phase ions, with electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) being the two most popular methods for protein analysis using MS (chapter 1.5.1.). After the

ionization step, gas phase ions are passed to the mass analyzer (chapter 1.5.2.), where they are separated according to their m/z values. Over the past decades, mass analyzers such as orbitrap (chapter 1.5.2.1.), quadrupole (chapter 1.5.2.2.), and time-of-flight (TOF), and ion mobility (chapter 1.5.3.) have become popular in the field of protein MS. Hybrid instruments, where two types of mass analyzers are combined, such as quadrupole time-of-flight (Q-TOF) (chapter 1.5.2.3.)^[80], have opened up a new dimension of protein characterization. After ionization and separation of ions according to their m/z , ions reach the detector and mass spectra are recorded.

Different MS methods can be applied depending on the protein system to be studied, which makes it possible to analyze peptides containing only a few amino acid residues or intact protein assemblies by MS^[81]. Chemical cross-linking combined with mass spectrometry (XL-MS) allows the elucidation of site-specific interactions within one protein or between interacting proteins^[82]. By targeting different chemical reactivities within a protein and with the combination of a distant constraint obtained by the chemical cross-links, structural modeling of proteins and protein complexes can be achieved (chapter 1.5.7.). With native MS, measurements of protein samples in their native state can be achieved. By preserving weak non-covalent interactions, structural information, such as the stoichiometry of protein complexes and the composition of protein subunits is obtained (chapter 1.5.6.)^{[23][83]}.

1.5.1. Ionization methods

As mentioned above, the first step in MS analysis is to generate ions that are transferred into the gas phase (Figure 10). In the early days, MS was only able to measure small and thermostable compounds^[84]. Ionization methods, such as electron ionization (EI) uses high energy electrons to generate radical cations^[85] and chemical ionization (CI) involves transferring a proton to generate a protonated ion^[86]. The disadvantage of above-mentioned “hard ionization” methods is the requirement of a large amount of energy to generate ions from the sample. Additionally, the restrictions, such as only being applicable to small masses (<500 Da), the sample requirement to be highly volatile as well as extensive fragmentations in EI-MS, were factors contributing to make the characterization of larger molecules with MS almost impossible for a long time^[87]. Protein MS was made possible in the late 1980s, when the invention of the “soft ionization” techniques, ESI and MALDI opened up new possibilities of studying proteins with MS^{[88][89]}. With ESI and MALDI, large molecules with masses beyond 1000 Da could be transformed into charged gas phase ions. The highly improved ionization efficiency of larger molecules has made these techniques the two most popular ionization methods for studying proteins^[90].

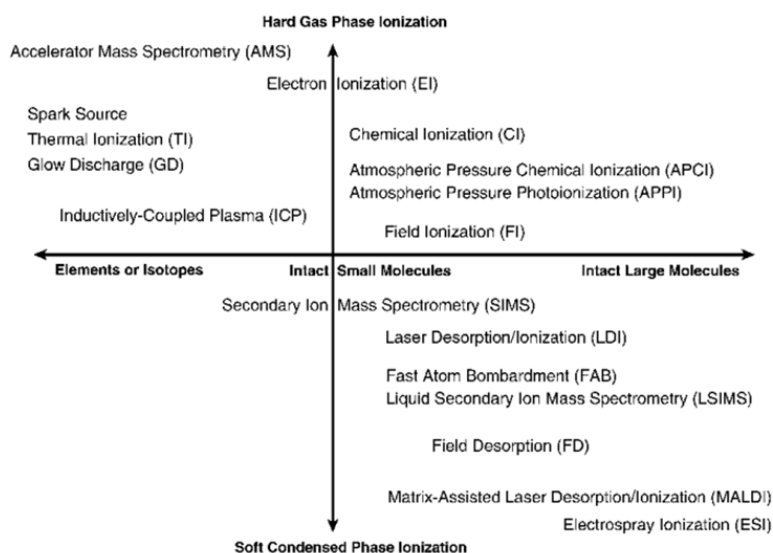


Figure 10. Criteria for selecting ionization method. "Hard ionization" methods represent the early developments of the technique and are most suitable for small and thermostable chemical compounds. The "soft ionization" techniques were developed in the 1980s have been optimal for studying protein samples as little to no fragmentations are generated. Figure adapted from Gross^[91]. Reproduced with permission from Springer Nature.

1.5.1.1. Electrospray ionization

The first observation of electrospray ionization (ESI) was discovered by Sir Geoffrey Taylor and Malcolm Dole^[92]. However, it was John Fenn's research group who made a breakthrough with this technique by showing the possibility of transferring large molecules from liquid aerosols to gas phase ions without fragmentation^[90]. The ionization process begins with a high voltage (>1 kV) being applied to the analyte solution (Figure 11). The surface tension pulls the liquid back into the nozzle, while the electrostatic Coulomb attraction pulls the liquid to the counter electrode. Due to the two competing forces, the elliptical shape transforms into a pointed cone is formed, known as the "Taylor cone". With the combination of the hydrodynamic and electrostatic forces, a spray of charged droplets is emitted from the tip of the cone^[93]. Immediately after the formation of the charged droplets, evaporation of the solvent as neutral molecules takes place, causing an increase in the field density at the surface of droplets and creating charged gas phase ions^[94].

There are two proposed models to describe the final stages of the process where liquid droplets are transformed into gas phase ions. The Ion Ejection Model (IEM), developed by Iribarne and Thomson, suggest that the increasing field strength at the surface causes droplet shrinkage and thereby ejecting the solvated ions from the liquid droplets^[95]. As elucidated by Charge Residue Model (CRM), droplets evaporate in fission cycles into smaller droplets that carry only one analytical ion. It has been suggested by Konermann and coworkers that the two different models work for analytes with different MWs and

sizes in principle^[96]. The IEM model applies to species with a lower MW and the CRM model can be applied to larger and globular analytes^[97]. ESI is known as a “soft-ionization” technique as the voltage used generates little to no fragmentation, which allows measuring intact and weakly bound protein complexes^[98].

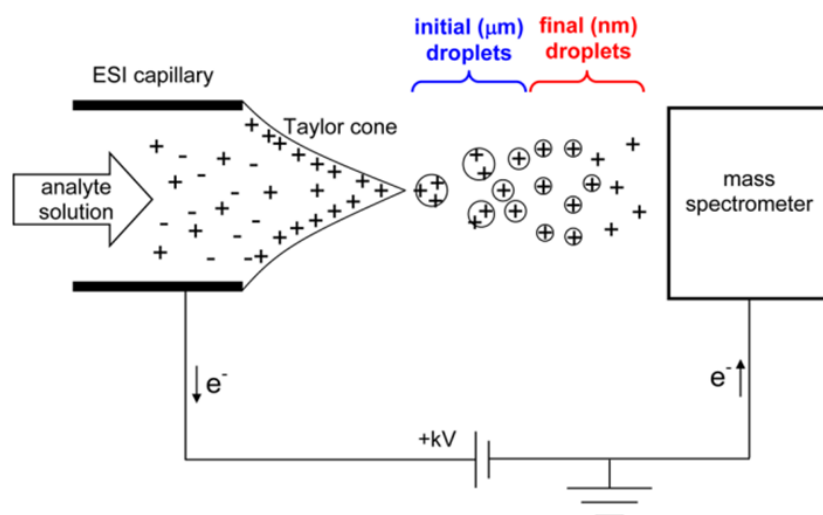


Figure 11. The ESI process in positive ionization mode. A high voltage of several kV is applied to the analyte solution. The “Taylor cone” is formed due to the two competing forces, the surface tension and the electrostatic Coulomb attraction. A spray of charged droplets is generated due to electrostatic and hydrodynamic forces. Gas phase ions are generated due to solvent evaporation. Figure adapted with permission from Konermann et al. Copyright 2013 American Chemical Society^[99].

1.5.1.2. Nano-electrospray ionization

Nano-electrospray ionization (nano-ESI) offers various advantages in comparison to ESI. As suggested by the name, nano-ESI is a miniature version of ESI. It is performed with a lower solution flow rate (1-5 nL/min), a lower applied voltage (0.5 – 1.5 kV), and the sample is introduced by a thinner capillary (1 μM inner diameter). All these conditions effectively reduce the size of the charged droplets during the spraying process^[100]. Smaller droplets offer an advantage, as they come with a higher surface-to-volume ratio, making a larger proportion of analyte available for desorption^[101].

Listed below are the benefits that have made nano-ESI the optimal choice for protein MS. The requirement of smaller sample volumes makes nano-ESI a better choice as often limited amounts of protein sample are available^[102]. It has been shown that nano-ESI is much more tolerant towards medium to high salt concentrations in protein samples compared to conventional ESI^[103]. With the improved desolvation in nano-ESI, the sensitivity and mass resolving power have both been improved^[101].

1.5.2. Mass analyzers

1.5.2.1. Orbitrap mass analyzer

First implemented by Kingdon in 1923, a Kingdon trap was equipped with a stretched wire along the axis of an outer cylinder with flanges enclosing the trap^[104]. When the ions were introduced into this trap, the voltage applied between the wire and the cylinder introduced an electrostatic field and only the ions with enough tangential velocity would stay in the orbit and survive^[105]. The orbitrap mass analyzer, built on the principle of trapping ions with an electrostatic field, was invented by Makarov in the late 1990s^{[105][106]}. Ions enter the orbitrap mass analyzer and experience a strong radial and axial electrostatic field. The radial field sets ions on a circular orbit around the central electrode and the axial field accelerates the ions^[107]. There was a shortcoming for the early orbitrap mass analyzer, which was the lack of control over the number of ions injected into the mass analyzer. This was later overcome with the invention of a C-trap^[108] (Figure 12a). With the addition of the C-trap, ions can be accumulated and stored prior to injection into the orbitrap. The combination of the C-trap with a continuous ion source, such as the ESI source, and the orbitrap mass analyzer has seen great success and has since been applied to modern orbitrap instruments^[109]. The orbitrap mass analyzer has a high mass accuracy of less than 2 ppm. A high resolving power is also an advantage of the mass analyzer, as up to 100,000 full width at half maximum (FWHM) has been reported. However, the resolving power decreases with increasing m/z ^{[110][111]}. A hybrid configuration is common in commercial orbitrap instruments (chapter 1.5.5.2.). For example, a linear ion trap combined with an orbitrap (LTQ-Orbitrap, Thermo Fisher Scientific) or a combination of a linear ion trap, a quadrupole, and an orbitrap (Orbitrap Fusion Tribrid, Thermo Fisher Scientific) are some of the widely popular commercial instruments for proteomics applications^[112]. The mass-to-charge ratio (m/z) of the orbitrap mass analyzer can be calculated from the given force constant of the potential (k) and the measured axial oscillation frequency (ω) (Figure 12b).

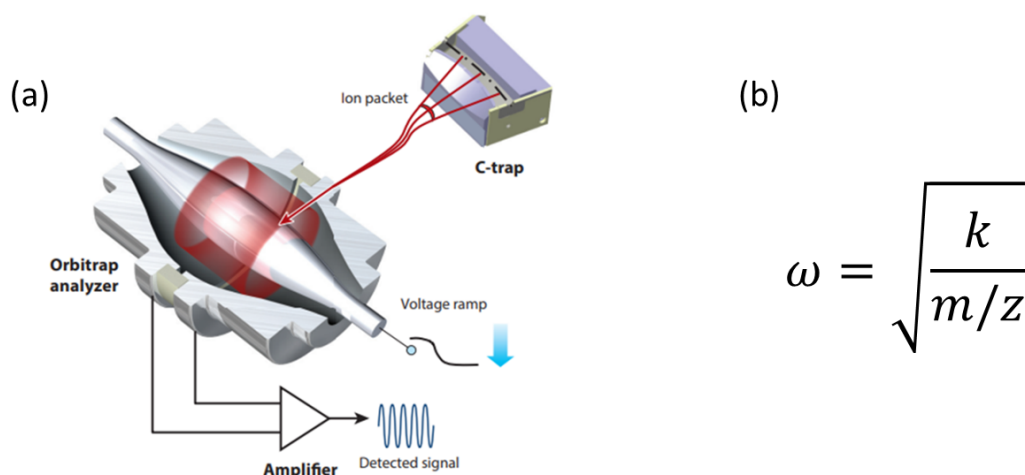


Figure 12. Schematic representation of an orbitrap mass analyzer and a C-trap in a commercial instrument (Thermo Fisher Scientific). (a) Ions are introduced into the orbitrap mass analyzer via the C-trap during a voltage ramp. The formation of rings from the injected ions induces a current in the outer electrodes that is detected by the amplifier. Figure adapted from Eluik et al^[107]. (b) The mass-to-charge ratio (m/z) can be calculated from the axial oscillation frequency (ω) and the force constant of the potential (k) measured with the orbitrap analyzer.

1.5.2.2. Quadrupole mass analyzer

The quadrupole, first described by Paul and Steinwedel^[113], was originally used as a tool for isotope separation^[114]. The linear quadrupole was developed for MS by Paul and the invention has since been widely used as an ion guide and a mass filter^[115]. Composed of four cylindrical-shaped rod electrodes extending in the z -direction, two pairs of opposite electrodes are held at the same electric potential (Figure 13). As ions enter the quadrupole in the z -direction, an electric force is exerted on the ions. The attraction and the repulsion creates a trajectory and causes the ions to oscillate along the x - and y -directions^[116]. Only ions of a certain m/z value with a stable trajectory will reach the detector. Those with an unstable trajectory will collide with the rods and not be detected^[117].

Recent developments have seen more rods added to the mass analyzer, making it a hexapole (six rods) or an octapole (eight rods). A higher number of electrodes provide steeper electric potential wells, offering better ion-guiding capabilities and a of m/z range^[118]. Many modern instruments contain more than one quadrupole, such as triple quadrupole (QqQ) (1.2.4.3) or Qq-TOF instruments (chapter 1.5.5.1.).

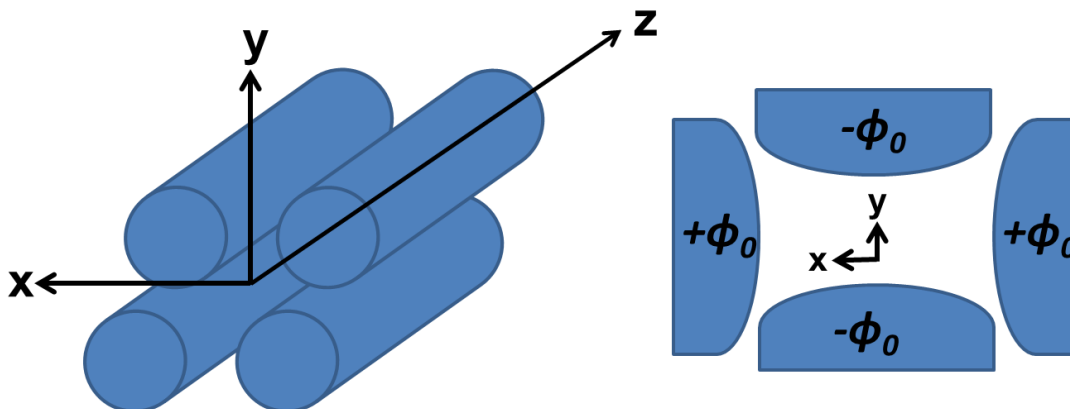


Figure 13. Schematic representation of a linear quadrupole. Ions enter the quadrupole along the z-axis and oscillate along the x- and the y-axes. Ions with a stable trajectory will reach the detector, while the ones with an unstable trajectory will collide with the rods.

1.5.2.3. Quadrupole time-of-flight mass analyzer

The Qq-TOF (or Q-TOF) mass analyzer (Figure 14) comprises two quadrupoles (Q1 and Q2) and a TOF analyzer. Q1 acts as a mass filter, Q2 acts as the collision cell, and the TOF analyzer separates the ions based on their m/z values. The reflectron allows ions with the same m/z values, but different kinetic energies, to arrive at the reflector detector at the same time^{[119][120]}.

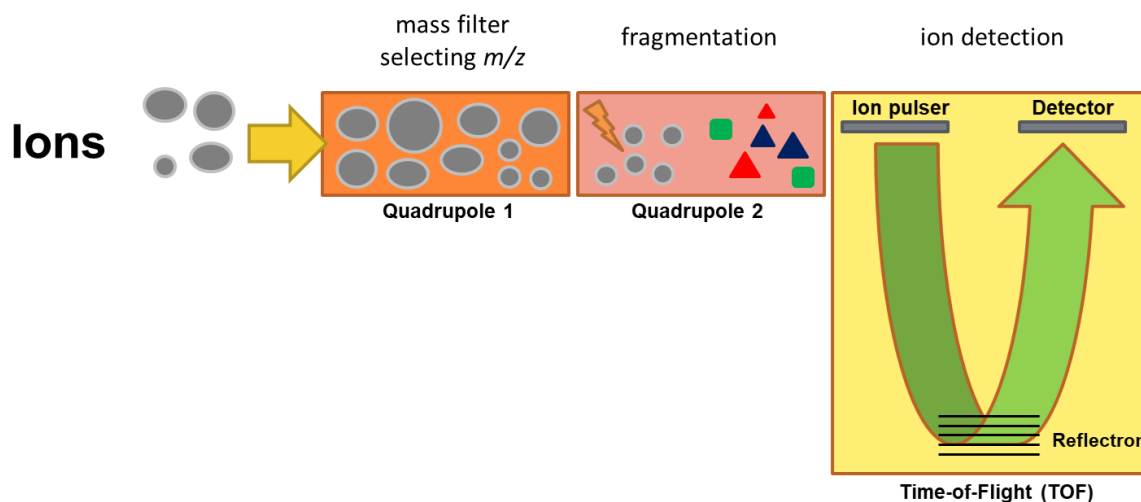


Figure 14. Schematic representation of a quadrupole time-of-flight mass spectrometer. The Qq-TOF mass analyzer comprises a quadrupole 1 (Q1), a quadrupole 2 (Q2), and a TOF mass analyzer. For a full MS scan, Q1 and Q2 act as ion guide. For tandem MS (CID-MS/MS), Q1 acts as a mass filter, Q2 acts as the collision cell, and the TOF analyzer separates the ions based on their m/z values.

1.5.3. Ion mobility spectrometry

Ion mobility spectrometry (IMS) separates ions based on their mobility with the help of a buffer gas^[121]. While IMS instruments can be used as stand-alone devices, they can also be coupled to high-performance liquid chromatography (HPLC) and MS as an additional dimension to separate ions based on their size, shape, and charge state^[122] (Figure 15a). IMS can be calculated by the equation shown in Figure 15b^[123]. Ions with a smaller CCS experience a smaller separation field strength with the buffer gas and move through the chamber faster than larger ions^[124]. Also, ions carrying a higher charge state will have much greater separation field strength and will move across the chamber faster than the ones with a lower charge state^[125].

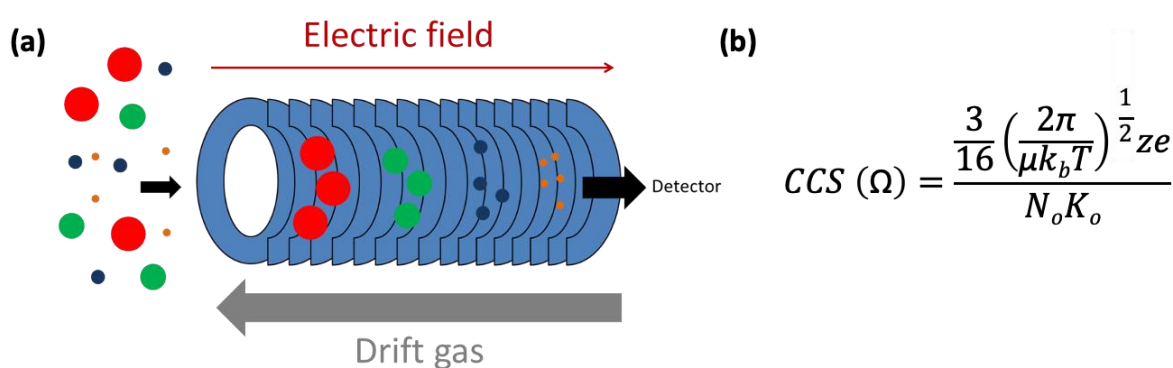


Figure 15. (a) Schematic of the ion-mobility spectrometry (IMS) and (b) Calculation of ion-mobility with the Mason-Schamp equation^[126]. Ions are separated based on their size, shape, and the charge state. $CCS(\Omega)$: collisional cross section of the ion in the drift gas, e : charge of an electron, z : ion charge, N_o : buffer gas density, K_o : ion mobility, k_b : Boltzmann's constant, T : drift region temperature, μ : reduced mass of the collision partners.

Operating with the same principle of ion mobility, small variations among different modes of ion mobility allow optimization based on different applications^[127]. In drift time ion mobility spectrometer (DTIMS), ions move through the chamber carrying a homogeneous electric field filled with neutral gas molecules, usually nitrogen^[128]. In travelling-wave IMS (TWIMS), opposite phases of radio-frequency voltages are applied to the ring electrodes generating a traveling wave. Ions advancing through the traveling wave result in higher transmission efficiency^[129]. High-field asymmetric waveform IMS (FAIMS) exploits the differences of ion mobility in higher electric fields and implements a periodic asymmetric waveform to separate ions^[130]. For trapped ion mobility spectrometer (TIMS), the ion mobility chamber is divided into different parts. A non-uniform electric field is applied to hold the ions stationary and trapped ions are subsequently released with decreasing electric field based on their collisional-cross-section-to-charge ratios^[131]. Many modern hybrid mass spectrometers incorporate IMS as a part of the instrument, such as ion mobility quadrupole mass spectrometry (IM(q)MS) or ion mobility time-of-flight mass spectrometry (IM-TOF-MS)^{[132][133]}.

1.5.4. Fragmentation

While soft ionization techniques, such as ESI (chapter 1.5.1.1.) extended the application of MS to protein studies, structural information at the peptide level is still inadequate when only masses of intact proteins and peptides are measured^[134]. Tandem mass spectrometry (MS/MS or MS²) overcomes this obstacle^[135]. With MS/MS, fragmentation of the precursor protein or peptide ion is made possible through multiple collisions with inert gas atoms, generating peptide fragments where backbone cleavages at different positions are observed (Figure 16). Based on the position of the backbone cleavage, the nomenclature of peptide sequencing is as follows: a-, b-, and c-type fragment indicate the peptide fragment with the charge retained at the *N*-terminus, while in x-, y-, and z-type fragments the charge resides at the *C*-terminus^[136]. For protein studies, collisional activation methods are most commonly used for fragmentation^[137].

The two main collisional activation methods are collisional-induced dissociation (CID) and higher energy collision-induced dissociation (HCD). CID is a resonance-type dissociation, meaning that mainly the precursor ion is fragmented. The resulting fragment ion mass spectrum indicates the fragments generated from one precursor ion^[138]. HCD is a beam-type dissociation where all ions within the chamber are activated, permitting multiple collisions. As the number of ions that can be activated increases, richer fragment ion mass spectra are generated, exhibiting more internal peptide fragments and neutral loss fragments^[139].

A shortcoming of CID is the cleavage of labile bonds^[140]. Information regarding the fragile bonds such as post-translational modifications (PTMs) can sometimes be lost during collisional activation. Electron-based dissociation methods, such as the electron-transfer dissociation (ETD) and electron-capture dissociation (ECD), can overcome this issue. With ETD and ECD, fragmentation induced by electrons cleaves the N-C_α bond and generates c- and z-type peptide fragments^[141]. ECD and ETD are non-ergodic process, meaning that bond cleavage takes place before the vibrational energy is redistributed. With that, labile bonds, like PTMs, are usually preserved during this process^[142]. As the collision-induced and the electron-based fragmentation methods yield complementary information, a combination of the two provides a powerful approach for proteomics studies and greatly reduces the risk of misidentification of peptides and proteins^[143].

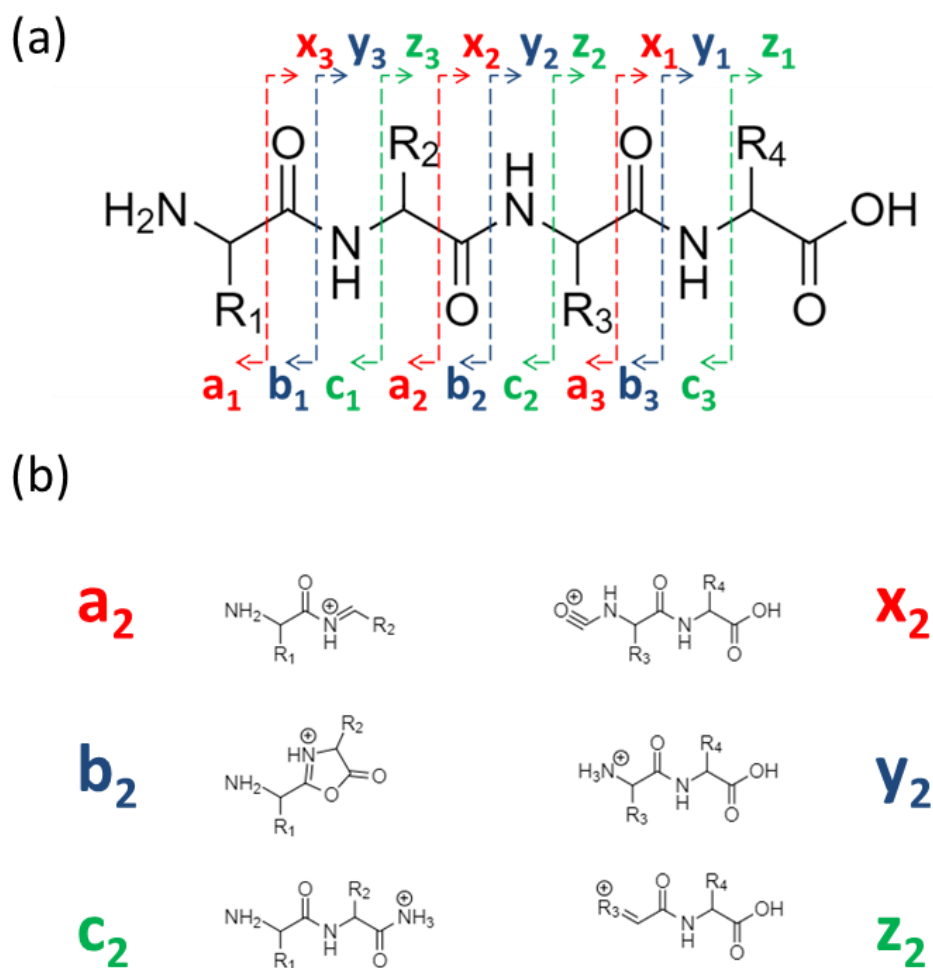


Figure 16. (a) Nomenclature of MS/MS peptide fragments as proposed by Roepstorff, Biemann, Fohlman, and Johnson^[144]. (b) Proposed structures of peptide fragment ions. The amino acid position of the bond cleavage is given as number; a-, b-, and c-type ions indicate fragments with the charge retained at the *N*-terminus; x-, y-, and z-type ions indicate fragments with the charge retained at the *C*-terminus.

1.5.5. Mass spectrometers used for this work

1.5.5.1. Q-TOF mass spectrometer

Constructed in 1996, the High-Mass Q-TOF II mass spectrometer (Micromass/MS Vision) (Figure 17) is a hybrid instrument that contains two quadrupoles and an orthogonal TOF analyzer. The instrument was first used to investigate the structural properties of major histocompatibility complexes (MHC)^[145] and it is a suitable instrument for native MS (chapter 1.5.6.) or top-down proteomics. With the optimization of analyzing ions with high masses, it is possible to study large intact protein complexes in terms of stoichiometry and composition with this instrument. The lower radiofrequency of the

quadrupole extends the range of the measurable m/z values and the high pressure in the collision cell has made dissociation of the protein complexes with collision of neutral gas molecules possible^[146]. A MS full-scan allows visualization of all the species present in the sample. In order to get insights into the composition of protein complexes and their topology, MS/MS experiments can be performed with the High-Mass Q-TOF II mass spectrometer by collisional activation. Here, the selected species is kept in the collision cell and by collision with an inert collision gas (nitrogen gas or argon gas)^[147] causing the protein subunits to be ejected from the complex^[148].

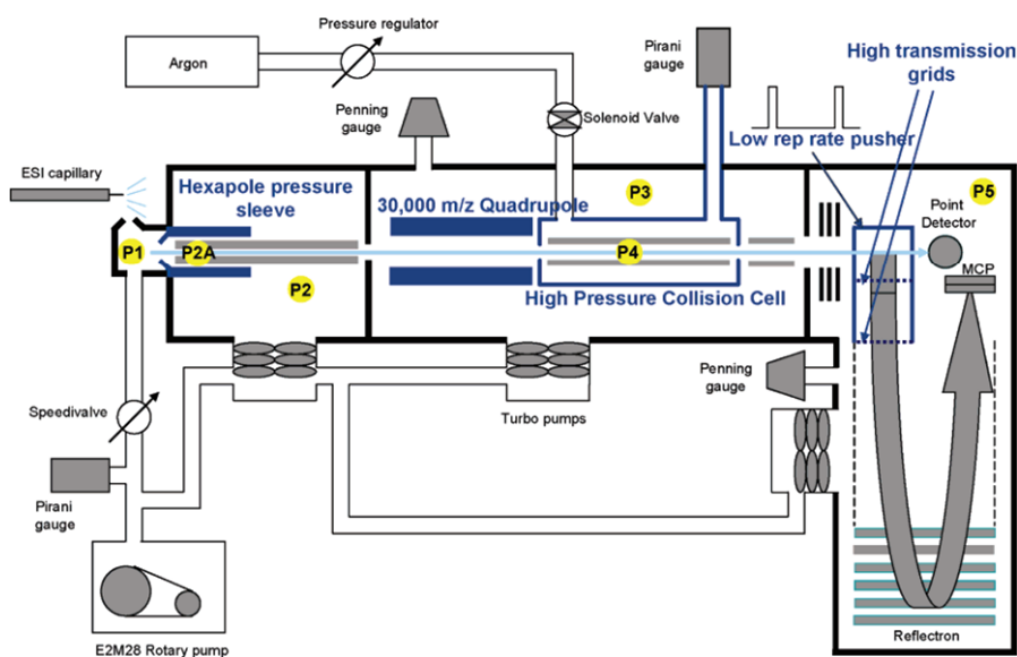


Figure 17. Schematic layout of the High-Mass Q-TOF II mass spectrometer (Micromass/MS Vision). The sample is introduced into the instrument via an ESI source at atmospheric pressure and the ions enter the first quadrupole, followed by the second quadrupole, and then proceed to the TOF analyzer where they are detected. MCP: multi-channel plate, P1: the first vacuum stage, P2: second pumping stage, P2A: estimated average pumping pressure: 8×10^{-3} mbar, P3: third vacuum chamber, P4: collision cell, P5: TOF vacuum chamber. Figure adapted from van den Heuvel et al^[149].

1.5.5.2. Orbitrap mass spectrometer

The Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) is a hybrid instrument combining a quadrupole, an orbitrap, and a linear ion trap (Figure 18). Ions are first injected via an ESI source. After passing through the ion beam guide where neutral particles are removed, ions arrive at the quadrupole mass filter where precursor ions are either transferred or selected. Various fragmentation techniques can be performed in this instrument type. HCD takes place at the ion routing multipole where the beam-type fragmentation results in secondary fragmentations and peptide backbone fragments^[150], providing good sequence coverage for protein identification. CID and ETD

occur at the high pressure cell of the linear ion trap^[151]. A module can be also added at the back of the instrument to perform ultraviolet photo-dissociation (UVPD)^[152]. For proteomics studies, data-dependent acquisition (DDA) is performed via a so-called “TopN” approach where *N* qualifying precursor ions with the highest intensities are fragmented. Specific precursor ions can also be selected based on their *m/z* values and their charge states using an inclusion list to perform targeted fragmentation^{[84][153]}.

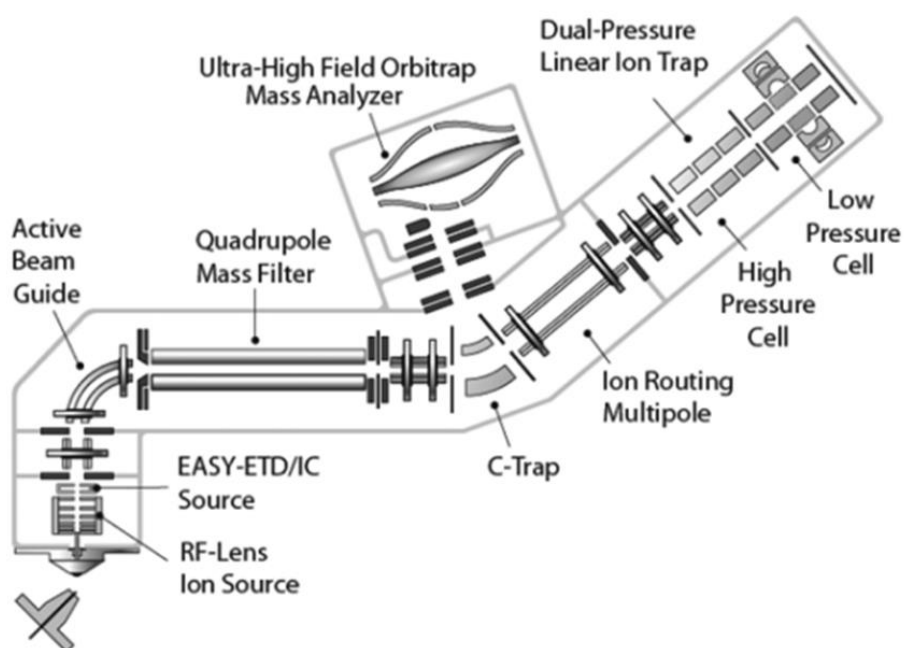


Figure 18. Schematic layout of the Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). This instrument contains a quadrupole, an orbitrap, and an ion trap. Ions can be stored at C-trap prior to injection. The high pressure cell performs CID. RF: radio frequency, ETD: electron transfer dissociation, IC: internal calibration. Figure adapted with permission from Senko et al. Copyright 2013 American Chemical Society^[154].

1.5.5.3. Triple quadrupole mass spectrometer

The triple quadrupole is a combination of three quadrupole mass analyzers where each of the quadrupoles plays a different role (Figure 19). The first quadrupole (Q1) is a mass filter. Only ions with defined *m/z* values can pass through Q1. The second quadrupole (Q2) acts as a collision cell. Here, with the aid of an inert gas, ions are fragmented. The third quadrupole (Q3) either scans the product ions after fragmentation or it can also be set to monitor ions of just a particular *m/z* values^[155]. Four scan modes are possible in triple quadrupole analyzers: i) in product ion scan, Q1 is set to a fixed mass of ions that can pass through the first quadrupole, these ions are then fragmented in Q2, and an entire *m/z* range of product ions is scanned at Q3. This mode allows the identification of the precursor ions based on its ion fragments. ii) In the precursor ion scan, Q1 is set to scan all precursor ion masses,

while Q3 is set to a fixed m/z value of a product ion. In this mode, particular functional groups can be identified. iii) For the neutral loss scan, Q1 and Q3 are both set at fixed masses. With a known constant mass difference, this mode allows selective recognition of the loss of the neutral fragments, for example, a loss of water ($-H_2O$) or ammonium ($-NH_3$), upon fragmentation in Q2. iv) In Selected Reaction Monitoring (SRM) or Multiple Reaction Monitoring (MRM) modes, both Q1 and Q3 can be configured to more than one m/z value, thereby allowing observation of different precursors and the detection of different product ions^{[156][157]}.

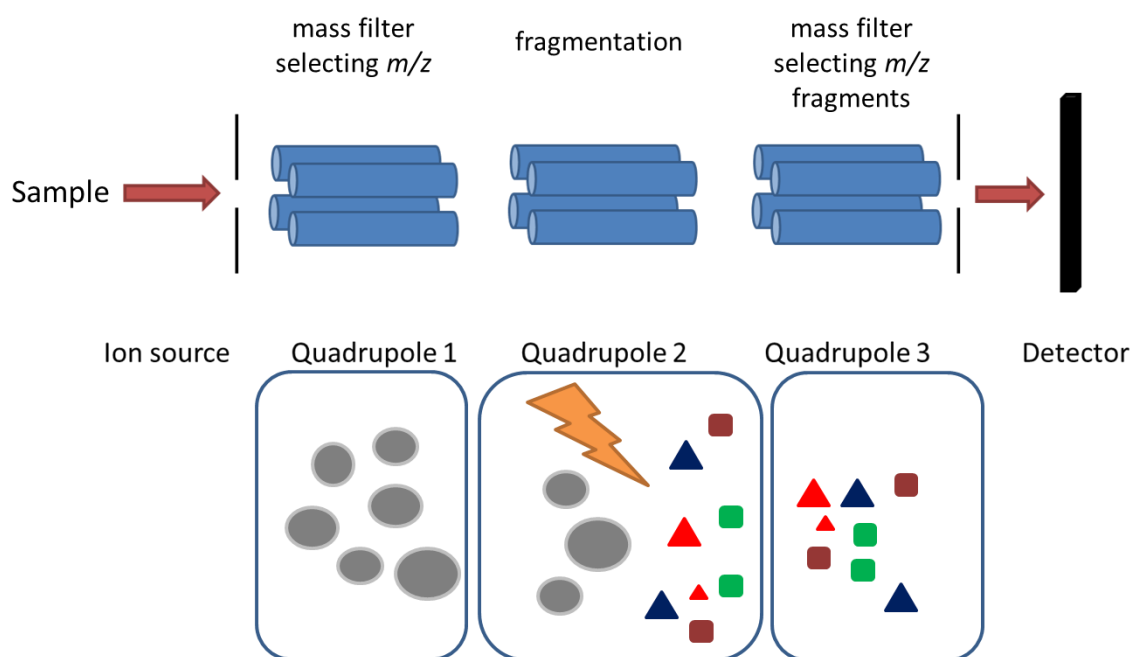


Figure 19. Schematic representation of a triple quadrupole mass spectrometer (TQD). In full scan mode, all quadrupoles act as ion guides. In different scan modes, Q1 and Q3 act as mass filters for fixed m/z values, while Q2 acts as collision cell for ion fragmentation.

1.5.5.4. Ion mobility Q-TOF mass spectrometer

The timsTOF Pro mass spectrometer (Bruker Daltonik) is a hybrid instrument that consists of a trapped ion mobility (TIMS) chamber, quadrupoles, and a TOF mass analyzer (Figure 20). A novel method, termed “Parallel Accumulation Serial Fragmentation (PASEF)” (Figure 21), optimizes the ion separation by combining precursor ion selection with TIMS separation. Ions are introduced into the instrument with the ESI source then deflected by an angle of 90° before entering the TIMS chamber, where they are trapped and stored with the applied voltage. Once a sufficiently high number of ions is accumulated, the decreasing voltage releases the ions sequentially, based on their masses and mobilities^[158]. In contrast to traditional IMS, the release of ions works in the inverse order in TIMS where the ions with a higher mass and lower mobility are released first, followed by the ions with a lower mass and a higher

mobility^[159]. Upon exiting the TIMS analyzer, ions are transferred through the ion transfer multipole, enter the quadrupole mass filter for precursor ion selection, and then the quadrupole collision cell for fragmentation. In PASEF mode, fast switching times of the quadrupoles are required in order to keep up with the pace of ion separation from the TIMS chamber. The main difference of PASEF in comparison to the conventional TIMS method is that it performs ion trapping and release in parallel^[160]. Due to maximizing the number of precursor ion scans in the first quadrupole prior to fragmentation in the collision cell, it was possible to increase the efficiency of the measurement without sacrificing the sensitivity. The TOF analyzer at the end of the instrument allows efficient acquisition speed in the millisecond time scale.

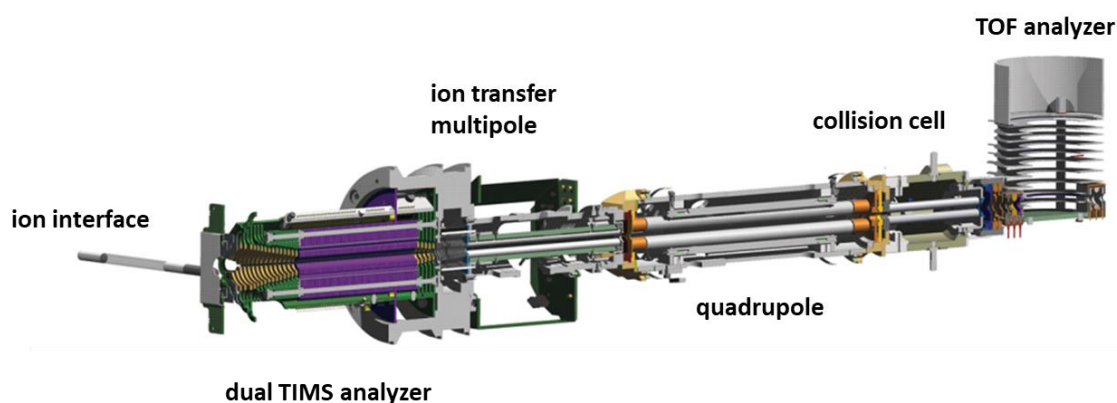


Figure 20. Schematic layout of the timsTOF Pro mass spectrometer (Bruker Daltonik). The instrument contains a TIMS analyzer, an ion transfer multipole, a quadrupole, and a TOF mass analyzer. Adapted from Meier et al. Copyright 2015 American Chemical Society <https://pubs.acs.org/doi/10.1021/acs.jproteome.5b00932>. Further permissions related to the material excepted should be directed to the ACS.^[161]

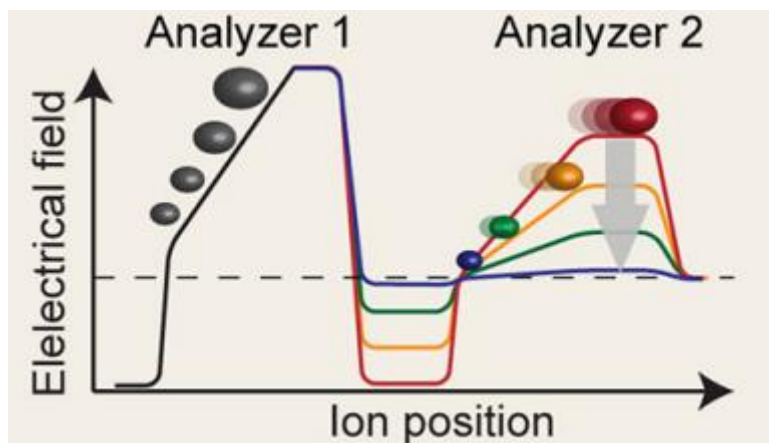


Figure 21. Illustration of a Parallel Accumulation Serial Fragmentation (PASEF) method in timsTOF Pro mass spectrometer. Two steps are involved in a PASEF scan. Ions enter the TIMS chamber and are trapped and stored with an applied voltage. Figure adapted from Meier et al^[162].

1.5.6. Native mass spectrometry

Under non-denaturing condition, native MS has become a widely popular technique to study the stoichiometry and the composition of protein complexes. The term “native” describes the native biological state of the protein samples in solution prior to the ionization into the gas phase ions^[163]. The non-covalent interactions such as hydrogen bonds, electrostatic interactions, and π - π interactions play a huge role in large biomolecules and they are also responsible for the folding of protein and the tertiary and the quaternary structures^[164]. The “soft” ionization technique ESI preserves these weak non-covalent interactions and is therefore applied in native MS^[165]. Protein samples are prepared at a low concentration (1-10 μ M) in a volatile buffer, commonly ammonium acetate, at nearly neutral pH. This should allow retaining the non-covalent interactions that are present at physiological conditions.

Protein analyses with native MS require instruments that can analyze ions with extended m/z values. This was enabled by the introduction of the TOF mass analyzer, which theoretically has no upper mass range limit^{[166][167]}. Hybrid instruments equipped with a combination of different mass analyzers, such as the Q-TOF instruments (chapter 1.5.5.1.), have become the most widely used mass spectrometers to study protein assemblies up to the Megadalton range (MDa) by native MS due to the high sensitivity and its good resolving power^{[168][169]}.

Tandem MS or MS/MS (chapter 1.5.4.) is commonly applied in native MS measurements. As the precursor ion is activated, dissociation of protein assemblies is observed. With that, structural information such as the stoichiometry and composition of the protein subunits is obtained^{[170][171]}.

1.5.7. Cross-linking mass spectrometry

XL-MS has evolved as a popular technique, especially in integrative approaches, to study protein-protein interactions. By introducing a covalent bond with a distance constraint between proximal amino acid residues, information on three-dimensional protein structures and protein interaction sites can be gained. XL-MS has become a widely used approach for structural proteomics as it offers many advantages, such as the requirement of only femtomole amounts of sample, the possibility to study heterogeneous protein samples, and unlimited protein mass^{[172][173][82]}.

In principle, two experimental approaches can be performed in XL-MS. i) The top-down approach where intact cross-linked protein complexes are analyzed; and ii) the bottom-up approach where samples are digested with proteases, such as trypsin, to yield peptides that might then be analyzed. The top-down approach was first reported with Fourier transform ion cyclotron resonance (FTICR)-MS, however, it has remained useful for proteins with a smaller size, thereby limiting the applicability for XL-MS^[174]. The bottom-up approach proves to be more popular and flexible, as the analyses of peptides are faster and many different software solutions are available for an automated assignment of the cross-linked products. First, a cross-linker is introduced into the protein samples and incubated. Then, the cross-linking reaction is quenched, the proteins are digested with proteolytic enzymes, and the resulting peptide mixtures are separated by HPLC. The outlet of the HPLC is coupled to mass spectrometers via an ESI source. Precursor ions of peptides are selected and fragmented with collisional activation (CID and HCD), the amino acid sequences are determined, and the site-specific interactions are mapped.

Over the past few decades, XL-MS has evolved from studying single proteins to studying larger protein interaction networks. Nowadays, XL-MS can even be applied to whole proteomes, cells, tissues, and organisms^{[175][176][177]}. Cross-linkers can be divided into different chemical reactivities and spacer lengths. Cross-linkers that connect residues with the same reactivity are known as homo-bifunctional cross-linkers, for example, ureido-4,4'-dibutyric acid bis(hydroxysuccinimide) ester (disuccinimidyl dibutyric urea, DSBU), which cross-links mainly lysine residues (Figure 22). Those cross-linkers possessing one type of reactivity on one end and another type of reactivity on the other end are called hetero-bifunctional cross-linkers, such as sulfosuccinimidyl 4,4'-azipentanoate (sulfo-SDA) (Figures 24 and 25). A combination of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and sulfo-N-hydroxysulfosuccinimide (EDC/sulfo-NHS) (Figure 23), connecting lysine residues and acidic residues, such as glutamic and aspartic acid, presents a so-called "zero-length" cross-linker. The above-mentioned cross-linkers can be further divided into two categories, chemical- and photo-cross-linking. Chemical cross-linking, like DSBU and EDC/sulfo-NHS, refers to the connection of two functional groups by introducing a covalent bond^[178]. Photo-cross-linking with sulfo-SDA involves a cross-linking reaction

with a diazirine moiety. Upon ultraviolet irradiation, a carbene or a diazo moiety is generated and amino acid residues in close proximity are joined^[179]. Cross-linkers that connect two residues that also possess a third functional group that allows enrichment of the cross-linked samples are known as enrichable trifunctional cross-linkers. Examples for trifunctional cross-linkers are the class of Protein Ion Reporters (PIRs) or 1,1-bis{4-[(2,5-dioxopyrrolidin-1-yl)oxy]-4-oxobutyl}-4-ethynylpiperidin-1-ium (PAC4)^{[180][181]}. PIRs contain a biotin group as an affinity handle and PAC4 utilizes the cycloaddition click chemistry for enrichment. PIRs are being used for studying protein interaction networks in cells, organisms, and tissues, while PAC4 has not experienced a widespread distribution so far^{[182][183]}.

1.5.7.1. The cross-linker disuccinimidyl dibutyric urea (DSBU)

In this work, DSBU was one of the mainly used cross-linkers. DSBU targets amino groups in proteins (lysines and the *N*-termini), but exhibits a side reactivity towards hydroxyl residues (serines, threonines, and tyrosines). DSBU belongs to the class of MS-cleavable cross-linkers^{[184][185]}.

MS-cleavable cross-linkers exhibit a labile covalent bond and show specific fragmentation in the spacer chain upon collisional activation^[186]. In DSBU, fragmentation of the NH-CO bond at the central urea group in the spacer generates characteristic 26-u doublets in the fragment ion mass spectrum. The advantage of DSBU as a cross-linker is that the energy required for peptide backbone fragmentation is roughly the same as the energy required for cleavage of the urea moiety. This allows the observation of the peptide fragments and the signature 26-u doublets already at MS/MS without the need to perform MS³ experiments (Figure 22).

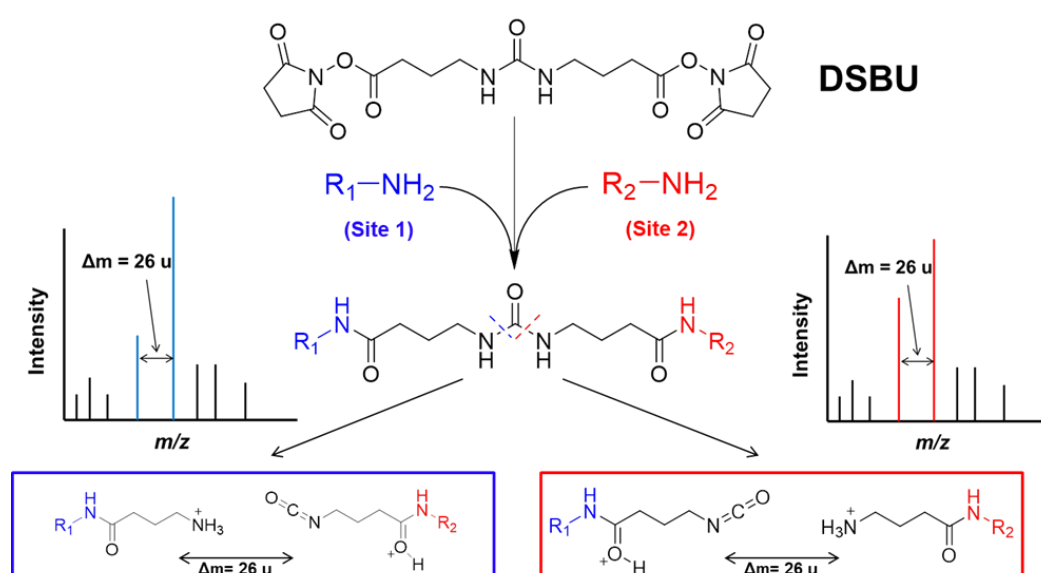


Figure 22. Cross-linking chemistry of DSBU MS-cleavability by collisional activation. Upon CID (or HCD), two doublet signals with a mass difference of 26 Da are observed (blue and red) in the fragment ion mass spectra. The characteristic doublets along with the peptide backbone fragmentation at the MS/MS level is utilized by data analysis software MeroX^{[187][188]}.

1.5.7.2. The cross-linker 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and sulfo-N-hydroxysulfosuccinimide (EDC/sulfo-NHS)

Another chemical cross-linking principle used in this thesis to target acidic and amine-reactive residues is the combination of EDC and sulfo-NHS (Figure 23). As mentioned above, carboxylic acids need activation to undergo a cross-linking reaction. This activation can be circumvented by the addition of the carbodiimide EDC, which targets carboxylic groups of acidic residues (glutamic and aspartic acids) and forms an ester intermediate. In the presence of a strong nucleophile it undergoes a nucleophilic substitution and forms an unstable O-acylisourea intermediate that is prone to hydrolysis^[189]. With the addition of sulfo-NHS, a more stable reaction intermediate, an NHS-ester, is formed. This can react with amine groups (in lysines), thereby directly connecting carboxylic acid and amine groups in proteins^[190]. A two-step coupling of amine and acidic residues with EDC/sulfo-NHS has several advantages over a one-step strategy as generally higher cross-linking yields have been observed^[191]. One drawback of cross-linking with EDC is its optimum pH range at low pH values, which might be problematic if proteins have to be studied under native pH conditions.

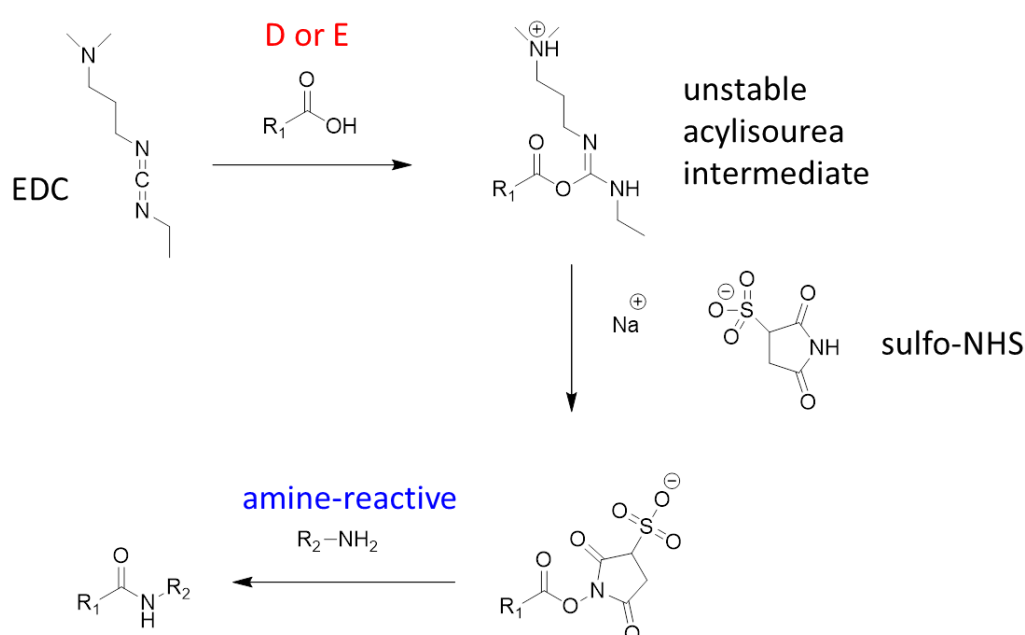


Figure 23. Cross-linking chemistry of EDC/sulfo-NHS cross-linkers. Carboxylic acid groups are activated by the addition of the carbodiimide EDC. Upon formation of the unstable O-acylisourea intermediate, the addition of sulfo-NHS allows forming a more stable intermediate that is able to react with a lysine residue (D = aspartic acid, E = aspartic acid).

1.5.7.3. The cross-linker sulfosuccinimidyl 4,4'-azipentanoate (sulfo-SDA)

The third cross-linker used in this work is sulfo-SDA, which is activated upon UV irradiation and allows targeting different reactive groups in proteins. Diazirine groups are a common feature of photo-cross-linkers that are widely used for photo-affinity labeling of proteins^[192]. Upon absorbing UV-A light (wavelengths between 350 nm and 380 nm), two competitive reaction pathways occur simultaneously. Either the diazirine undergoes photolysis to a singlet carbene by the loss of nitrogen (carbene pathway)^[193], or isomerization takes place and the diazirine turns into a linear diazo compound (diazo pathway) (Figure 24)^{[194][195]}. The carbene pathway involves the insertion of polar bonds as the carbene intermediate is electrophilic. Insertions of C-H bonds can also occur in the absence of better nucleophiles. This pathway is considered to account for the minority of the cross-linking reactions with diazirines. In the diazo pathway, the carboxylic side chains of proteins are targeted and converted to esters, which accounts for the majority of the cross-linking reactions with diazirine. Sulfo-SDA is a hetero-bifunctional cross-linker. On one end, sulfo-SDA contains the amine-reactive NHS-ester with the ability to target lysine residues and the *N*-termini in proteins. A photo-reactive diazirine group on the other end of the cross-linker sulfo-SDA, allows reactions to all 20 proteinogenic amino acids, with a preference for acidic amino acids (aspartic and glutamic acids). Typically, targeting carboxylic acid residues involves a step of activation, usually with a carbodiimide, such as EDC (chapter 1.5.7.2.) that requires low pH which is not suited to keep physiological or native conditions for proteins. On the other hand, using the diazirine moiety from sulfo-SDA allows the cross-linking reaction to proceed at physiological conditions with no prior activation step is required. This enables the preservation of the native fold of proteins and protein complexes. Furthermore, software analysis of sulfo-SDA cross-linked products is achievable with the MeroX software. The MS-cleavability of diazirine cross-links formed with acidic groups (Figure 25) shows that the fragmentation of the cross-linked product upon collisional activation generates a peptide with a characteristic +82 u (+C₅H₆O) modification and another peptide representing the acidic residue with no mass modification^[196].

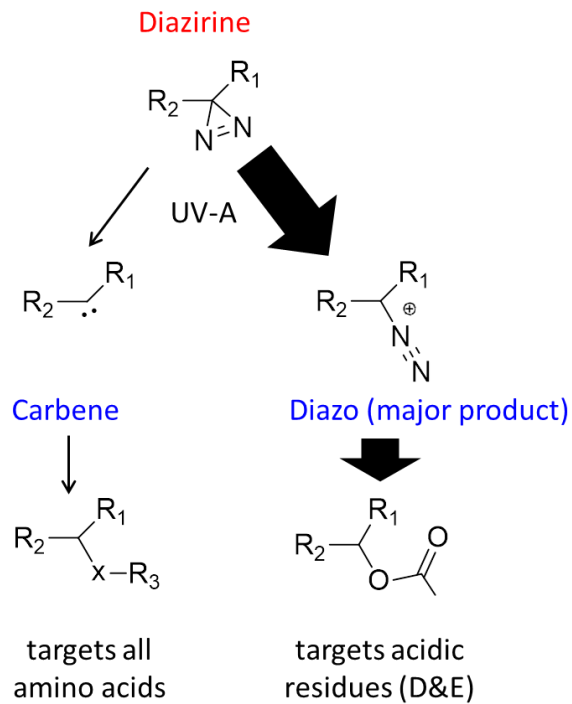


Figure 24. Cross-linking chemistry of diazirine cross-linkers. Upon UV photo-activation, there are two possible pathways for the diazirine. The major diazo pathway esterifies carboxylic acid residues (aspartic acid and glutamic acid) and is the major reaction pathway of diazirines. Targeting all 20 amino acids via a carbenes is the minor pathway.

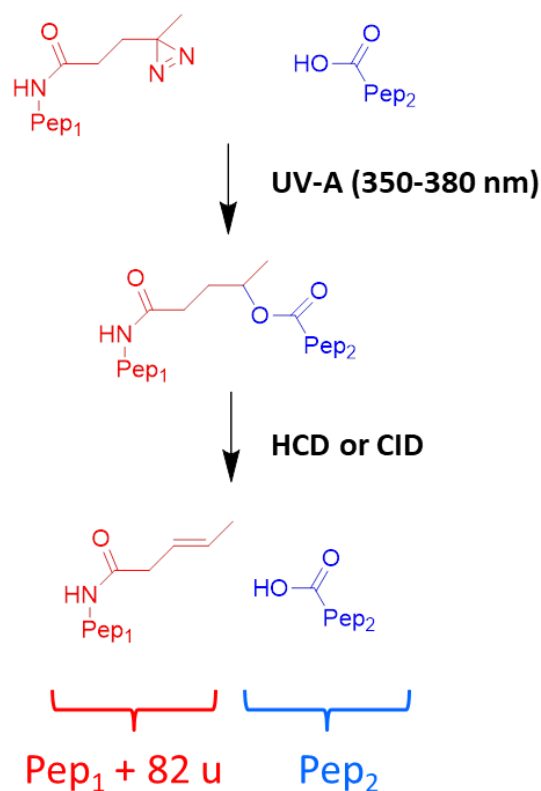


Figure 25. Dissociation of a sulfo-SDA cross-linked product via the diazo pathway. The reaction of sulfo-SDA proceeds first with the amine-reactive NHS ester, then the photo-cross-linking of the diazirine group is induced by UV-A irradiation. Upon collision-induced dissociation (HCD or CID), the cross-linked product is fragmented to yield peptide Pep₁ (+82 u modification) and Pep₂.

1.5.8. COMPetitive Pairing StatisticS

Traditional XL-MS approaches allow visualization of site-specific interactions in a heterogeneous system that contains different proteins. For cross-links identified in the same protein, it has been proven to be challenging to differentiate intra-protein and inter-protein crosslinks. For a homodimeric protein, cross-links within the same monomer are known as intra-protein cross-links, while cross-links between two different monomers are known as inter-protein cross-links (Figure 26).

S100 β forms a homodimer under native conditions. In order to differentiate inter- and intra- protein cross-links, software that differentiates the two types of cross-links have to be used. Previously, Novak et al has developed the Linx software^[197]. Here, an integrative method with the COMPetitive Pairing StatisticS (COMPASS)^[198] that had recently been developed by Dr. Claudio Iacobucci was applied to quantitate the two different types of cross-links (Figure 27). To that, an isotopically labeled version of S100 β (¹⁵N) was purified, and XL-MS was performed between the “light” (¹⁴N) and the “heavy” (¹⁵N) forms of S100 β . The COMPASS approach was derived from the observation of the mixed isotopic species in the mass spectra. Inter-protein cross-links are represented by the ¹⁴N/¹⁵N and the ¹⁵N/¹⁴N signals, while the intra-protein cross-links originate from ¹⁴N/¹⁴N and ¹⁵N/¹⁵N peaks. From the signal

intensities, it is possible to differentiate the two types of cross-links, quantitate the abundance, and draw conclusions on the intra-/inter-protein character of a specific cross-link.

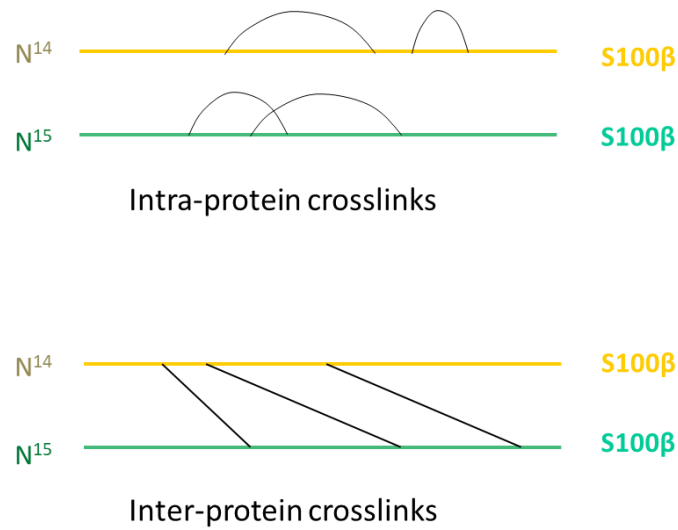
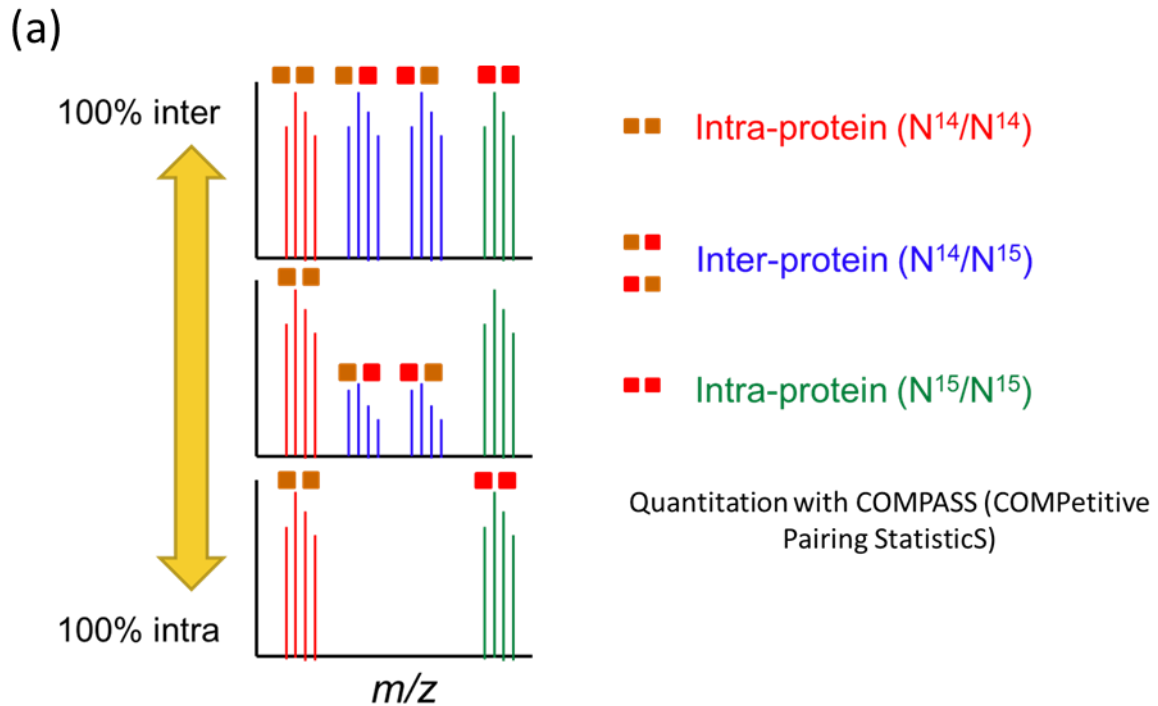


Figure 26. Schematic representation of intra/inter-protein cross-links. Two different monomers of S100β are shown in yellow and green. For a homodimeric protein, intra-protein cross-links connect two amino acid residues within the same monomer. Inter-protein crosslinks connect two amino acid residues between two different monomers. By introducing isotope labeling, the cross-links within a homodimer can be differentiated.

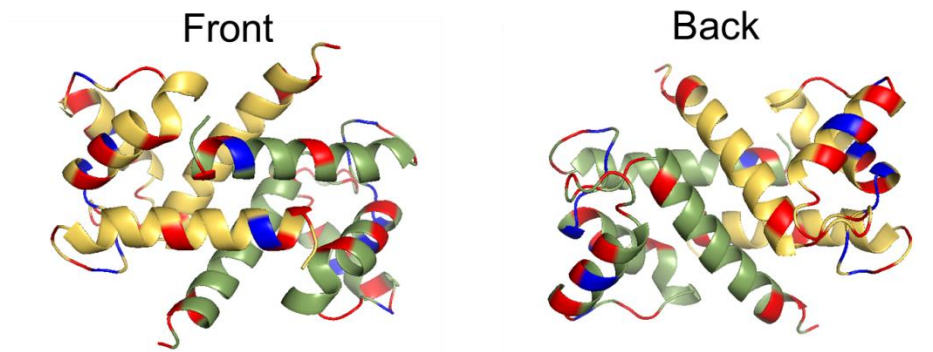


(b)

$$\text{Ratio of interprotein links} = 100 * 4 * \frac{\frac{1}{2}(LH + HL)}{(LL + LH + HL + HH)}$$

Figure 27. Schematic illustration of quantitation with the COMPASS approach. (a) Inter-protein cross- and intra-protein cross-links can be differentiated and quantitated with equation (b). Light (L) represents signals from the ^{14}N -S100 β and heavy (H) are signals from the ^{15}N -S100 β .

Due to the lack of lysine residues in S100 β , XL-MS with DSBU as amine-reactive cross-linker gives only limited structural information (Figure 28). In order to efficiently cross-link S100 β and to gain further structural insights, additional cross-linkers that also target the acidic residues of the protein, such as sulfo-SDA and EDC/sulfo-NHS, were used. S100 β is known to undergo a conformational change upon calcium binding as it belongs to the family of EF-hand proteins (Figure 5). Using COMPASS for cross-links quantitation and comparing the calcium-bound and -depleted states of S100 β provided structural insights on the p53 binding behavior of S100 β .



```

MSELEKAMVA LIDVFNHQYSG REGDKHKLKK SELKELINNE
LSHFLEEIKE QEVVDKVMET LDNDGDGECDFQEFMAFVAM
VTTACHEFFE HE

```

Figure 28. Crystal structure (PDB: 3CZT) and amino acid sequence of S100 β . Residues colored in blue represent the lysine residues, while residues colored in red represent acidic residues (aspartic acid and glutamic acid).

1.6. Surface plasmon resonance

Thermodynamic parameters, such as the dissociation constant (K_D), can be measured with surface plasmon resonance (SPR). In SPR, a non-radiative polarized light hits a sensor surface. The energy of photons from the polarized light excites the free electrons on the sensor surface and generates an electron density wave, known as surface plasmons (Figure 29a). To perform SPR, one of the interacting partners is first immobilized on a sensor surface. The interaction between one immobilized interaction partner and a ligand can be observed with the change in the intensity of the reflected light (Figure 29b), which results in the change of the SPR angle (Figure 29c)^[199].

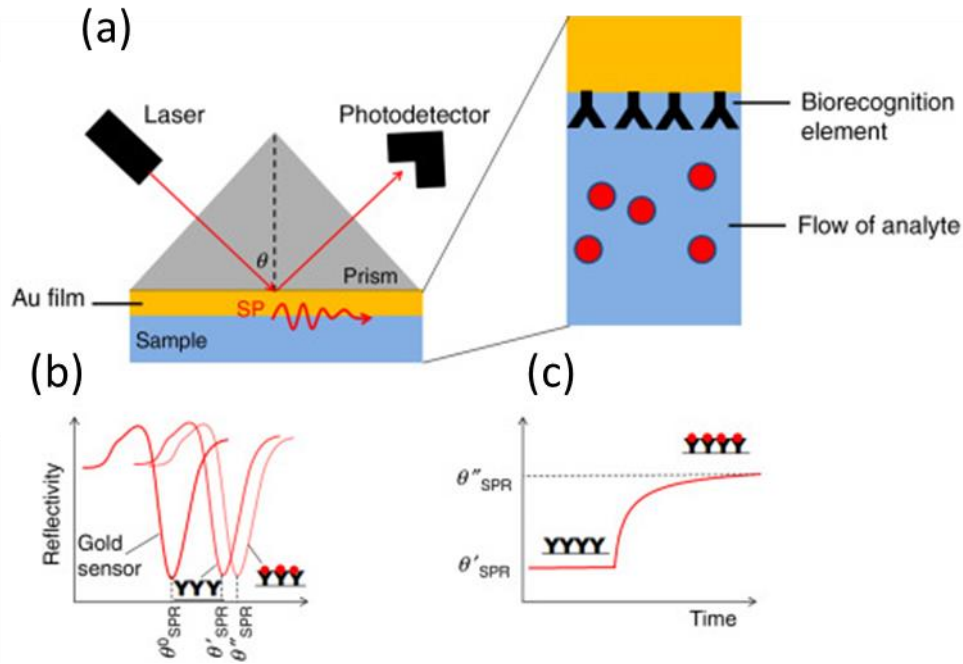


Figure 29. Illustration of SPR experimental setup. (a) A polarized laser light strikes the gold sensor surface. The electrons on the surface are excited and generate an electron charge density wave (surface plasmons). (b) The reflectivity indicates the change of intensity on the reflected light. Analyte bound to the immobilized ligand causes a change in the reflectivity. (c) An example of an SPR sensogram. Upon binding of analyte, a change in the refractive index causes a shift of the SPR angle. Figure adapted from Miyazaki et al.^[200].

Several methods of coupling chemistry can be applied, including amine, thiol, and aldehyde reactivities. In the case of immobilizing one binding partner using amine-coupling chemistry, surface is first activated with a combination of EDC/sulfo-NHS (chapter 1.5.7.2.) and lysine residues of the protein under investigation are immobilized covalently at the chip surface. After immobilization, multiple rounds of injections with the analyte are performed at increasing concentrations. The binding affinity can then be acquired based on the responses from multiple rounds of the analyte injection.

SPR measures the real-time interaction between two binding partners. The immobilized binding partner allows repeated measurements with rounds of analyte injection and surface regeneration. As SPR is a label-free and highly sensitive technique, it is a popular method to characterize the binding affinities between a protein and a small-molecular analyte (drug) or between two proteins^[201].

2. Materials and Methods

2.1. Materials

2.1.1. Chemicals

Chemicals	Manufacturer
¹⁵ N Ammonium chloride	Sigma-Aldrich
4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Carl Roth GmbH & Co. KG
Acetic acid	Carl Roth GmbH & Co. KG
Acetone	Roth
Acetonitrile (HPLC, LC/MS grade)	VWR
Acrylamide/bisacrylamide	Carl Roth GmbH & Co. KG
Ammonium acetate	Sigma-Aldrich
Ammonium bicarbonate	Sigma-Aldrich
Ammonium persulfate (APS)	Sigma-Aldrich
β-Mercaptoethanol	Merck KGaA
Calcium chloride (CaCl ₂)	Carl Roth GmbH & Co. KG
Cesium iodide (CsI)	Sigma-Aldrich
cOmplete Protease Inhibitor Cocktail (EDTA-free)	Roche
Coomassie Brilliant Blue G250	Thermo Fisher Scientific
Coomassie Brilliant Blue R250	Thermo Fisher Scientific
Dimethylsulfoxide (DMSO)	Thermo Fisher Scientific
Disodium hydrogen phosphate (Na ₂ HPO ₄)	Sigma-Aldrich
Dithiothreitol (DTT)	AppliChem GmbH
Ethanol (HPLC, LC/MS grade)	Merck KGaA
Ethanolamine	Thermo Fisher Scientific
Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA)	Carl Roth GmbH & Co. KG
Formic acid (FA)	Carl Roth GmbH & Co. KG
Glycerol	Grüssing GmbH
Hydrochloric acid (HCl)	Merck KGaA
Imidazole	Merck KGaA
Iodoacetamide (IAA)	Merck KGaA

Isopropyl β -D-1-thiogalactopyranoside (IPTG)	Carl Roth GmbH & Co. KG
Kanamycin sulfate	Carl Roth GmbH & Co. KG
Magnesium chloride ($MgCl_2$)	Merck KGaA
Methanol (HPLC, LC/MS grade)	VWR
Monopotassium dihydrogen phosphate (KH_2PO_4)	Sigma-Aldrich
N,N,N',N'-Tetramethylethylenediamin (TEMED)	VWR
PageRuler Protein Ladder (Pre-stained Plus)	Thermo Fisher Scientific
Polysorbate 20 (Tween-20)	Thermo Fisher Scientific
Potassium chloride (KCl)	Carl Roth GmbH & Co. KG
Rotiphorese 10x SDS	Carl Roth GmbH & Co. KG
Sodium dodecyl sulfate (SDS)	Carl Roth GmbH & Co. KG
Sodium acetate	Sigma-Aldrich
Sodium chloride	Carl Roth GmbH & Co. KG
Sodium hydrogencarbonate ($NaHCO_3$)	Grüssing GmbH
Sodium hydroxide (NaOH)	Carl Roth GmbH & Co. KG
N-Hydroxysulfosuccinimide (Sulfo-NHS)	Thermo Fisher Scientific
Trifluoroacetic acid (TFA)	Merck KGaA
Tris(2-carboxyethyl)phosphine (TCEP)	Carl Roth GmbH & Co. KG
Tryptone	AppliChem GmbH
Urea	AppliChem GmbH
Water (HPLC, LC/MS grade)	VWR
Yeast Extract	Carl Roth GmbH & Co. KG
Zinc chloride ($ZnCl_2$)	Sigma-Aldrich

2.1.2. Plasmids

Plasmid	Insert	Remarks	Antibiotic resistance	Source
HLT_p53FL_pET28a	wild-type tetrameric p53	N-terminal HLT (histidine-lipoyl tobacco etch virus) tag	Kanamycin	Bastian Bräuning, Weizmann Institute, Rehovot, Israel
HLT_p53_L344A_pET28a	L344A variant dimeric p53			Dr. Christian Arlt and Wiebke Schultze, MLU Halle-Wittenberg
HLT_p53_L344P_pET28a	L344P variant monomeric p53			Dr. Christian Arlt and Wiebke Schultze, MLU Halle-Wittenberg
pET28a-S100β	S100β			Dr. Christian Arlt and Maximilian Hans Clare, MLU Halle-Wittenberg

2.1.3. Bacterial strains

Strain	Genotype
<i>E. coli</i> BL21(DE3)	<i>E. coli</i> , F- <i>ompT gal dcm lon hsdSB</i> (r-Bm-B) λ(DE3 [<i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i>])
<i>E. coli</i> DH5α	<i>E. coli</i> , F- <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG</i> Φ80 <i>dlacZΔM15 Δ(lacZYA-argF) U169, hsdR17(rK- mK+)</i> , λ-

2.1.4. Media

Antibiotic	Concentration
Kanamycin	stock concentration 30 mg/mL in H ₂ O, working concentration 30 µg/mL

Medium	Composition
2 x YT	16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, pH 7.0
LB	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.2

Minimal medium	1 M phosphate buffer solution (225 mM KH ₂ PO ₄ + 775 mM K ₂ HPO ₄ , pH 7.0), 10 g/L glucose, 1 g/L ¹⁵ N NH ₄ Cl, 1 mM MgSO ₄ , 2 mL trace element solution I, 200 µL trace element solution II
SOC medium	10 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 20 mM glucose
Trace element solution I (1,000x)	11.3 g/L CaCl ₂ x 2 H ₂ O, 372.2 g/L Na ₂ EDTA, 270.3 g/L FeCl ₂ x 6H ₂ O
Trace element solution II (10,000x)	250 g/L CuSO ₄ x 5 H ₂ O, 170 g/L MnSO ₄ x H ₂ O, 287.5 g/L ZnSO ₄ x 7 H ₂ O, 238 g/L CoCl ₂ x 6 H ₂ O

2.1.5. Protein

Protein	Source
S100β Human (PRO-2312)	Prospec-Tany TechnoGene Ltd.

2.1.6. Buffers and solutions

Buffer	Composition
IMAC Binding Buffer (Cell Lysis Buffer)	50 mM Na ₂ HPO ₄ , 300 mM NaCl, 2.5 mM TCEP, 20 mM imidazole, pH 8.0
IMAC Elution Buffer	50 mM Na ₂ HPO ₄ , 300 mM NaCl, 2.5 mM TCEP, 500 mM imidazole, pH 8.0
SEC Buffer	50 mM HEPES, 300 mM NaCl, 2.5 mM TCEP, 10% (v/v) glycerol, pH 7.2
IPTG	238.31 g/mol IPTG in Milli-Q water
Zinc chloride (ZnCl ₂)	0.136 g/L ZnCl ₂ in Milli-Q water
Gel staining solution	1% (w/v) Coomassie Brilliant Blue R250 in 10% (v/v) acetic acid, 25% (v/v) methanol
Gel destaining solution	10% (v/v) acetic acid, 25% (v/v) methanol

Phosphate Buffer Saline (PBS)	8 g NaCl, 0.2 g KCl, 1.44 g Na ₂ HPO ₄ , 0.24 g KH ₂ PO ₄ in 1 L Milli-Q water, pH 7.4
Sodium bicarbonate	0.2 M sodium bicarbonate, pH 9.0
SPR running buffer	50 mM HEPES, 300 mM NaCl, 2.5 mM TCEP, 1 mM CaCl ₂ , pH 7.2
SPR regeneration buffer	10 mM NaOH, 10 mM EGTA
Peptide extraction solution	2:1 (v/v) acetonitrile and 5% (v/v) TFA

2.1.7. Cross-linking reagents

<u>Cross-linker</u>	<u>Supplier</u>
DSBU (Ureido-4,4'-dibutyric acid bis(hydroxysuccinimide) ester)	CF Plus Chemicals
EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide)	ThermoFisher Scientific
Sulfo-NHS (<i>N</i> -hydroxysulfosuccinimide)	ThermoFisher Scientific
Sulfo-SDA (sulfosuccinimidyl 4,4'-azipentanoate)	ThermoFisher Scientific

2.1.8. Enzymes

<u>Enzyme</u>	<u>Supplier</u>
Benzonase nuclease	EMD Millipore Corp.
Endoproteinase AspN	New England BioLabs Inc.
Endoproteinase GluC	New England BioLabs Inc.
ProTEV Plus	Promega
Sequencing grade modified trypsin	Promega
Thrombin	Sigma-Aldrich Chemie GmbH

2.1.9. Lab consumables

Item	Supplier
Amicoin Ultracentrifugal filters (10k, 30K)	Merck Millipore
Cuvettes (polystyrene, 1.5 to 3.0 mL)	Roth
Deepwell Plate 96, Protein LoBind (1000 µL)	Eppendorf
Deepwell Plate 96 square well (2000 µL)	VWR
Eppendorf tubes (0.5, 1,5, 2 mL)	Eppendorf
Falcon Tubes, 50 mL	Corning
Membrane Filters, 20 mm	Sartorius
Membrane Filters, 47 mm	Nylon
Needles	Terumo
Pasteur pipettes	Roth
Pipettes (2 µL ,10 µL, 200 µL, 1000 µL, 5000 µL)	Eppendorf
Precise Tris-HEPES (4-20%) gradient gels	Thermo Fisher Scientific
SPR Sensorchip, CMD-2D	Xantec
Syringes (2 mL, 5 mL, 10 mL)	Braun
Universal Indication Paper (pH 1-14)	Roth

2.1.10. Instruments

2.1.10.1. Laboratory equipment

Equipment	Model	Company
Autoclave	Systec VX-75	SYSTEC GmbH
Centrifuge	5804R	Eppendorf
Centrifuge	MICROSTAR 17R	VWR
Gel electrophoresis chamber	Mini-Protean Tetra Cell	Bio-Rad Laboratories GmbH

Heatblock thermostat		HLC
Magnetic stir bars		Framo/Heidolph
MP-SPR Navi 2000 OTSO		BioNavis
Nanodrop 2000c		ThermoFisher Scientific
pH meter	Seven Compact	Mettler Toledo
Rotors	45 Ti, JLA 9100	Beckman Coulter
See-saw rocker	SSL4	Stuart
Sonicator	Vibra-Cell 75185	VWR
Sonifier	SFX 250	Branson
SpeedVac Concentrator	Savant SPD1010	ThermoFisher Scientific
Thermo mixer C		Eppendorf
Ultracentrifuge	Avanti™ J-20 XP	Beckman Coulter™
Ultrasonication cleaner	Emmi 08ST H	EMAG
UV/Vis Spectrometer	Nanodrop 200C	Thermo Fisher Scientific
UV Curing System	Aicure UJ30	Panasonic
Vortex mixer 7-2020		neoLab
Gel imaging	Gel Doc XR	Bio-Rad Laboratories GmbH

2.1.10.2. Chromatography

ÄKTA Pure FPLC System		GE Healthcare
UltiMate 3000 Nano-HPLC System		Dionex
Consumables		Supplier
Acclaim PepMap C18		Thermo Fisher Scientific
HisTrap FF, 1 mL		GE Healthcare
HisTrap FF, 5 mL		GE Healthcare
HiLoad 26/600 Superdex 200 pg		GE Healthcare
HiLoad 16/600 Superdex 200 pg		GE Healthcare

HiLoad 16/600 Superdex 75 pg

GE Healthcare

μPAC 200 cm

Pharma Fluidics

2.1.10.3. Mass spectrometers

Instrument	Supplier
High-Mass Q-TOF II mass spectrometer	Micromass/MS Vision
Orbitrap Fusion Tribrid mass spectrometer with Nanospray Flex Nano-ESI source	Thermo Fisher Scientific
timsTOF Pro mass spectrometer	Bruker Daltonik
Easy-Spray ESI source	Thermo Fisher Scientific
Xevo TQD mass spectrometer	Waters

2.1.11. Software

Software	Function
Adobe Illustrator 24	Generation of figures
MassLynx 4.1 (Waters)	MS data acquisition and data analysis
Origin 2021 9.8	Data plotting
ProteinProspector (UCSF)	<i>In silico</i> digestion of proteins
Proteome Discoverer 2.5 (Thermo Fisher Scientific)	MS data analysis
Protparam (www.expasy.ch)	Computation of physical and chemical parameters for protein sequences
PeptideMass (www.expasy.ch)	Calculation of peptide masses after enzymatic digestion
PyMol 2.3.2	Visualization of protein structures
Snappgene 2.8.1	DNA cloning

TraceDrawer (BioNavis)	SPR data analysis
MeroX 2.0.1.7. (Michael Götze)	Analysis and visualization of XL-MS data
MP-SPR Navi Control (BioNavis)	SPR data acquisition
MP-SPR Navi Data Viewer (BioNavis)	SPR data viewing and export
Quantity One 4.6.2 (BioRad)	Gel imaging
UNICORN 5 (GE Healthcare)	ÄKTA Pure FPLC system data acquisition
xCalibur 2.2 (Thermo Fisher Scientific)	LC-MS data acquisition and analysis
xiNET (Rappsilber lab)	Visualization of XL-MS data

2.2. Methods

2.2.1. Protein expression and purification

2.2.1.1. Protein expression and purification of full-length p53

The protocols of full-length, wild-type p53 expression and purification were established by Dr. Christian Arlt^[202]. Dimeric L344A and monomeric L344P p53 variants were initially expressed and purified by Dr. Christian Arlt and Wiebke Schultze in the Sinz lab at the MLU Halle-Wittenberg^[203]. Tetrameric wild-type, dimeric L344A, and monomeric L344P contain an *N*-terminal His-lipoyl-tobacco etch virus (HLT) tag. The HLT tag contains an affinity polyhistidine tag, followed by a lipoyl domain from *Bacillus stearothermophilus* for enhanced protein solubility, and a tobacco etch virus (TEV) cleavage site that can be cleaved by ProTEV protease. The competent cells *E. coli*. BL21(DE3) cells of each p53 variant was stored as glycerol stocks (50% (v/v) glycerol) at -80 °C.

Cell culture and incubation

A 5 mL pre-culture in LB medium with kanamycin (30 µg/mL) was incubated overnight at 37 °C, 120 rpm. The pre-culture was then transferred to a 2-L flask containing 1 L of 2xYT medium and kanamycin (30 µg/mL). The culture was incubated at 37 °C, 120 rpm until the OD reached a value of 0.8. After, the culture was then incubated at 18 °C, 120 rpm until the OD reached a value of 1.2. Zinc chloride (0.1 M) and IPTG (1 mM) were added to induce the protein expression before the culture was incubated at 18 °C, 120 rpm overnight.

Cell harvest

The entire procedure was performed at 4 °C. The cells were harvested by centrifugation (30 min, 4000 x g). The IMAC binding buffer (cell lysis buffer, see Chapter 2.1.6.) was added at a ratio of 1 g cells to 5 mL buffer plus protease inhibitor cocktail (cOmplete, EDTA-free). Cell lysis was performed with a sonicator (35% amplitude, 30-sec on/off pause interval, six times). Benzonase (1 µL) was added and the cell lysate was incubated for 30 min before it was centrifuged (45 min, 35,000 x g). The soluble supernatant fraction was separated from insoluble components and cell debris. Only the soluble supernatant fraction was used for further purification steps.

Purification of HLT-tagged p53 via immobilized metal-ion affinity chromatography (IMAC)

The first purification step was performed using IMAC. IMAC was performed with the ÄKTA Pure FPLC system using a HisTrap column (HisTrap FF, column volume 1 mL). Prior to sample injection, the column was equilibrated with IMAC binding buffer (5 column volumes). Following that, the soluble supernatant fraction was injected on the column at a flow rate of 1 mL/min. After sample injection, the column was washed with IMAC binding buffer with 10% (v/v) of the IMAC elution buffer containing 50 mM imidazole (10 column volumes, 10 mL). After washing, the percentage of the IMAC elution buffer was increased to 40% (200 mM imidazole) followed by 100% (500 mM imidazole) to elute the immobilized tagged protein. All fractions were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (10% resolving gel) to analyze the efficiency of protein purification.

Purification of HLT-tagged p53 via size-exclusion chromatography (SEC)

Size-exclusion chromatography (SEC) was performed directly after IMAC purification step to remove p53 from aggregates and other impurities. SEC was performed with the ÄKTA Pure FPLC system using a Superdex 16/600 200 prep grade column (column volume 120 mL). The eluted fractions containing p53 from the IMAC were collected and injected onto the Superdex column. The separation of different species based on their hydrodynamic radii was performed at a flow rate of 1 mL/min and each fraction was checked with SDS-PAGE and native MS to verify the molecular mass of the species.

TEV protease cleavage

The tag cleavage was performed with the TEV protease (ProTEV Plus, Promega). TEV protease was added to purified HLT-tagged p53. Proteolysis was performed at 4 °C overnight. The success of tag removal was checked with SDS-PAGE (10% resolving gel).

Purification of untagged p53 via size-exclusion chromatography (SEC)

Following the overnight tag removal with TEV protease, another round of SEC was performed to separate the untagged from tagged p53. SEC was performed at a flow rate of 1 mL/min and fractions

were monitored by SDS-PAGE (10% resolving gel) and native MS. Purified p53 was aliquoted (500 µL) and stored at -20 °C for no a maximum of two months.

2.2.1.2. Protein expression and purification of S100β

The cloning, expression, and purification protocol was previously established by Maximilian Hans Clare and Dr. Christian Arlt in the Sinz lab at the MLU Halle-Wittenberg^[204]. A glycerol stock of *E. coli* BL21(DE3) competent cells that had been transformed with pET28a-S100β was stored at a -80 °C. The construct of S100β contains an *N*-terminal hexa-histidine tag followed by a thrombin cleavage site. The “light” ¹⁴N-S100β was expressed in 2xYT medium as described below. The “heavy” ¹⁵N-labeled S100β was expressed in 1 L minimal medium.

Cell culture and incubation

A 5 mL pre-culture in LB medium with kanamycin (30 µg/mL) was incubated overnight at 37 °C, 120 rpm. The pre-culture was transferred to a 2-L flask containing 1 L of 2xYT medium and kanamycin (30 µg/mL). The culture was incubated at 37 °C, 120 rpm until the OD reached a value of 1.2. IPTG (1 mM) was added to the culture to allow protein expression. Following this, the culture was incubated overnight at 37 °C, 120 rpm.

Cell harvest

The procedure was performed in the same manner as described for p53 (chapter 2.2.1.1.).

Purification of tagged S100β with immobilized metal-ion affinity chromatography (IMAC)

The procedure was performed in the same manner as described for p53 (chapter 2.2.1.1.).

Thrombin enzymatic cleavage of tagged S100β

Buffer exchange was performed with an Amicon centrifugal filter units (MWCO 10 kDa) was performed to remove the high concentration of imidazole in the IMAC fractions. The filter was equilibrated with IMAC binding buffer (5000 x g, 15 min). The buffer of the IMAC fractions containing the tagged S100β was exchanged to IMAC binding buffer (5000 x g, 15 min, 5 times). For tag removal, 1 U of thrombin was added to the tagged S100β. Proteolysis was performed at 4 °C overnight. The success of tag removal was verified by native MS due to the small difference of tagged (12 kDa) and untagged S100β (10.9 kDa).

Purification of untagged S100β via immobilized metal-ion affinity chromatography (IMAC)

A second purification step with IMAC was performed to separate the untagged S100β from tagged S100β. After overnight thrombin cleavage, the sample was loaded onto a HisTrap 1 mL FF column and

separated with an ÄKTA Pure FPLC system. The untagged S100 β eluted in the flow through fraction in IMAC binding buffer, while the tagged S100 β eluted with a higher concentration of imidazole. Only the untagged S100 β was collected for further use.

Purification of untagged S100 β via size-exclusion chromatography (SEC)

Following the IMAC step to purify the untagged S100 β , the separation of thrombin from the purified untagged protein was performed with SEC. The sample was loaded onto a Superdex 16/600 75 prep grade column (column volume 120 mL) to allow the removal of thrombin (36 kDa) from the untagged S100 β (10.9 kDa) using SEC buffer. The success of S100 β purification was verified with SDS-PAGE (15% resolving gel) and native MS. Protein samples were aliquoted (500 μ L) and stored at -20 °C for a maximum of two months.

2.2.2. Buffer exchange

2.2.2.1. Manual buffer exchange – Amicon filtration units

Prior to injection into the mass spectrometer, buffer exchange of the samples was performed to 500 mM ammonium acetate (pH 6.8). Amicon Ultra centrifugal filter units (30 kDa MWCO for p53 samples and 10 kDa MWCO for S100 β) and a microcentrifuge were used to perform the buffer exchange. The filter unit was first equilibrated with 500 mM ammonium acetate (at 4 °C, at a speed of 12.4 RPM for 10 minutes). A volume of 100 μ L of sample and 400 μ L of ammonium acetate were mixed and centrifuged (at 4 °C, at a speed of 12.4 RPM for 10 minutes). This step was repeated six times to ensure complete buffer exchange to 500 mM ammonium acetate.

2.2.2.2. Online buffer exchange using a self-packed column

The protocol used for the Online Buffer Exchange (OBE) was developed by Dr. Christian Ihling and Florian Otto in the Sinz lab at the MLU Halle-Wittenberg^[205]. For OBE, a self-packed polyacrylamide P6 column was prepared. Prior to sample injection, the loading pump and the column were equilibrated for 15 min at a flow rate of 100 μ L/min with 500 mM ammonium acetate. Samples were loaded onto an autosampler in a glass sample vial. A volume of 5 μ L was loaded onto the column and buffer was exchanged to 500 mM ammonium acetate at a flow rate of 100 μ L/min prior to injecting into the Q-TOF mass spectrometer for analysis. The UV lamp was ignited and the UV intensity at 280 nm was monitored.

2.2.3. SDS-PAGE of proteins

The result of each round of protein purification was verified with SDS-PAGE. Proteins are separated according to their apparent MWs. The larger the MW of the protein, the slower it migrates through the pores of the gel. Each SDS-PAGE contains a stacking gel and a running gel (Table 2). For optimal gel

electrophoresis, a suitable gel percentage must be selected. For p53, a gel percentage of 10% was selected and for S100 β , a gel percentage of 15% was selected to allow optimal separation.

Table 2. Composition of stacking and resolving gel for SDS-PAGE.

Components	Stacking Gel		Running Gel	
	5%	10%	15%	
0.5 M Tris-HCl, pH 6.8	1300 μ L	-	-	-
1.5 M Tris-HCl, pH 8.0	-	5000 μ L	7180 μ L	
Acrylamide/Bisacrylamide Solutions 40% (w/v)	2500 μ L	6000 μ L	7500 μ L	
Milli Q H₂O	6000 μ L	8680 μ L	7180 μ L	
10 % SDS	100 μ L	200 μ L	200 μ L	
10 % APS	50 μ L	100 μ L	100 μ L	
TEMED	20 μ L	20 μ L	20 μ L	

Prior to loading, a 1:1 ratio (v/v) of Laemmli buffer and protein sample was mixed. The sample was then heated to 95 °C for 10 minutes to allow complete denaturation, thereafter, sample was loaded onto the stacking gel. Separation was performed by applying a voltage of 200 V for 50 minutes. The staining of the gel was done for 4-10 hours at room temperature with Coomassie Brilliant Blue staining solution and then destained with destaining solution overnight. For cross-linking samples, due to the different mass species present in the sample, a commercially available gradient gel (Precise Tris-HEPES, 4-20%, Thermo Fisher Scientific) was used instead to allow better separation of species with different apparent MW.

2.2.4. ESI mass spectrometry of intact cross-linked p53-S100 β complexes

All native MS experiments were performed at 4 °C with the SEC buffer (Figure 30)^[206]. Each p53 variant (tetrameric wild-type, dimeric L344A, and monomeric L344P) was used at a final concentration of 6 μ M, S100 β was used at a final concentration of 6 μ M, and DSBU with a final concentration of 0.12 mM (20-fold molar excess of p53). 1 mM of calcium chloride was added to calcium-containing DSBU-modified samples, while for calcium-depleted DSBU-modified samples 1 mM of the chelating agent ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA) was added. Quenching of cross-linking reactions was performed by adding 20 mM of ammonium bicarbonate after 60 minutes.

Buffer exchange to 500 mM ammonium acetate was conducted immediately after the addition of ammonium bicarbonate as described in section 2.2.2.

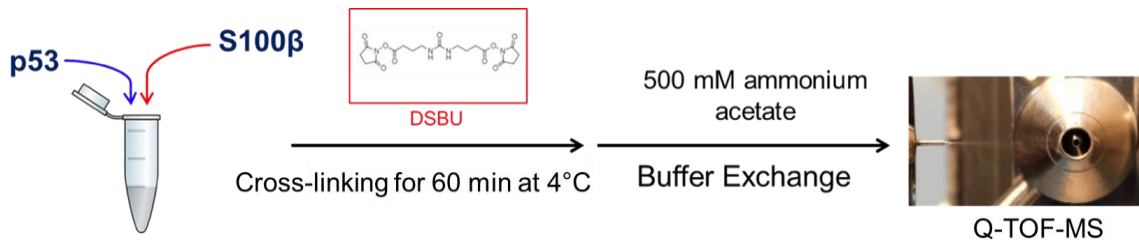


Figure 30. Protocol for ESI-MS of intact cross-linked protein complexes. Cross-linking with DSBU was performed and the buffer exchange with the sample to 500 mM ammonium acetate was performed. Then, sample was injected into the mass spectrometer via a capillary.

2.2.5. Cross-linking mass spectrometry

The protocol of XL-MS was adapted from Iacobucci et al (Figure 31)^[186]. Cross-linking reaction using different cross-linkers (chapter 2.2.5.1.) was followed by enzymatic digestion (chapter 2.2.6.2.), and the digested peptides were injected into the mass spectrometer with the application of a LC/MS gradient (chapter 2.2.7.3.). MS data was evaluated with the MeroX software (chapter 2.2.7.4.) and a summary of cross-links of the protein-protein interactions was generated.

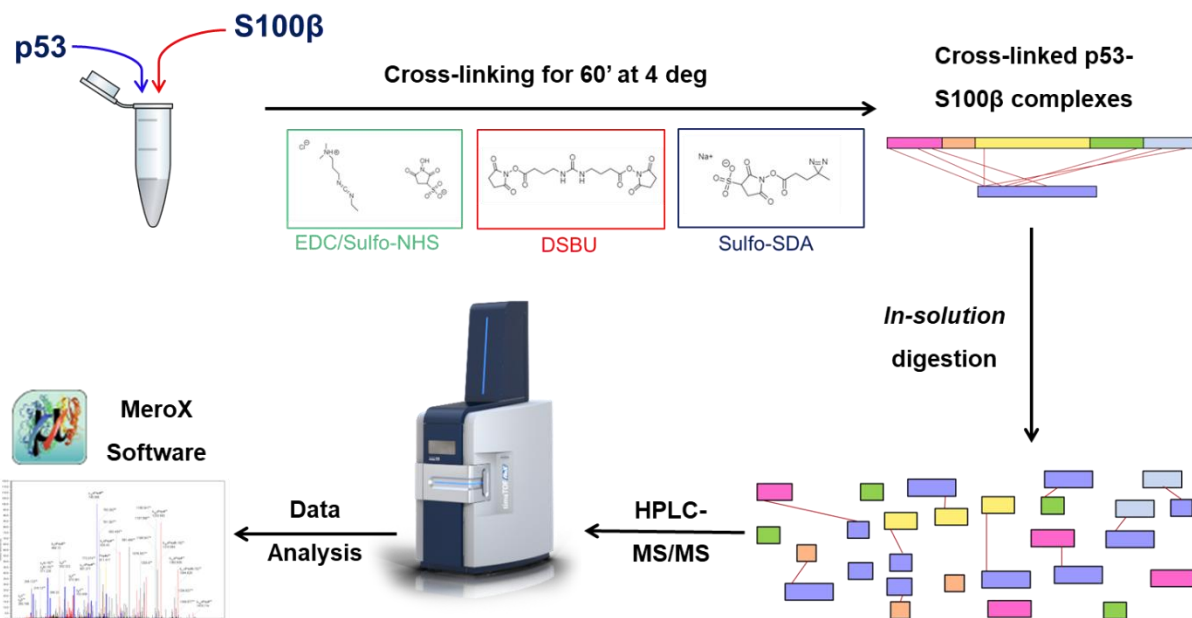


Figure 31. Cross-linking mass spectrometry (XL-MS) experimental workflow. Cross-linking reaction was performed with three different cross-linkers, DSBU, sulfo-SDA, or EDC/sulfo-NHS in independent experiments. After the cross-linking reaction, in-solution digestion was performed with AspN and

Trypsin. Digested peptides separated by nano-HPLC and analyzed with the timsTOF Pro mass spectrometer. Cross-linked products were assigned with the MeroX software.

2.2.5.1. Cross-linking reactions

All proteins were purified and stored in the SEC buffer and all cross-linking experiments were performed in the same buffer. Three replicates were performed for each cross-linking experiment and were analyzed individually with the MeroX software (settings of software are described in chapter 2.2.7.4).

DSBU cross-linking

In a 1.5 mL protein LoBind tube 6 μ M of S100 β was mixed with 6 μ M of each variant of p53. To compare the calcium-bound and the calcium-depleted state, 1 mM of calcium chloride or 1 mM of the chelating agent EGTA was added accordingly. A 100-fold molar excess (0.6 mM) of DSBU was added and the reaction was incubated at 4 °C for one hour. To stop the cross-linking reaction, quenching was performed by adding 20 mM ammonium bicarbonate to raise the pH and to deactivate the amine-coupling chemistry.

Sulfo-SDA cross-linking

In a 1.5 mL protein LoBind tube 6 μ M of S100 β was mixed with 6 μ M of each variant of p53. To compare the calcium-bound and the calcium-depleted state, 1 mM of calcium chloride or 1 mM of the chelating agent EGTA was added accordingly. A 100-fold molar excess (0.6 mM) of sulfo-SDA was added into the protein mixture and incubated at 4 °C for 2 hours. After the incubation, the sample was irradiated for 15 seconds at 365 nm with a UV laser and subsequently quenched with 20 mM of ammonium bicarbonate.

EDC/sulfo-NHS cross-linking

In a 1.5 mL protein LoBind tube 6 μ M of S100 β was mixed with 6 μ M of each variant of p53. To compare the calcium-bound and the calcium-depleted state, 1 mM of calcium chloride or 1 mM of the chelating agent EGTA was added accordingly. A 500-fold molar excess (3 mM, dissolved in DMSO) of EDC and a 500-fold molar excess (3 mM, dissolved in SEC Buffer) of sulfo-NHS (3 mM) were added. Sample was incubated for 2 hours at 4 °C and was subsequently quenched with 20 mM ammonium bicarbonate.

2.2.6. Enzymatic digestion

To investigate protein-protein interactions with bottom-up MS, proteins have to be enzymatically digested to peptides. Prior to performing enzymatic digestion of protein, optimal combination of proteases was determined with software tools PeptideMass and MS-Digest. Protein samples can be enzymatically digested either in-gel or in-solution as described below^[186].

2.2.6.1. In-gel proteolysis for protein identification

The gels were destained with the destaining solution for two hours and changed to water for 30 minutes prior. Hereon, gloves and over-sleeves were always worn and all steps were performed in a laminar flow hood to avoid contamination of keratin. To maximize the surface area of contact, gel bands that correspond to the MW of interest were excised with a scalpel into 1 mm³ cubes and transferred to a protein LoBind tube. Gel pieces were dehydrated with 500 µL of acetonitrile and incubated for 10 minutes at room temperature. To ensure dehydration, the tube was inverted and mixed for a few times, the gel pieces visibly shrank and lost colour. Acetonitrile was removed from the tube and the gel pieces were left to air dry for 1 minute. To reduce the cysteine residues of the proteins, 50 µL of 10 mM DTT was added and incubated at 55 °C for 30 minutes. After, 500 µL of acetonitrile was added into the gel pieces and incubated for 10 minutes to dehydrate. It was then removed prior to the next step. For alkylation of the free and reduced cysteine residues in the protein, 50 µL of 55 mM iodoacetamide solution was added into the tube and left to incubate for 20 minutes in the dark. After, 500 µL of acetonitrile was added to dehydrate the gel pieces and incubated for 10 minutes. The solution was subsequently discarded. A volume of 100 µL of a mixture a 1:1 (v/v) ratio of ammonium bicarbonate and acetonitrile was added to the gel pieces to ensure complete destaining of the gel pieces followed by an addition of 500 µL of acetonitrile to dehydrate them.

To ensure maximum cleavage of proteins and complete sequence coverage, a combination of two proteases were used to perform enzymatic digestion. AspN/trypsin and GluC/trypsin were applied to the protein with a final molar ratio of 1:50 (enzyme to protein). The combination of the endoproteinase AspN or GluC was added to cover up the excised gel pieces at 37 °C for 16 hours in the incubator. After, trypsin was added and incubated for 4 hours at 37 °C. Thereafter, 100 µL of peptide extraction solution (2:1 (v/v) acetonitrile and 5% TFA) was added into the gel pieces and incubated at 37 °C for 15 minutes to extract the peptide from the gel pieces. This step was repeated once more to ensure complete extraction. Prior to injecting samples for LC-MS/MS analyses, removal of acetonitrile in the extracted solution was performed with vacuum centrifugation.

2.2.6.2. In-solution proteolysis for cross-linked protein samples

To avoid contamination of keratin, gloves and over-sleeves were worn at all times to perform all the steps of in-solution digestion in a laminar flow hood. Protein samples were dried down with the vacuum centrifuge until only 5 µL of residual volume remains. 25 µL of 400 mM ammonium bicarbonate and 8 M of urea were added to the protein and the sample was sonicated for 5 minutes. 5 µL of DTT (45 mM) was added and heated at 55 °C for 30 minutes. The sample was cooled down to room temperature for 10 minutes, and then 5 µL of iodoacetamide (100 mM) was added and incubated in the dark at room temperature for 30 minutes. To quench the residual iodoacetamide, 10 µL of DTT

(45 mM) was added and incubated at room temperature for 10 minutes. To dilute the total concentration of urea to 1 M, 160 μ L of Milli Q water was added.

Protein samples were digested with a different combination of enzymes (AspN/trypsin and GluC/trypsin). All enzymes were added with a molar ratio of 1:50 enzyme to protein. AspN or GluC was added and incubated at 37 °C for 16 hours and trypsin was added for 4 hours. To quench the enzymatic digestion, 10 μ L of 10% TFA was added, and the pH value of the sample was tested to ensure the complete quench of the reaction.

2.2.7. Mass spectrometry

2.2.7.1. Electrospray ionization mass spectrometry

ESI-MS experiments were performed with a High-Mass Q-TOF II instrument. The source pressure of the instrument was set at 0.1 mbar and the collisional cell pressure was set at 1.0×10^{-2} mbar. For MS experiments, the capillary voltage was set at 1.2 – 1.4 kV (for manual buffer-exchanged samples) or 2.0 kV (for OBE), sample cone voltage was set to 60 – 80 V, the collision voltage was set to 30 – 100 V, and the transmission range off-set at a mass range from m/z 1000 – 8000. For CID-MS/MS experiments the collisional energy was set at 120-140 V and the collisional cell pressure was increased to 1.2×10^{-2} mbar.

2.2.7.2. Nano-HPLC-MS/MS-Orbitrap mass spectrometry

An UltiMate 3000 RSLC nano-HPLC system (Thermo Fisher Scientific) was coupled to the Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). The peptides after enzymatic digestion were first injected into a 5 mm precolumn (precolumn Acclaim Pepmap 100, 300 μ m x 5 mm, 5 μ m, 100 Å, Thermo Fisher Scientific) and rinsed with 0.1% (v/v) trifluoroacetic acid (TFA) for 15 minutes. After rinsing, samples were then transported to a separating column (C18) and rinsed with Solution A (0.1% formic acid (v/v) in water) at with a flow rate of 0.6 μ L/min. After that, a linear gradient was applied to the separating column. During the first 10 min, the concentration of Solution B (100% acetonitrile with 0.08% formic acid in water) was increased from 0% to 5.5%. For the next 35 min, an increase of the concentration of solution B from 5.5% to 40% was applied. After, a 5-minute gradient where the concentration of solution B was increased from 40% to 100% was applied. After the linear gradient, a washing step was performed. The separating column was subsequently rinsed with 100% Solution B for 5 minutes to prepare for the next sample injection.

For the orbitrap settings, a full-scan mass spectrum with m/z of 375-1000 for precursor ions and a mass resolution of $R=120000$ were applied. At full-scan, precursor ions with a signal intensity higher than 3.0×10^4 were subjected to higher-energy dissociation (HCD) with a normalized collision energy (NCE) set at 27%. The fragments were detected with the Orbitrap mass analyzer set at a resolution of

R=15000. For precursor ions with a signal intensity greater than 5.0×10^3 , fragmentation with a normalized collision energy (NCE) of 35% with collision induced dissociation (CID) was performed.

2.2.7.3. Nano-HPLC-MS/MS-trapped ion mobility mass spectrometry

After enzymatic digestion, the samples were injected into an UltiMate 3000 RSLC nano-HPLC system (Thermo Fisher Scientific), which was coupled to a timsTOF Pro mass spectrometer with a CaptiveSpray source (Bruker Daltonik). A C18 reversed phase column (precolumn Acclaim Pepmap 100, 300 μm x 5 mm, 5 μm , 100 Å, Thermo Fisher Scientific) was used to trap the peptides and a μPAC column (50 cm, PharmaFluidics) was used for peptides separation with a 90-minute linear gradient from 3% (v/v) to 40% (v/v) acetonitrile. During the first 15 minutes of elution, a flow gradient of 15 minutes was employed from 900 nL/min to 600 nL/min and then held constant at 600 nL/min. After the 90-minute gradient was completed, a washing step was performed. The column was rinsed with a 5-minute gradient of 40% (v/v) acetonitrile to 85% (v/v), then followed by a 5-minute gradient of 85% (v/v) acetonitrile to 3% (v/v).

For mass spectrometric data acquisition, Parallel Accumulation Serial Fragmentation (PASEF) method for standard proteomics was used. The mobility-dependent collision energy ramping was set to 95 eV at an inversed reduced mobility ($1/k_0$) of 1.6 $\text{V} \cdot \text{s}/\text{cm}^2$ and 23 eV at 0.73 $\text{V} \cdot \text{s}/\text{cm}^2$. The value of target intensity for each individual PASEF precursor was set to 20000. A scan range between 0.6 and 1.6 $\text{V} \cdot \text{s}/\text{cm}^2$ with a ramp time of 166 ms was set. A total number of 14 PASEF MS/MS scans were triggered per cycle of 2.57 seconds, where a maximum number of seven precursors per mobilogram was selected. The m/z range for precursor ions was set from 100 and 1700 and the charge states from +2 to +8 were selected for fragmentation. The active exclusion was enabled for 0.4 minutes with a mass width of 0.015 Th and a calibration run of the PASEF +2 to +8 (charge states) method was performed before injection of samples.

2.2.7.4. Data analysis with MeroX

Analysis and identification of cross-links was performed with MeroX 2.0.1.7^[196]. The setting for proteolytic cleavage for AspN/trypsin digested samples was C-terminal of lysine and arginine residues and N-terminal of aspartic acid and glutamic acid; for GluC/trypsin digested samples, the setting was C-terminal of lysine, arginine, aspartic acid, and glutamic acid. Number of missed cleavages was set to three. The length of peptide was set from five to thirty amino acid residues. For fixed modification, alkylation of cysteine residues modified by iodoacetamide was considered; while for variable modifications, oxidation of methionine residues was taken into consideration. For samples cross-linked with DSBU, cross-linked positions included lysine, serine, threonine, tyrosine, and the N-terminus. For sulfo-SDA, lysine, aspartic acid, glutamic acid, N-terminus, and the C-terminus were included for the cross-linking sites. For EDC/sulfo-NHS cross-linking, lysine, aspartic acid, glutamic acid, N-terminus, and

C-terminus were included. In the Mass Comparison tab, precursor precision (MS1) was set at 20 ppm and the fragment ion precision (MS2) was set at 30 ppm. The signal-to-noise ratio was set at greater than 1.5. For analysis mode, quadratic mode was used to search for samples cross-linked with EDC/sulfo-NHS, and RISEUP Mode was used for samples cross-linked with sulfo-SDA and DSBU. For the Score & FDR setting, the score cut-off was set at 50, false discovery rate (FDR) cut-off was set at 1% , and the intensity cut-off of precursor was set at 10%. Mass recalibration was performed for all searches.

2.2.8. Data analysis

2.2.8.1. Data analysis with MeroX

For identification and quantitation of intra- and inter-protein cross-links, an integrative method, COMPetitive Pairing StatisticS (COMPASS) was performed (Figure 32).

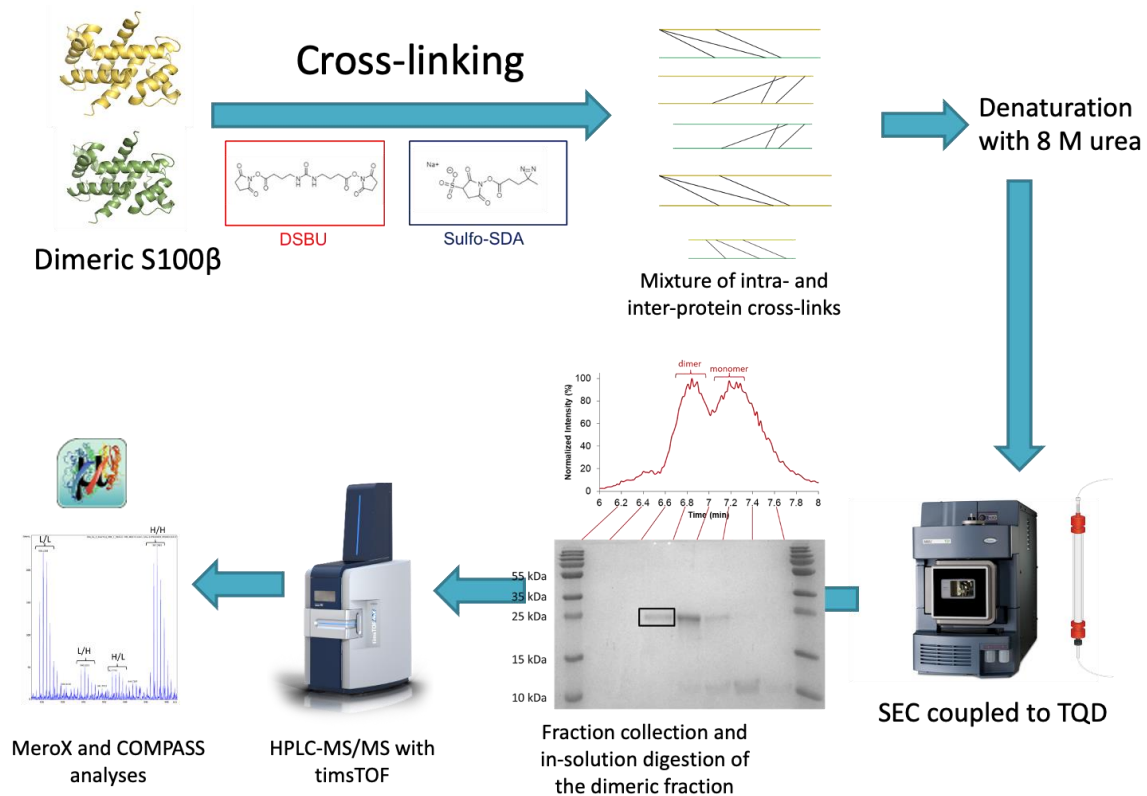


Figure 32. Protocol of SEC enrichment and cross-link quantitation with COMPASS. Cross-linking reaction was performed with ^{14}N and ^{15}N S100 β with three different cross-linkers, DSBU, sulfo-SDA, and EDC/sulfo-NHS to target different amino acid residues. Then, sample was denatured with urea and loaded on the autosampler. With SEC, cross-linked dimeric S100 β was separated from monomeric S100 β . Only the dimeric S100 β were digested and measured with timsTOF Pro instrument. MS result files were analyzed with MeroX and quantitation was performed with the COMPASS protocol.

Cross-linking of ¹⁴N and ¹⁵N S100β

For the cross-linking reaction of S100β, 20 μM of ¹⁴N S100β and 20 μM of ¹⁵N S100β were mixed in a 1.5-mL protein LoBind tube. To compare the calcium-bound and the calcium-depleted state, 1 mM of calcium chloride or 1 mM of the chelating agent EGTA was added accordingly. The experimental conditions of cross-linking with the addition of DSBU and sulfo-SDA followed the experimental conditions described in section 2.2.5.1. Three technical replicates of both the calcium-loaded and the calcium-depleted states were performed to allow a more accurate quantitation of cross-links.

Denaturation and resuspension of sample

The solvent of the cross-linked S100β samples were first evaporated and removed in a vacuum centrifuge until a residual volume of 5 μL remained. To denature and to re-suspend the cross-linked samples, 30 μL of 8 M urea in 400 mM ammonium bicarbonate was added.

Enrichment with TQD mass spectrometer

An enrichment step with SEC was performed to separate the cross-linked dimeric and monomeric S100β. The samples were loaded into a glass vial and placed in an autosampler to be injected onto a SEC column (ACQUITY UPLC Protein BEH SEC column, 200Å, 1.7 μm, 4.6 mm x 300 mm, 10-500 kDa, Waters). The elution was performed with a 15-min elution gradient with 0.2% formic acid in water with a flow rate of 300 μL/min. The outlet of the SEC column was coupled to a XEVO TQD mass spectrometer for direct analysis of the eluants from the SEC column. For MS analysis, the following settings of the XEVO TQD were applied. The capillary voltage was set to 2 kV, sample cone voltage set to 45 V, desolvation temperature set to 400 °C. MS data was acquired in full scan mode with a *m/z* from 500-2000. Two successive rounds of sample injections were performed. First, 10 μL of injected samples were coupled to the TQD mass spectrometer to acquire the chromatogram and to check the mass spectra of the injected samples. For the second round of the 10 μL injection, the size exclusion column was uncoupled from the TQD mass spectrometer for fraction collections. Elution from 6.2 – 7.8 min was collected where each fraction accounted for 0.2 min (12 sec). In order to ensure accurate analyses, only fractions that contained the cross-linked dimeric fractions verified by SDS-PAGE were digested and analyzed by injecting into timsTOF for measurement (see section 2.2.7.3).

Sample analysis with MeroX and COMPASS

The result files acquired from the timsTOF Pro mass spectrometer were first analyzed with MeroX to pick out identified inter-protein cross-links (see section 2.2.7.4). The corresponding mass spectra were then extracted from the result file according to the retention time and quantified with the COMPASS approach (see section 2.2.7.5).

2.2.9. Surface plasmon resonance

SPR experiments were performed with MP-SPR Navi 200 OTSO system.

2.2.9.1. Buffer and sample preparation

All buffers used for SPR experiments were degassed and filtered for thirty minutes prior to use. For immobilization, a coupling buffer consisting of 10 mM sodium acetate was prepared. The pH value of the coupling buffer was adjusted with hydrochloric acid (HCl) and three different pH values were prepared to perform pH scouting and to find the optimal pH value of ligand immobilization. Buffer exchange was performed with all three variants of the p53 analytes. Prior to analyte injection, SEC with ÄKTA was performed to ensure complete buffer exchange.

2.2.9.2. Protein immobilization

A pH scouting step and a pre-concentration test was performed in order to find the optimal pH for ligand immobilization. Prior to sensor surface activation, the optimal pH value was determined via the surface electrostatic interaction. A volume of 300 μL of 5 μM S100 β was injected into the instrument and then subsequently removed with 1 M NaOH for another round of injection. The fastest and the highest response were observed at a pH value of 3.9 and S100 β was immobilized on the sensor surface with sodium acetate buffer at the given condition (0.5 M, pH = 3.9).

To activate the sensor surface of the CMDP sensor chip, a mixture of 0.1 M EDC/0.2 M sulfo-NHS was injected three times consecutively. After, S100 β (5 μM) was injected three times to ensure maximum immobilization. The sensor surface was then deactivated with 1 M ethanolamine in order to quench all the unreacted NHS-esters. The sensor surface was rinsed with the SPR running buffer for ten minutes at a flow rate of 100 $\mu\text{L}/\text{min}$ prior to analyte injection.

2.2.9.3. Analyte injection and surface regeneration

Various concentrations of wild-type tetrameric p53 (0.5 μM , 1 μM , 3 μM , 5 μM , 7 μM), L344A dimeric variant (0.8 μM , 1 μM , 4 μM , 5 μM , 7 μM , 8 μM), and L344P monomeric variant (0.2 μM , 0.5 μM , 0.8 μM , 3 μM , 5 μM , 8 μM) were injected at a flow rate of 100 $\mu\text{L}/\text{min}$ for 120 seconds.

After each round of analyte injection, regeneration (10 mM NaOH and 10 mM EGTA) was performed to allow another round of injection of sample. The binding affinity (K_D value) was calculated with the software TraceDrawer.

3. Results

3.1. Expression and purification of p53 and S100 β

3.1.1. Expression and purification of three p53 variants

All expression, purification, and quality control steps were performed according to established protocols from the Sinz lab at MLU Halle-Wittenberg^[171]. A comprehensive and detailed description can be found in chapter 2.2.1.1.

A visual representation of the complete expression and purification workflow can be found in the following figure. Full-length tetrameric wild-type, dimeric L344A, and monomeric L344P p53 were purified using immobilized metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC). Tag removal with ProTEV protease was performed overnight and SEC was performed thereafter. For each purification step, the efficiency of impurity removal and the apparent molecular weights of the proteins were assessed by SDS-PAGE. After sonication and ultracentrifugation of the cell lysate, the supernatant was chromatographically purified. For the wild-type p53 and dimeric L344A p53, SEC was performed with the eluted fractions from the IMAC step, and multiple peaks were observed. After successful tag removal of the monomeric L344P p53 using ProTEV protease, SEC was performed to obtain the purified protein (Figure 33). To confirm the oligomeric state and activity of purified p53, native mass spectrometry (native MS) was performed (chapter 3.2).

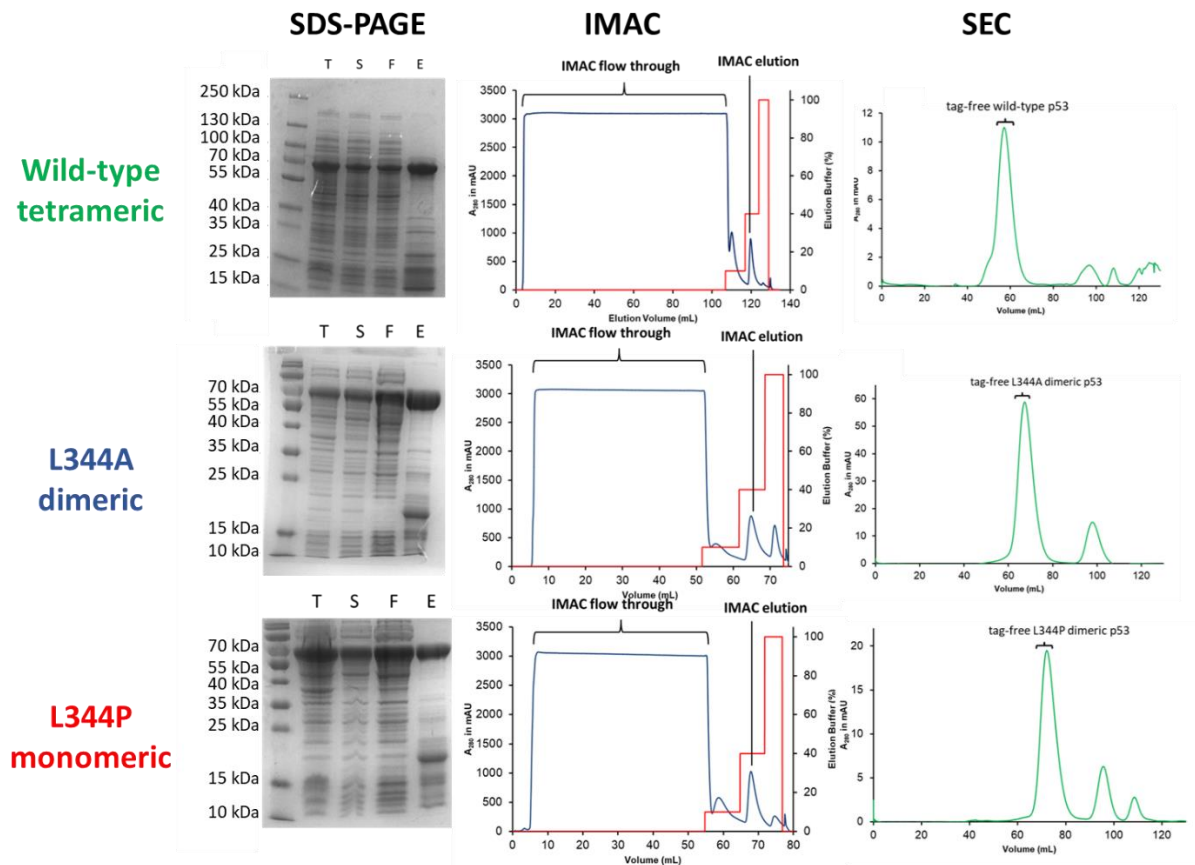


Figure 33. Summary of the expression and purification of three variants of p53. SDS-PAGE, T: total cell lysate, S: supernatant, F: flow through from IMAC, E: eluted fraction with 40% imidazole from IMAC. The chromatograms of the IMAC (2nd column) and SEC (3rd column) purification steps are shown.

3.1.2. Expression and purification of ¹⁴N and ¹⁵N S100 β

Expression and purification of the ¹⁴N and ¹⁵N S100 β followed the protocol detailed in chapter 2.2.1.2^[204]. SDS-PAGE was used to monitor every step of the protein purification. After cell lysis and ultracentrifugation, the supernatant was loaded onto the HisTrap column and IMAC was performed. Tagged S100 β contains a thrombin cleavage site and tag removal was performed with overnight proteolysis with thrombin. Afterwards, the sample was loaded onto the HisTrap column and a second IMAC step was performed to collect the untagged S100 β . To remove excess thrombin from the protein sample, SEC was performed. A summary of the protein purification of S100 β can be found in Figure 34.

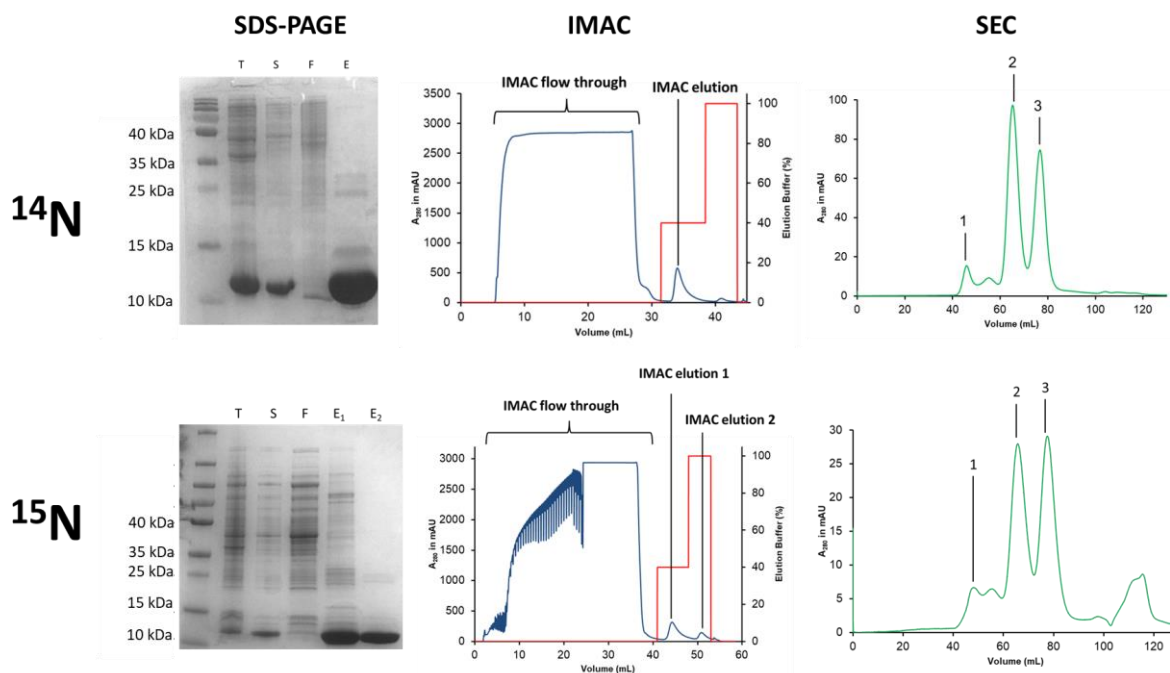


Figure 34. Summary of the expression and purification of the ^{14}N and the isotope-labelled ^{15}N S100 β . SDS-PAGE, T: total cell lysate, S: supernatant, F: flow through from IMAC, E: eluted fraction with 40% imidazole from IMAC. The chromatograms of the IMAC (2nd column) and SEC (3rd column) purification steps were shown.

3.2. Native mass spectrometry of p53 and S100 β

The protein activity and the oligomeric state of each p53 variant were verified with native MS. With the acquired m/z values, the corresponding charge states, and the MW of every species was calculated and the identity of each species was assigned in the native mass spectra.

3.2.1. Tetrameric wild-type p53

For wild-type p53, signals from m/z 3000 to 8000 were observed. The molecular mass of each identified species was calculated with the MassLynx software. The tetrameric state of the wild-type p53 was confirmed. Signals with a charge state from +25 to +30 with a calculated molecular mass of 175.4 kDa were observed, confirming the tetrameric state of the wild-type p53. Also observed in the mass spectrum were the monomeric, dimeric, and trimeric p53. For the monomeric p53, charge states from +11 to +13 with a mass range from m/z 3000 – 4000 were observed at a calculated molecular mass of 43.8 kDa. For the dimeric p53, charge states from +18 to +21, a mass range from m/z 4000 – 5000 were observed. The species with a MW of 87.7 kDa indicates the presence of the dimeric p53. A trimeric p53 species was found. The calculated molecular mass was 131.5 kDa and the corresponding signals were observed from charge states +23 to +26 at a mass range from m/z 5000 – 6000. In addition to the presence of p53 species, the chaperone protein, DnaK, was also detected from a charge state

between +14 and +16 and a mass-to-charge range from m/z 4000 to 5000. The calculated molecular mass of DnaK matches the theoretical molecular mass, 68.9 kDa (Figure 35).

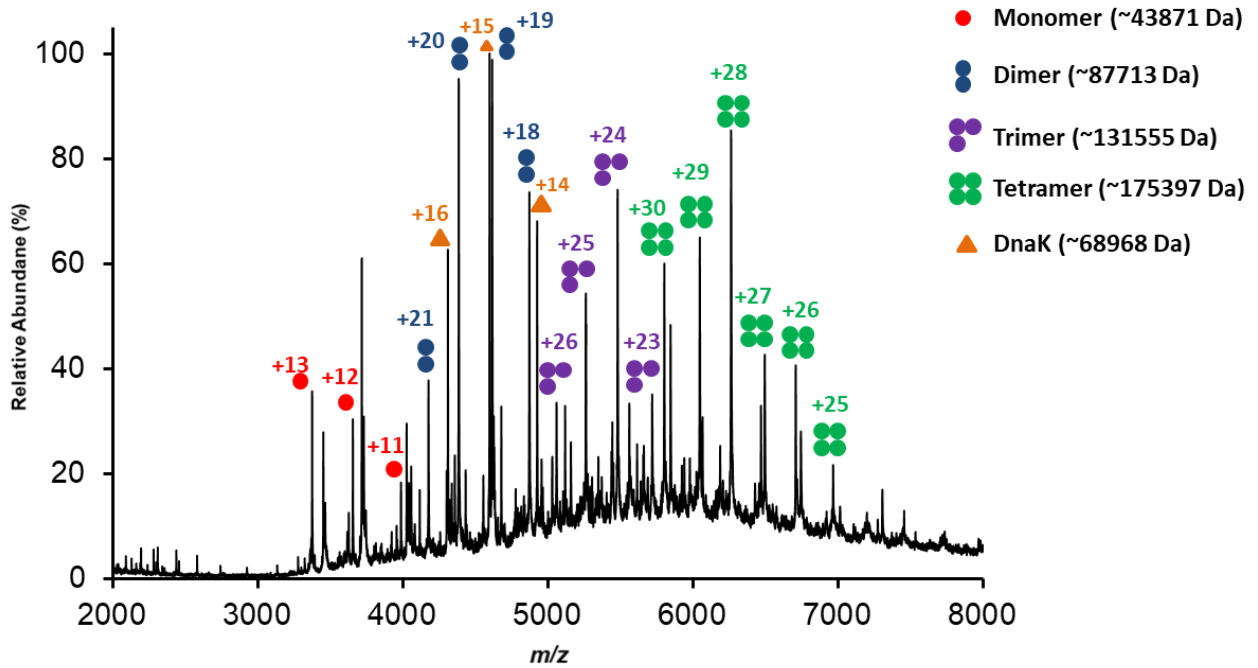


Figure 35. Native mass spectrum (positive ion mode) of the wild-type tetrameric p53. Different p53 species were observed, the monomeric (red circles, charge states from +11 to +13, MW = 43871 Da), dimeric (blue circles, charge states from +18 to +21, MW = 87713 Da), trimeric (purple circles, charge states from +23 to +26, MW = 131555 Da), and tetrameric (green circles, charge states from +25 to +30, MW = 175397 Da). The chaperone protein DnaK (orange triangles, charge states from +14 to +16, MW = 68968 Da) was also detected in the sample.

3.2.2. Dimeric L344A p53

For the L344A variant of p53, only two oligomeric states were observed, the dimeric and the monomeric p53 species. From m/z 2500 – 3600, monomeric p53 with a charge state distribution from +12 to +17 was observed. The calculated molecular mass of the monomeric p53 is 43.8 kDa, which correlates well with the theoretical molecular mass (43.7 kDa) of the monomeric species. From m/z 4000 – 5000 and charge states from +18 to +22, dimeric p53 was observed. The calculated molecular mass was 87.6 kDa and it matched well with the theoretical molecular mass of the dimeric p53 variant. Also, the chaperone protein DnaK with a charge state distribution from +14 to +17 was also observed. No signals corresponding to the tetrameric p53 were found. This confirms the single-point mutation (L344A) disrupts the tetramer formation as only the dimeric and monomeric species were observed for the L344A variant (Figure 36).

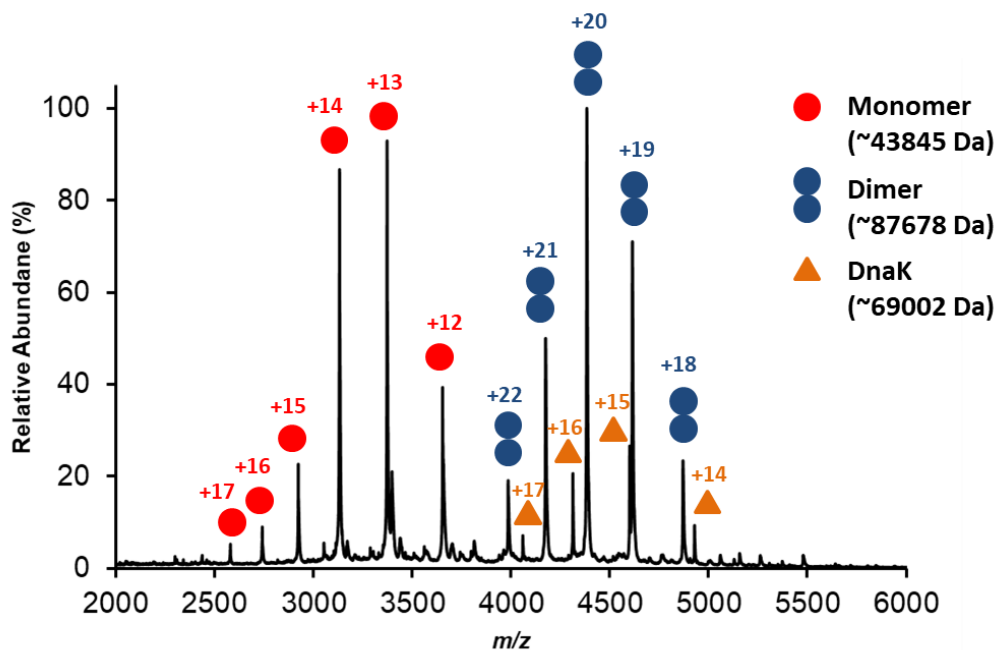


Figure 36. Native mass spectrum (positive ion mode) of the dimeric variant L344A p53. Species identified were the dimeric p53 (blue circles, charge states from +18 to +22, MW = 87678 Da), monomeric p53 (red circles, charge states from +12 to +17, MW = 43845 Da), and the chaperone protein DnaK (orange triangles, charge states from +14 to +17, MW = 69002 Da).

3.2.3. Monomeric L344P p53

The oligomeric state of the monomeric variant, L344P, was acquired by native MS. The monomeric p53 variant at a mass range from m/z 3000 – 4000 and charge states from +11 to +14 were observed. The calculated molecular mass of the species matched well with the theoretical molecular mass of monomeric p53, 43.8 kDa. Native MS confirmed the presence of only the monomeric species in the purified L344P variant, as p53 with higher oligomeric states (dimer, trimer, and tetramer) were not observed (Figure 7).

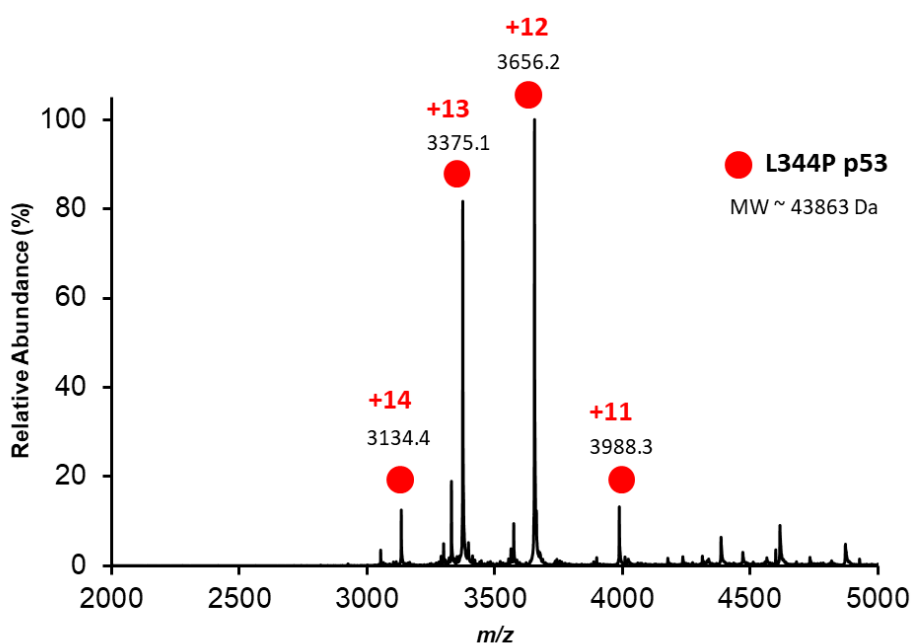


Figure 37. Native mass spectrum (positive ion mode) of the monomeric variant L344P p53. The species observed was the monomeric p53 (red circles, charges states from +11 to +14, MW = 43863 Da).

3.2.4. ^{14}N S100 β

For ^{14}N S100 β , native MS showed the presence of the monomeric S100 β with a charge state distribution of +3 to +6 with a m/z range between 1500 and 3700. The dimeric S100 β species with charge states from +7 to +9 and a m/z between 2400 and 3150 was present. The calculated molecular mass of monomeric S100 β is 11003 Da. The recombinantly produced S100 β as mentioned in chapter 2.2.1.2. possesses a thrombin cleavage site. Upon thrombin cleavage, three amino acids (Gly-Ser-His) remained at the *N*-terminus of S100 β , which is confirmed by native MS. The determined molecular mass of S100 β also indicated the presence of the artificial *N*-terminus (Figure 38).

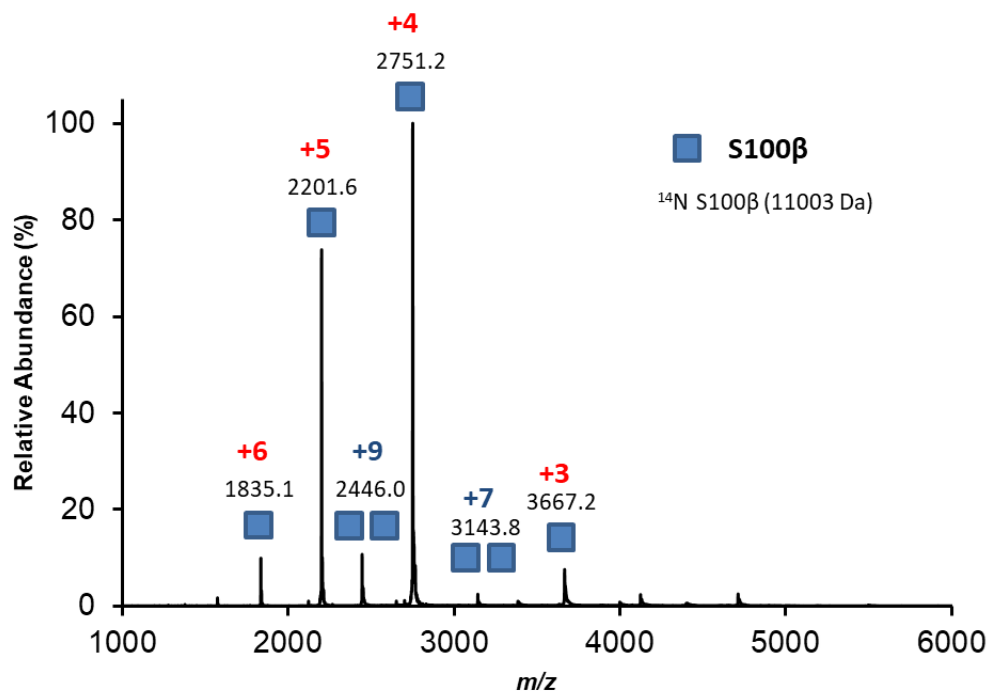


Figure 38. Native mass spectrum (positive ion mode) of ^{14}N S100 β . Monomeric S100 β is indicated by one blue square (charge states +3 to +6, MW = 11003 Da) and the dimeric S100 β (charge states +7 to +9, MW = 22002 Da) is indicated by two squares.

3.2.5. ^{15}N S100 β

For the isotopically labeled heavy ^{15}N S100 β , native MS showed the presence of the monomeric species with charge states from +3 to +6 with m/z from 1850 to 3700 and the dimeric species with charge states between +7 and +9 with m/z from 2400 to 3200. The monomeric S100 β has a calculated mass of 11127 Da. The recombinant S100 β protein contains 124 nitrogen atoms. From the mass spectrum acquired for ^{14}N S100 β , the molecular mass is 11003 Da. The mass difference between ^{14}N and ^{15}N was identical to the number of the nitrogen atoms, 124 Da. Native MS confirmed the successful production of ^{15}N S100 β . As was observed for ^{14}N S100 β , the three amino acids (Gly-Ser-His) were confirmed at the *N*-terminus (Figure 39).

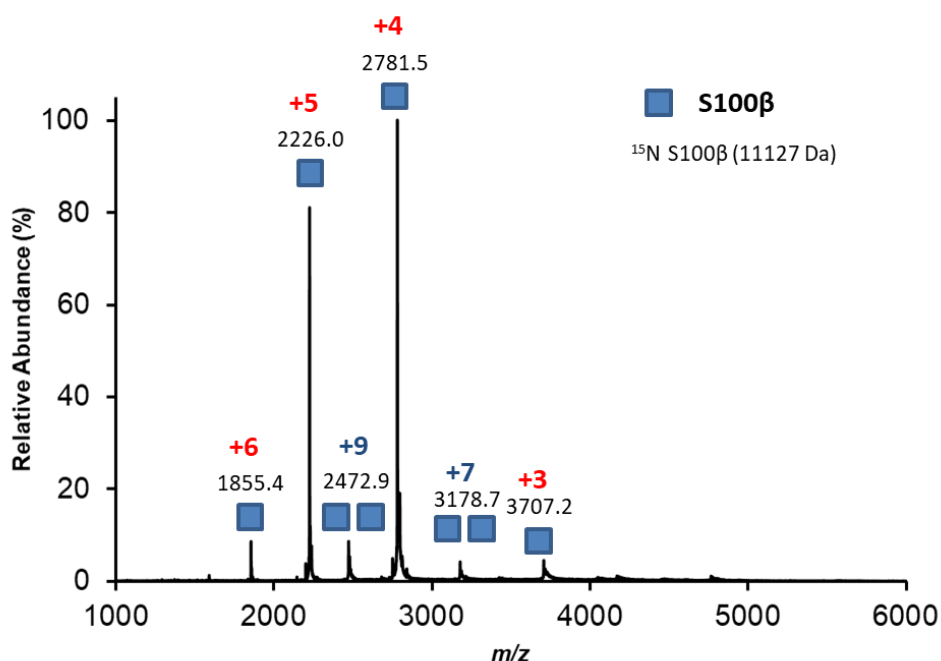


Figure 39. Native mass spectrum (positive ion mode) of ^{15}N S100 β . Monomeric S100 β is labeled with one blue square (charge states +3 to +6, MW = 11127 Da) and the dimeric S100 β (charge states +7 to +9, MW = 22244 Da) is labeled with two squares.

The charge state distribution of ^{14}N and ^{15}N were found to be in perfect agreement.

3.3. Covalent modification and mass spectrometry of p53-S100 β protein-protein interaction

Non-covalent interactions can be maintained in ESI-MS^[98]. In this study, attempts to investigate protein complex formation of p53-S100 β via native MS proved to be challenging. Native mass spectrum after buffer exchange showed inconclusive results (Supplementary Figure 5), which might be attributed to the low binding affinities^[78]. To counter this issue, a mild cross-linking reaction was performed to covalently connect the p53-S100 β complexes (chapter 2.2.4.) before ESI-MS was conducted.

3.3.1. Tetrameric wild-type p53 and S100 β

The protein-protein interaction of p53 and S100 β is calcium-dependent. Two experimental conditions, the calcium-loaded and the calcium-depleted states, were compared to verify the complex formation. The spectrum of the calcium-loaded state showed a complex formation with a charge state distribution from +29 to +32 at a mass range between m/z 6000 to 7000. The calculated molecular mass of the complex is 200 kDa, which matches the molecular mass of a p53 tetramer and an S100 β dimer (a zoomed-in version of the mass spectrum and calculation of the MW of protein complex shown in Supplementary Figure 6). In comparison, the complex was not observed in the mass spectrum of the calcium-depleted condition, confirming the calcium-dependent nature of this interaction. Other species that were present in both the calcium-loaded and calcium-depleted conditions were the monomeric (charge states of +6 and +7) and dimeric S100 β (charge states of +8 and +10), the chaperone protein DnaK from *E. coli* (charge states from +15 to +18), and the monomeric (+12 and +13), dimeric (+19 and +20), trimeric (+26 and +27), and tetrameric p53 (+23 to +31). Despite observing the presence of p53 with different oligomeric states (monomeric, dimeric, and trimeric), complex formation was only observed with the tetrameric p53 (Figure 40).

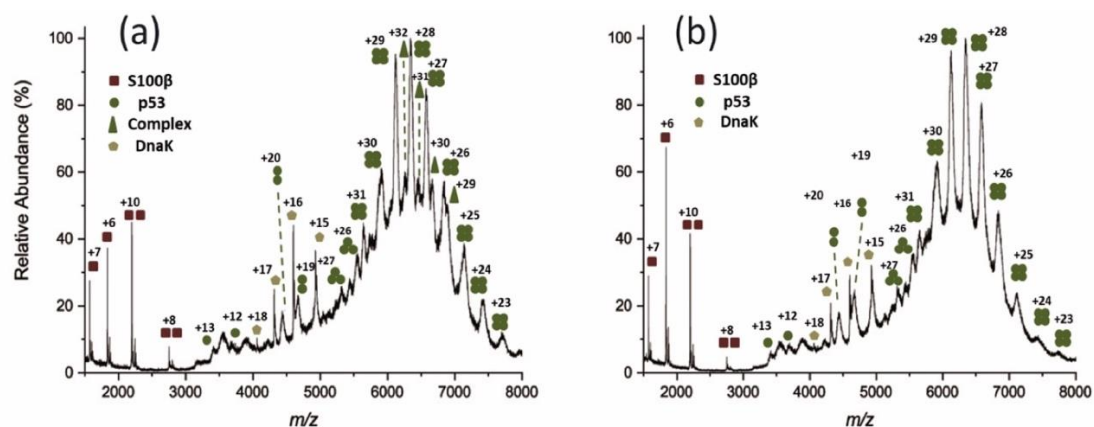


Figure 40. ESI mass spectra of intact, DSBU cross-linked S100 β and wild-type tetrameric p53 (a) with the addition of calcium (b) without calcium. S100 β monomer and dimer are indicated with brown

squares, the chaperone protein DnaK is indicated with beige pentagons, and the wild-type p53 is labeled with green circles. The p53-S100 β complex indicated with green triangles is only observed in the calcium-loaded state (a) but absent in the calcium-depleted state (b).

The composition of the complex was confirmed with CID-MS/MS. A precursor ion that indicated the formed wild-type p53-S100 β complex, m/z 6661 (+30), was selected. Upon CID-MS/MS, protein subunits and the peptide backbone fragments due to a higher collisional energy were observed. Monomeric S100 β (a charge state distribution from +5 to +7), monomeric p53 (a charge state distribution from +15 to +18), tetrameric p53 (charge states +23 and +24), and peptide backbone fragments were ejected from the p53-S100 β complex. Performing ESI-MS of an intact, mildly covalently-fixed protein complex followed by CID-MS/MS, the composition of the formed protein complex was confirmed to be composed of a p53 tetramer and an S100 β dimer (Figure 41).

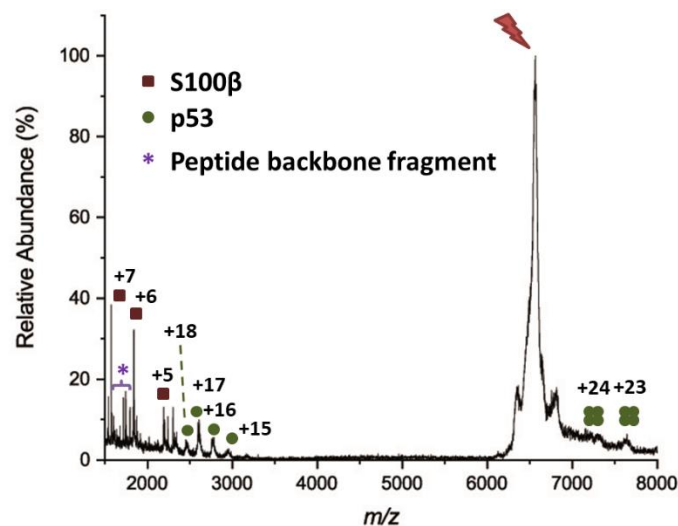


Figure 41. CID-MS/MS of the complex of wild-type tetrameric p53 and S100 β . CID-MS/MS confirmed the dissociation of the protein complex. Ion with m/z 6661 with a charge state of +31 was selected and dissociated (indicated by the red flash). Upon CID-MS/MS (collisional energy of 120 eV), monomeric S100 β , monomeric p53, and tetrameric p53 were observed.

3.3.2. Dimeric L344A p53 and S100 β

The calcium-loaded and the calcium-depleted conditions were compared for the complex formation of the dimeric L344A variant with S100 β . For the calcium-loaded state, signals with a charge state distribution from +20 to +23 at a mass range from m/z 4500 to 6000 were observed. The calculated molecular mass of the p53-S100 β complex was 111 kDa (a zoom-in region of the mass spectrum and the calculation of the protein complex is presented in Supplementary Figure 7), which corresponds to

the molecular weight of a p53 dimer and an S100 β dimer. For comparison, in the calcium-depleted state, signals of the p53-S100 β complex were not observed. Other species that were observed for both experimental conditions were the monomeric (charge states of +6 and +7) and dimeric S100 β (a charge state distribution from +8 to +10). The chaperone protein DnaK from *E. coli* (a charge state distribution from +15 to +18), and the monomeric p53 (+11 to +15) and dimeric p53 (+16 to +22) were observed. One interesting observation was that despite the monomeric p53 was observed, a complex was only formed with dimeric p53.

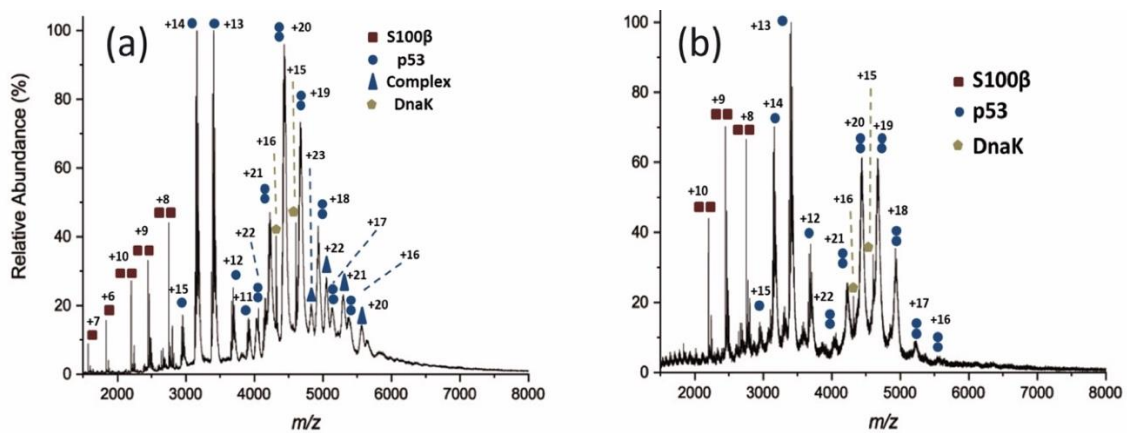


Figure 42. ESI mass spectra of DSBU intact, cross-linked S100 β and the dimeric L344A p53 variant (a) with calcium and (b) without calcium. S100 β monomer and dimer are indicated with brown squares, the chaperone protein DnaK is indicated with beige pentagons, and the monomeric and dimeric p53 are labeled with blue circles. The L344A p53-S100 β complex indicated with blue triangles is only observed in the calcium-loaded state (a), but is absent in the calcium-depleted state (b).

CID-MS/MS experiments were performed to verify the stoichiometry of the complex and the protein subunit's composition. The ion at m/z 5050 (+22) was selected and subjected to CID-MS/MS. Upon collisional activation, a monomeric S100 β was ejected from the complex. The remaining species comprised a S100 β monomer (charge state distribution from +5 to +7) and a p53 dimer (charge states +13 to +16). ESI-MS of the covalently-fixed L344A p53-S100 β complex and CID-MS/MS confirmed the composition of complex to be one p53 dimer and one S100 β dimer (Figure 43).

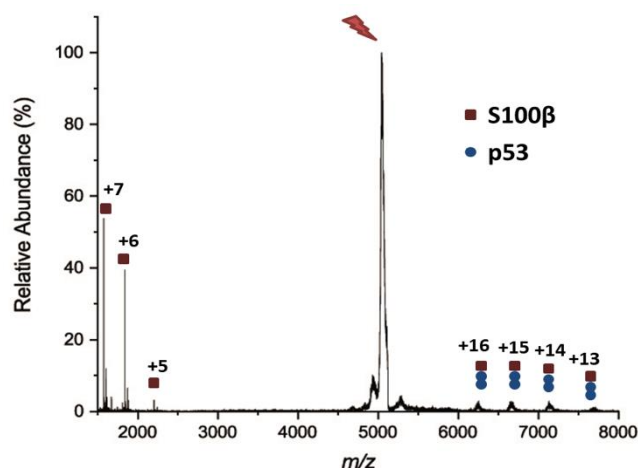


Figure 43. CID-MS/MS of the complex of the dimeric L344A p53 variant and S100β. CID-MS/MS observed the dissociation of the protein complex. Ion with m/z 5050 with a charge state of +22 was selected and dissociated as indicated by the red lightning. Upon CID-MS/MS (collisional energy of 120 eV), monomeric S100β and the complex with a dimeric p53 and monomeric S100β were observed.

3.3.3. Monomeric L344P p53 and S100β

The complex formation between monomeric p53 and S100β was compared in the calcium-loaded and the calcium-depleted states. Signals indicating the covalently-fixed protein complex between the monomeric L344P p53 and S100β were observed in the calcium-loaded condition. With a charge state distribution from +12 to +16 and a mass range of m/z 3800 to 5000, the calculated molecular mass of the complex, 66kDa, indicated the presence of a p53 monomer and S100β (a zoomed-in version of the mass spectrum and calculation of the molecular mass of protein complex shown in Supplementary Figure 8). In the absence of calcium, the complex was not observed in the calcium-depleted state. In addition to the p53 and S100β, an unidentified species with a molecular mass of 131 kDa was found. This molecular mass is similar to that of the p53 trimer. However, upon CID-MS/MS, results were inconclusive, and the unknown species remained unidentified (Figure 44).

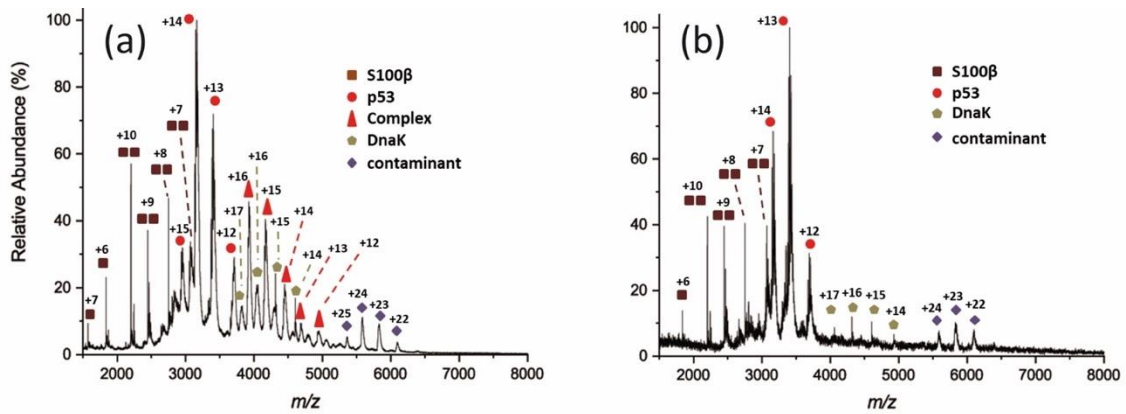


Figure 44. ESI mass spectra of DSBU intact, cross-linked S100 β and the monomeric L344P p53 variant (a) with the addition of calcium (b) without calcium. S100 β monomer and dimer were labeled with brown squares. Monomer of p53 was labeled with red circles. An unidentified contaminant was labeled with purple diamonds. The L344P p53-S100 β complex formation was indicated and labeled with red triangles and was only observed in the calcium-loaded state (a) but not the calcium-depleted state (b).

The ion indicating the formed complex with m/z 3927 (+17) was selected and CID-MS/MS was performed. Upon collisional activation, an S100 β monomer (a charge state distribution from +5 to +7) was ejected from the complex. The remaining complex (a charge state distribution from +9 to +12) comprised a p53 monomer and an S100 β monomer. ESI-MS and CID-MS/MS showed the stoichiometry and the protein subunits of the covalently-fixed monomeric L344P p53-S100 β protein complex (Figure 45).

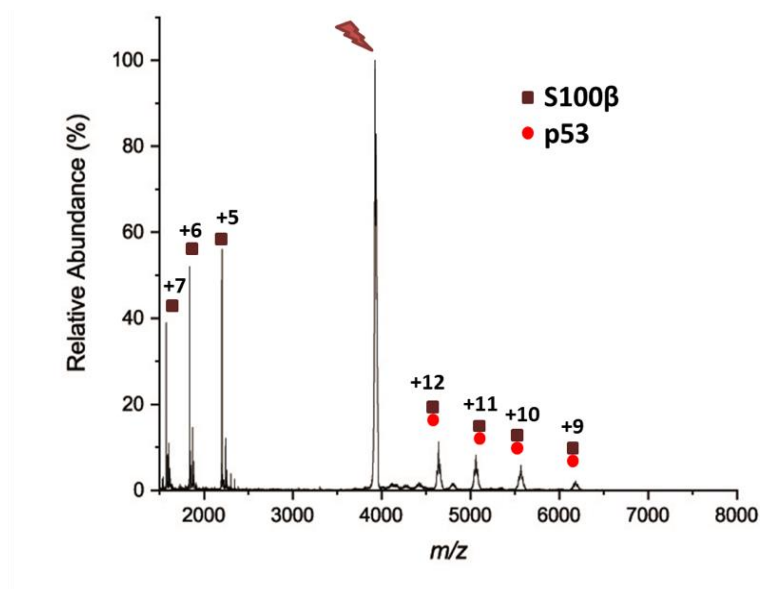


Figure 45. CID-MS/MS of the complex of the monomeric L344P p53 variant and S100 β . CID-MS/MS observed the dissociation of the protein complex. Ion with m/z 3927 with a charge state of +17 was selected and dissociated. Upon CID-MS/MS (collision energy at 120 eV), monomeric S100 β and the complex with a monomeric p53 and monomeric S100 β were observed.

3.4. Cross-linking mass spectrometry of the interaction between p53-S100 β

For XL-MS experiments, the calcium-loaded and the calcium-depleted states of S100 β were compared regarding their interaction with different p53 variants. To verify the influence of the p53 oligomeric state on the site-specific interaction with S100 β , all three oligomeric states, wild-type tetrameric, L344A dimeric, and L344P monomeric, were each investigated individually. Experiments with three cross-linkers, DSBU, sulfo-SDA, and EDC/sulfo-NHS were performed separately, and to each p53 variant, to ensure coverage of different chemical reactivities and different regions of p53 and S100 β .

3.4.1. Cross-linking with DSBU

Cross-linking experiments with DSBU between p53 and S100 β were performed and the results were verified with SDS-PAGE (Figure 46 and Supplementary Figure 1).

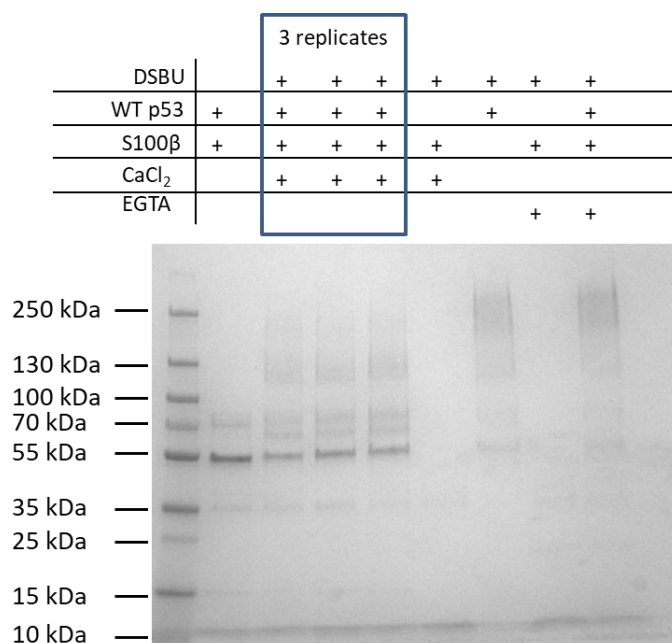


Figure 46. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100 β and wild-type p53 with DSBU. Gel band at 53 kDa indicates p53 and gel band at 10 kDa indicates S100 β . Three technical replicates were performed for the calcium-loaded environment. As a negative control, the calcium-depleted environment, with the addition of EGTA, was performed.

The calcium-depleted state of S100 β served as a negative control, and different behaviors of calcium-loaded and -depleted states were monitored by SDS-PAGE. Enzymatic proteolysis (chapter 2.2.6.2.) was performed to all cross-linked samples and digested peptides were analyzed by nano-HPLC/nano-

ESI-MS/MS. Subsequently, acquired mass spectrometric data were analyzed with MeroX software for cross-link identification.

For wild-type p53, one intermolecular cross-link was identified, for which the *N*-terminus of S100 β is cross-linked to K374 of p53. For the dimeric L344A variant, the same cross-link was found (*N*-terminus of S100 β to K374 of p53). In addition, the *N*-terminus of S100 β is cross-linked to K358 of p53. Another cross-link connecting K30 of S100 β and S379 of p53 was observed. For the monomeric L344P variant, cross-links that were observed in the dimeric variant were seen, where the *N*-terminus of S100 β is connected to K358 and K374 of p53.

A fragment ion spectrum analyzed and annotated with MeroX (Figure 47). One distinct cross-link was identified for all three p53 variants to S100 β , connecting the *N*-terminus of S100 β with a lysine at position 374 of p53 (Table 3). For DSBU, as described earlier (chapter 1.5.7.1.), the MS-cleavable doublets (+ 26 Da) were observed in the fragment ion spectrum. The XL-MS results with DSBU agree with the previously reported peptide studies^{[207][67]}, where the regulatory domain (Reg.) of p53 is the binding interface for S100 β . Here, all the p53 oligomeric states showed an identical binding behavior and the same cross-links towards S100 β .

The identified cross-links between p53 and S100 β are summarized in Table 3.

Table 3. Summary of the identified p53-S100 β cross-links with DSBU.

Wild-type tetrameric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
<u>K</u> GQSTSR	374	380	{GSHMSELEK	0	9	K374	<i>N</i> -terminus
L344A dimeric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
<u>K</u> GQSTSR	374	380	{GSHMSELEK	0	9	K374	<i>N</i> -terminus
DAQAG <u>K</u> EPGGSR	353	364	{GSHMSELEK	0	9	K358	S5
SKKGQST <u>S</u> R	372	380	H <u>K</u> LK	29	32	S379	K30
L344P monomeric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
DAQAG <u>K</u> EPGGSR	353	364	{GSHMSELEK	0	9	K358	<i>N</i> -terminus
<u>K</u> GQSTSR	374	380	{GSHMSELEK	0	9	K374	<i>N</i> -terminus

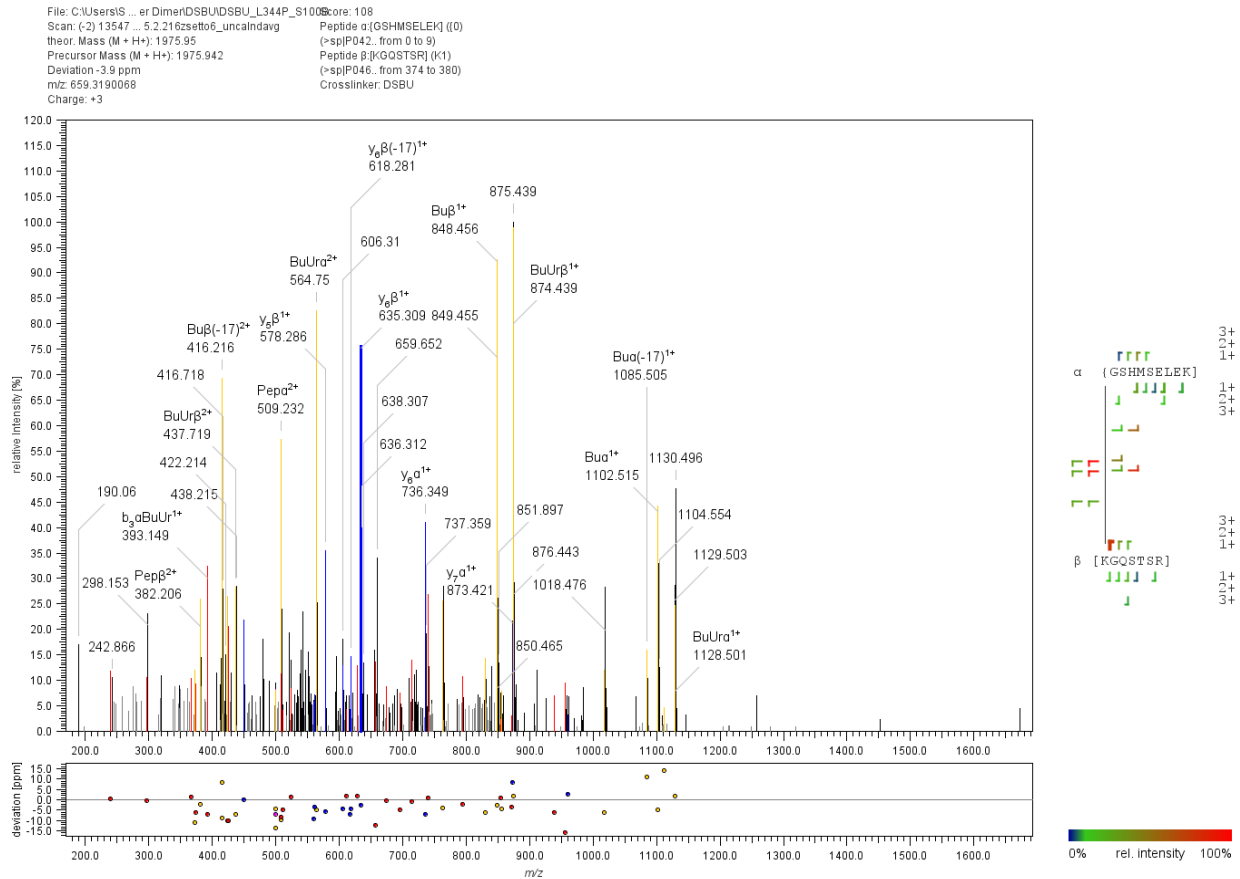


Figure 47. Fragment ion spectrum of the cross-linked product of the L344P monomeric variant with S100 β . The retention time of the precursor ion $[M+3H]^{3+}$ was 31.46 min with a m/z 659.319. Annotation was performed with MeroX.

3.4.2. Cross-linking with sulfo-SDA

Cross-linked products between p53 variants and S100 β generated with sulfo-SDA were compared in the presence and absence of calcium. All cross-linking samples were checked by SDS-PAGE (Figure 48 and Supplementary Figure 2).

		3 replicates						
Sulfo-SDA		+	+	+	+	+	+	+
WT p53	+	+	+	+		+		+
S100 β	+	+	+	+	+		+	+
CaCl ₂		+	+	+	+			
EGTA							+	+

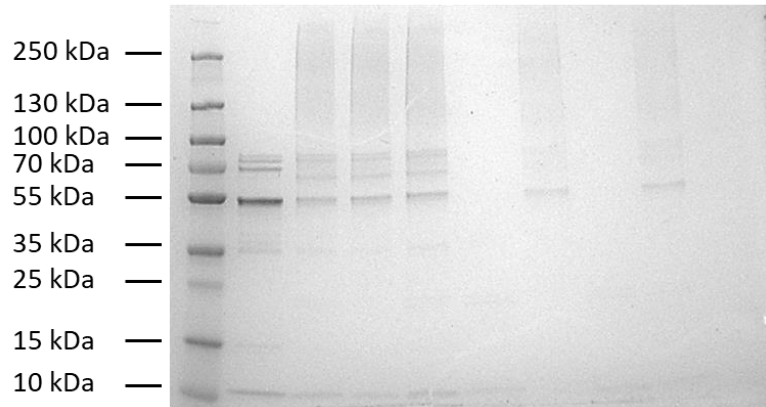


Figure 48. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100 β and wild-type p53 with sulfo-SDA. Gel band at 53 kDa indicates p53 and gel band at 10 kDa indicates S100 β . Three technical replicates were performed for the calcium-loaded state. Calcium-depleted state, with the addition of EGTA, was performed as a negative control.

With sulfo-SDA as the cross-linker, only one intermolecular cross-linked product was observed with wild-type p53, connecting S45 of S100 β to E344 of wild-type p53 (Table 4).

Table 4. Summary of the identified p53-S100 β cross-links with sulfo-SDA.

Wild-type tetrameric							
Peptide 1 (p53)	From	To	Peptide 2 (S100 β)	From	To	Site 1 (p53)	Site 2 (S100 β)
EMFRE <u>L</u> NEALEIK	340	352	EL <u>S</u> HFLEEIK	43	52	E344	S45

The corresponding fragment ion spectrum annotated with MeroX (Figure 49) showed the peptide backbone fragments and the modification of sulfo-SDA (α -peptide with +82 Da) (chapter 1.5.7.3). For the dimeric L344A and the monomeric L344P variant, despite showing different behavior, no cross-linked products were identified.

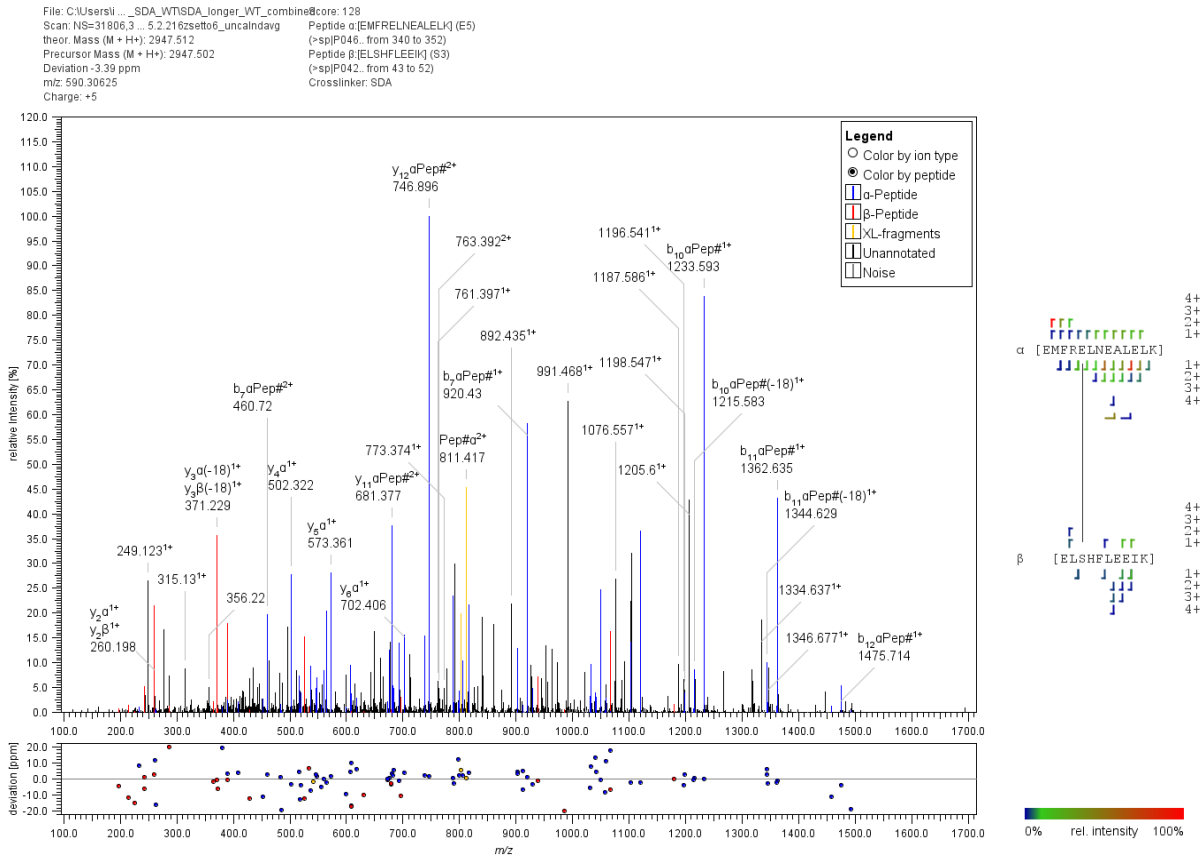


Figure 49. Fragment ion spectrum of the sulfo-SDA cross-linked product of the wild-type tetrameric p53 with S100β. The precursor ion $[M+5H]^{5+}$ has a retention time of 91.37 min, m/z 590.306. Annotation was performed with MeroX.

3.4.3. Cross-linking with EDC/sulfo-NHS

Products of cross-linking between p53 variants and S100β generated with EDC/sulfo-NHS were compared in a calcium-loaded and calcium-depleted state. SDS-PAGE was used to check all cross-linking samples (Figure 50 and Supplementary Figure 3).

EDC/Sulfo-NHS		3 replicates						
		+	+	+	+	+	+	+
WT p53	+	+	+	+		+		+
S100 β	+	+	+	+	+		+	+
CaCl ₂		+	+	+	+			
EGTA							+	+

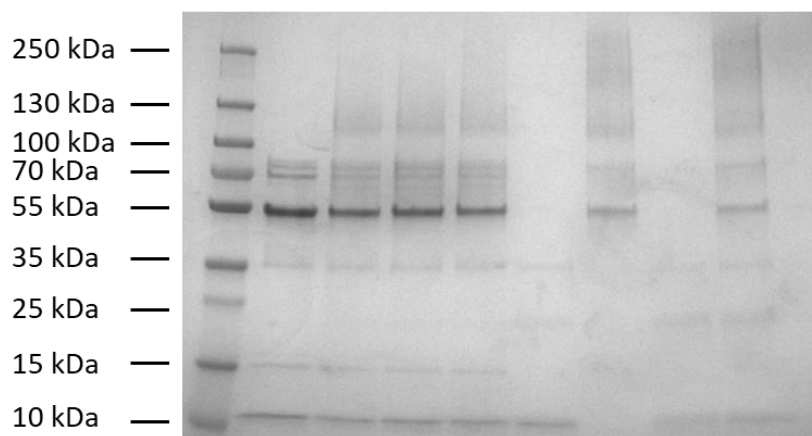


Figure 50. SDS-PAGE (4-20% gradient gel) of XL-MS experiments of S100 β and wild-type p53 with EDC/sulfo-NHS. Gel band at 53 kDa indicates p53 and gel band at 10 kDa indicates S100 β . Three technical replicates were performed for the calcium-loaded environment.

Intermolecular cross-links found with EDC/sulfo-NHS were identical for all three p53 variants. The *N*-terminus of S100 β was cross-linked to D50, D62, D392, and D394 of wild-type p53. For the dimeric L344A variant, the *N*-terminus of S100 β was cross-linked to D50, D62, D392, and D394 of p53. For the monomeric variant, the *N*-terminus of S100 β was cross-linked to D49, D50, D62, D392, and D394 of p53 (Table 5). One exemplary fragment ion mass spectrum is presented in Figure 51 showing a cross-link between E63 of the monomeric L344P p53 and S2 of S100 β .

Table 5. Summary of all identified cross-links with EDC/sulfo-NHS.

Wild-type tetrameric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
<u>D</u> IEQWFTE	50	57	{GSHMSELEK	0	9	D50	N-terminus
DPGP <u>D</u> EAPR	58	66	{GSHMSELEK	0	9	D62	N-terminus
TEGP <u>D</u> SD	388	394	{GSHMSELEK	0	9	D392	N-terminus
TEGP <u>D</u> SD	388	394	{GSHMSELEK	0	9	D394	N-terminus
L344A dimeric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
<u>D</u> IEQWFTE	50	57	{GSHMSELEK	0	9	D50	N-terminus
DPGP <u>D</u> EAPR	58	66	{GSHMSELEK	0	9	D62	N-terminus
TEGP <u>D</u> SD	388	394	{GSHMSELEK	0	9	D392	N-terminus
TEGP <u>D</u> SD	388	394	{GSHMSELEK	0	9	D394	N-terminus
L344P monomeric							
Peptide 1 (p53)	From	To	Peptide 2 (S100β)	From	To	Site 1 (p53)	Site 2 (S100β)
DDL <u>M</u> LSP <u>D</u>	42	49	{GSHMSELEK	0	9	D49	N-terminus
<u>D</u> IEQWFTE	49	57	{GSHMS	0	5	D49	N-terminus
<u>D</u> IEQWFTE	50	57	{GSHMSELEK	0	9	D50	N-terminus
DPGP <u>D</u> EAPR	58	66	{GSHMSELEK	0	9	D62	N-terminus
DPGP <u>D</u> EAPR	58	66	{GSHMSELEK	0	9	E63	N-terminus
LMFKTEGP <u>D</u> SD	384	394	{GSHMSELEK	0	9	D392	N-terminus
TEGP <u>D</u> SD	388	394	{GSHMSELEK	0	9	D394	N-terminus

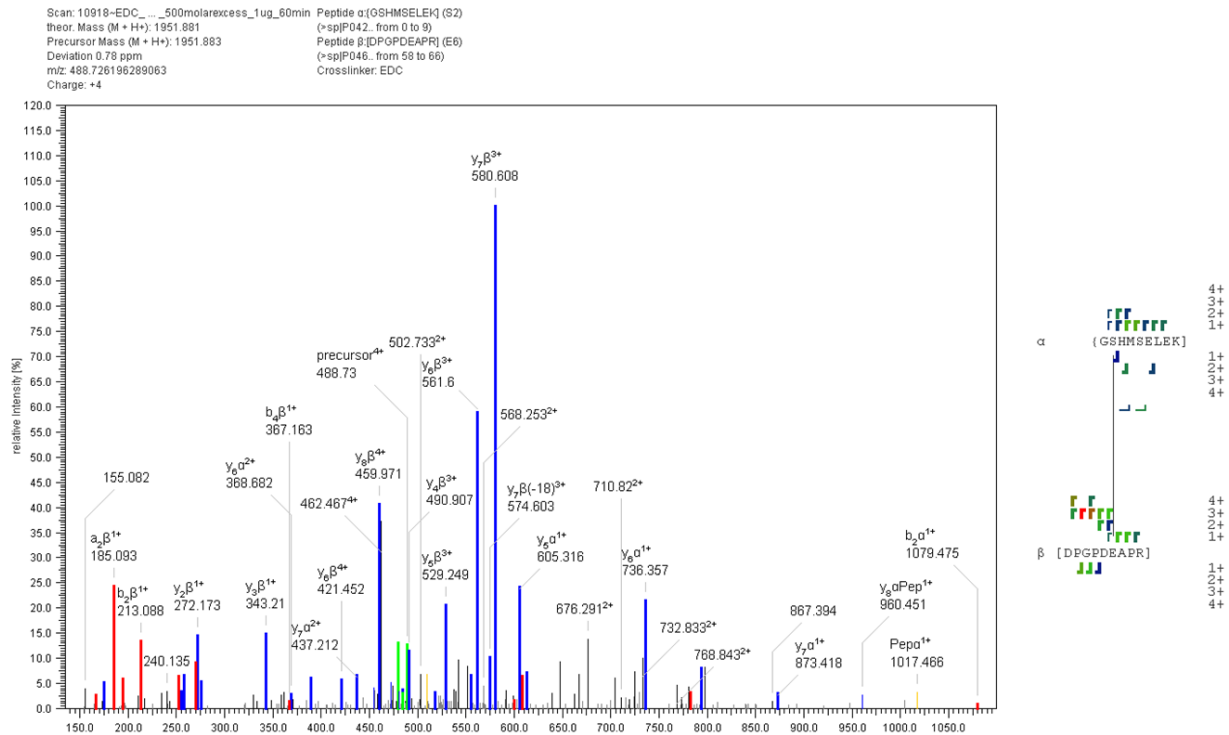


Figure 51. Fragment ion spectrum of the EDC/sulfo-NHS cross-linked product of the monomeric L344P p53 with S100 β . The precursor ion $[M+4H]^{4+}$ has a retention time of 54.65 min, m/z 488.726.

It should be noted that the recombinant S100 β fusion protein contains an artificial *N*-terminus of three amino acids (Gly-Ser-His), which is generated upon cleavage of the thrombin site. According to the experimental results, the majority of the cross-links were found between p53 and the *N*-terminus of S100 β . To exclude artificial cross-links due to these three additional amino acids, cross-linking reactions were repeated using untagged S100 β purified from bovine brain.

Three cross-links were identified from purified S100 β pointing to the same interaction region in p53 as had already been observed with the recombinant S100 β preparations (Table 6).

Table 6. Summary of DSBU cross-links with wild-type p53 and S100 β purified from bovine brain.

Wild-type tetrameric							
Peptide 1 (p53)	From	To	Peptide 2 (S100 β)	From	To	Site 1 (p53)	Site 2 (S100 β)
DAQAGK <u>EP</u> GGSR	353	364	{MSELEK	0	6	K358	<i>N</i> -terminus
AHSSHL <u>K</u> SK	365	373	{MSELEK	0	6	K371	<i>N</i> -terminus
KGQ <u>S</u> TSR	374	380	{MSELEK	0	6	T378	<i>N</i> -terminus

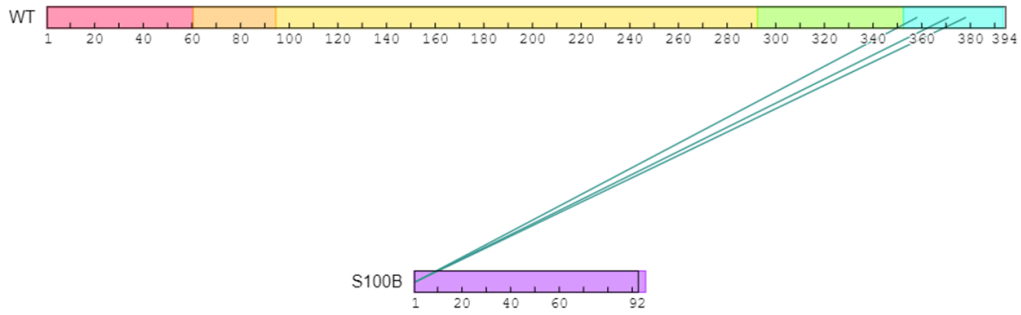


Figure 52. Mapping of cross-linked products of wild-type p53 and S100β purified from bovine brain.

3.5. COMPetitive PAirng StatisticS (COMPASS)

To assess the nature of the cross-links identified between S100β and the p53 variants, whether they originate from one or two different S100β monomers, isotope labeled protein (^{14}N and ^{15}N) was employed. Using the innovative COMPASS strategy that had been developed by Dr. Claudio Iacobucci allowed to unambiguously assign the cross-links found for S100β^[198].

3.5.1. Cross-linking of ^{14}N and ^{15}N S100β

Cross-linking of ^{14}N and ^{15}N S100β was performed with two different cross-linkers, DSBU and sulfo-SDA, to target the amine and carboxylic acid groups of the protein, and monitored by SDS-PAGE (Figure 53).

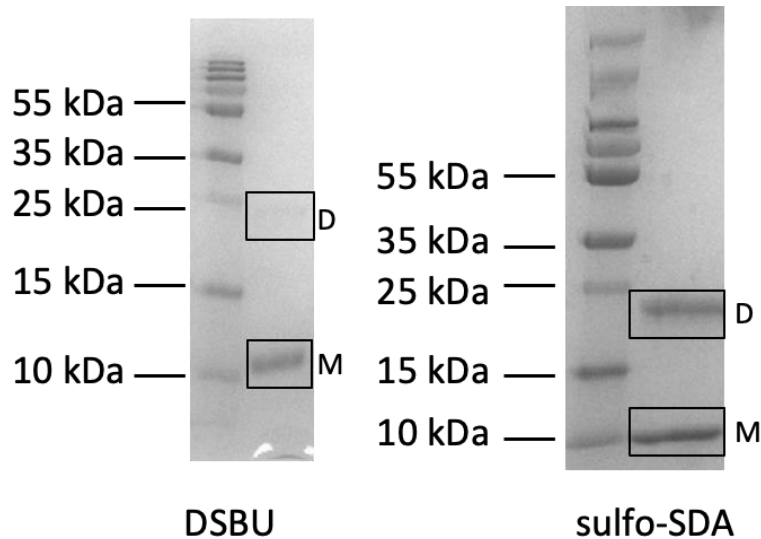


Figure 53. SDS-PAGE (15%) of the cross-links of ^{14}N and ^{15}N S100β with DSBU and sulfo-SDA. D: cross-linked dimeric S100β, M: cross-linked monomeric S100β.

3.5.2. Enrichment of S100β dimer with SEC

SEC separates the cross-linked dimeric and monomeric S100β. For the DSBU cross-linked sample, two elution peaks indicating S100β were observed. The mass spectra of the corresponding elution peaks

indicated the presence of dimeric S100 β (eluted at 6.6 – 7.1 min) and monomeric S100 β (eluted at 7.1 – 7.8 min). Due to the similar elution times of the S100 β monomeric and dimeric species, SDS-PAGE and ESI-MS were performed (Figure 54 and 55). To ensure only the dimeric S100 β was collected and analyzed, fractionation was performed with a 0.2 min increment from 6.2 until 7.6 min. For DSBU cross-linked S100 β , a fraction eluting from 6.8 – 7 min was collected, digested, and subjected to mass spectrometric analysis. Subsequent analyses with MeroX and the integrative COMPASS approach were performed for cross-link identification and quantitation.

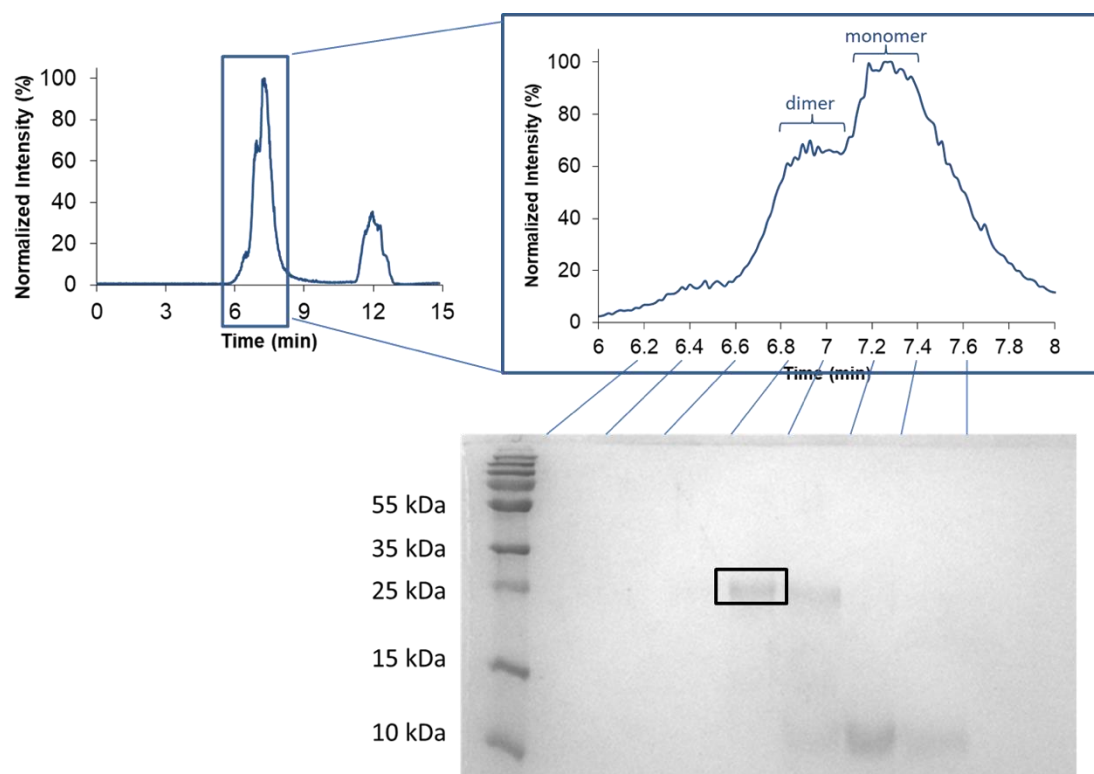


Figure 54. Enrichment with size exclusion chromatography of cross-linked S100 β dimer with DSBU. SDS-PAGE (15%) gel run was performed to check the dimeric and monomeric fractions. To ensure only the dimeric fraction was analyzed, fractionation was performed in an increment of 0.2 min. For MS-analyses, only fraction from 6.8 – 7 min (marked with a black rectangle) was digested and analyzed.

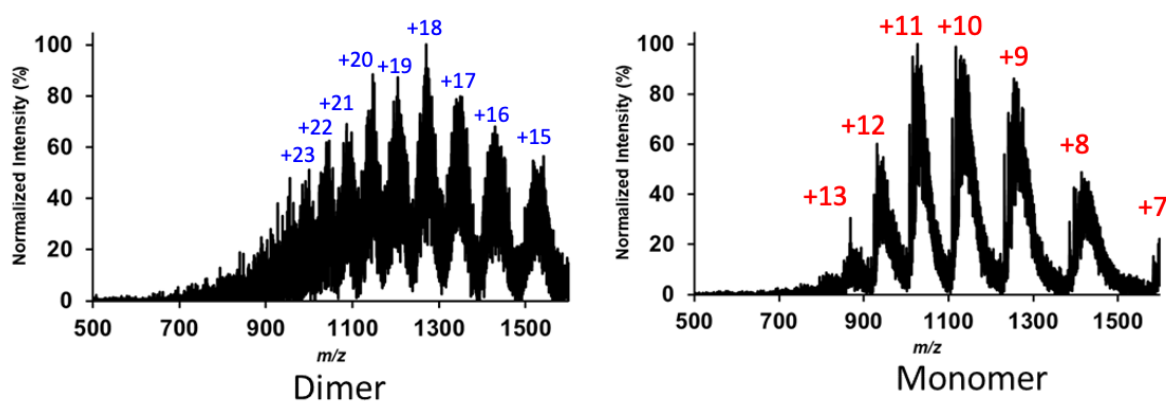


Figure 55. ESI-mass spectra of the dimeric and monomeric fractions from SEC enrichment of the DSBU cross-linked samples. Mass spectra acquired with TQD of (a) the dimeric S100 β (elution time 6.6-7.2 min) and (b) monomeric S100 β (elution time 7.2-7.6 min) with the charge states of all signal peaks labeled.

The same protocol was followed for sulfo-SDA cross-linked samples. The chromatogram and the SDS-PAGE indicated the presence of both the dimeric and monomeric S100 β species (Figure 56). The fraction eluted from 6.6 - 6.8 min was collected due to the presence of the dimeric species. Subsequent analyses was performed with MeroX and COMPASS as described (chapter 3.5.3).

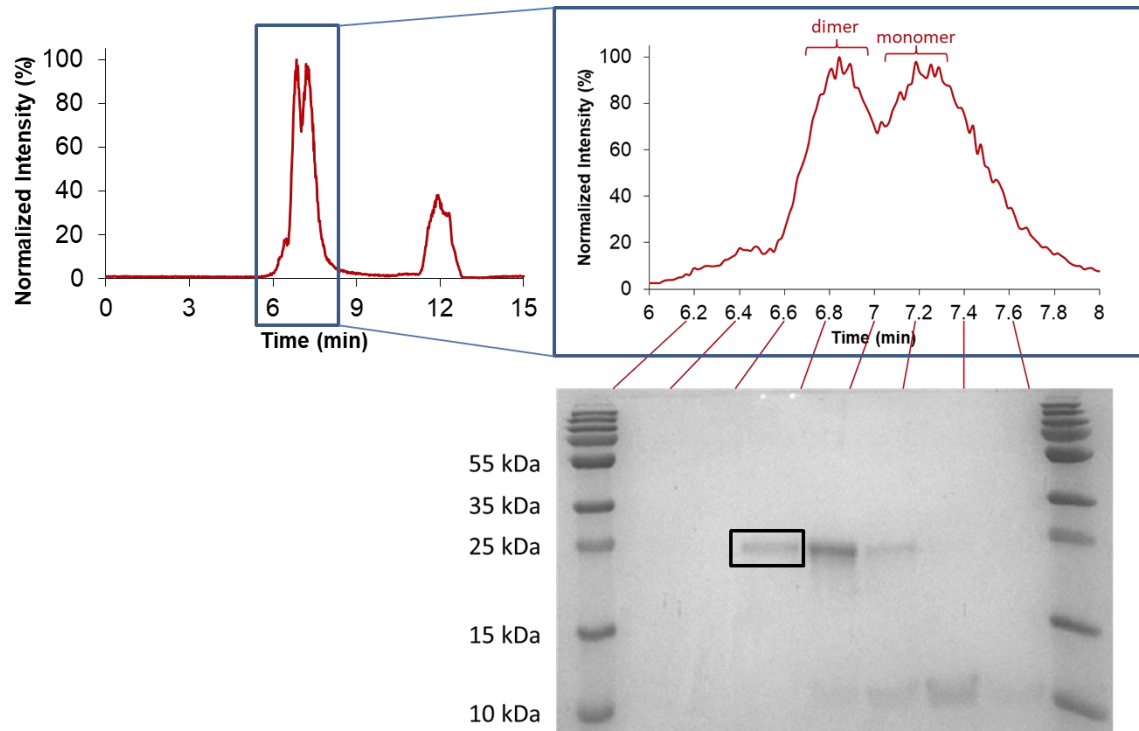


Figure 56. Enrichment with SEC of cross-linked S100 β dimer with sulfo-SDA. SDS-PAGE (15%) gel run was performed to check the dimeric and monomeric fractions. To ensure only the dimeric fraction was analyzed, fractionation was performed in an increment of 0.2 min. For MS-analyses, only fraction from 6.6 – 6.8 min (marked with a black rectangle) was digested and analyzed.

3.5.3. Data analyses with MeroX and COMPASS

For each identified cross-link, information regarding the retention time and the ion mobility were acquired with the MeroX software. With the given information of the cross-linked peptides from MeroX, signal intensities of each species (light/light (L/L): $^{14}\text{N}/^{14}\text{N}$, light/leavy (L/H): $^{14}\text{N}/^{15}\text{N}$, heavy/light (H/L): $^{15}\text{N}/^{14}\text{N}$, and heavy/heavy (H/H): $^{15}\text{N}/^{15}\text{N}$) were extracted and the inter-protein ratio was calculated as described (chapter 1.5.8.). One of the identified DSBU cross-linked peptides was $^9\text{KAMVALIDVFHQYSGR}^{24} - ^{51}\text{ikeqevvd}\underline{\text{k}}\text{vme}^{62}$, which connected K9 to K59 of the light and the heavy

S100 β monomers (Figure 57). The signal intensities of each isotope pattern were observed and the calculated inter-protein ratio (calculated with the equation given in Figure 27b) was 97.7% (Figure 58).

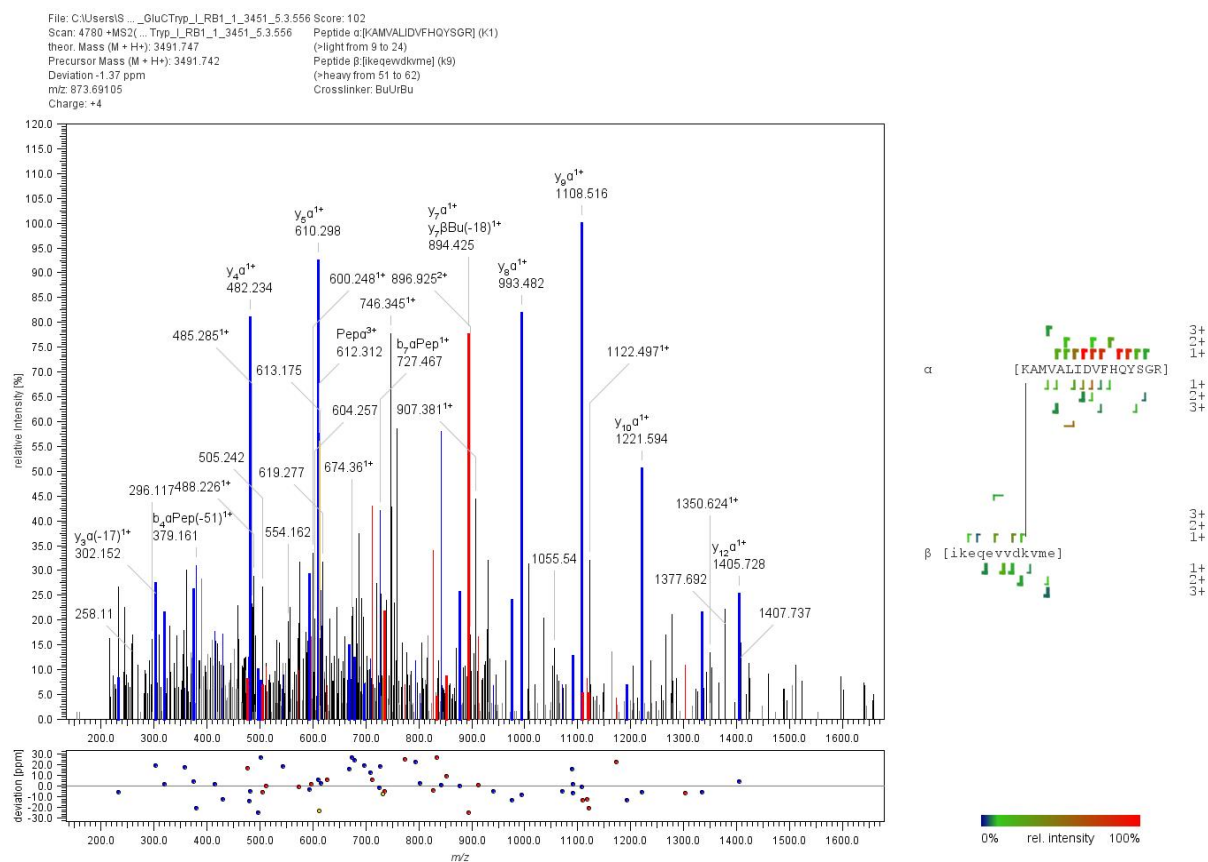


Figure 57. Fragment ion mass spectrum of the cross-linked product between ^{14}N and ^{15}N S100 β $^9\text{KAMVALIDVFHQYSGR}^{24}$ – $^{51}\text{ikeqevvdkvme}^{62}$. The precursor ion $[\text{M}+4\text{H}]^{4+}$ has a retention time of 50.43 min, m/z 873.691, Annotation was performed with MeroX. Capital letters of the residues represent the light ^{14}N and the lower case represents the ^{15}N heavy isotopically labeled.

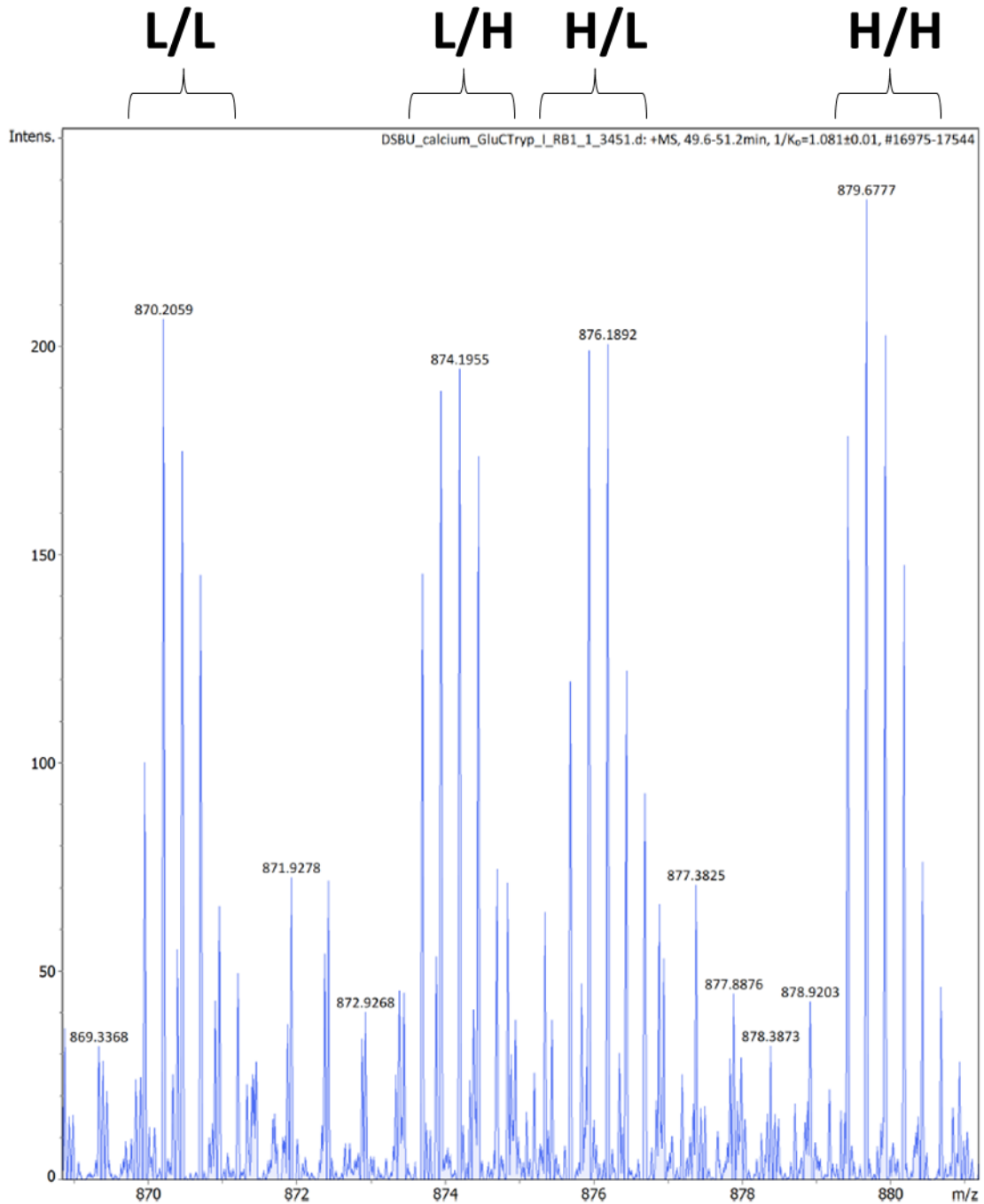


Figure 58. COMPASS cross-link analysis of the K9 to K59 cross-link. Full-scan mass spectrum showing signals of the DSBU cross-linked peptide between the K9 and K59 of ¹⁴N and ¹⁵N S100β with a retention time 49.6-51.2 min, scan number 16975-17544. The inter-protein ratio of this cross-linked peptide is 97.7%.

For the calcium-loaded state, an additional DSBU cross-link was identified connecting *N*term to K52. Two additional cross-links (*N*term-D16 and *N*term-E49) were identified using sulfo-SDA as the cross-linker. A schematic representation of all the acquired cross-links from both cross-linkers is illustrated in Figure 59. Cross-link quantitation (Equation in Figure 27b) is summarized in Table 7.

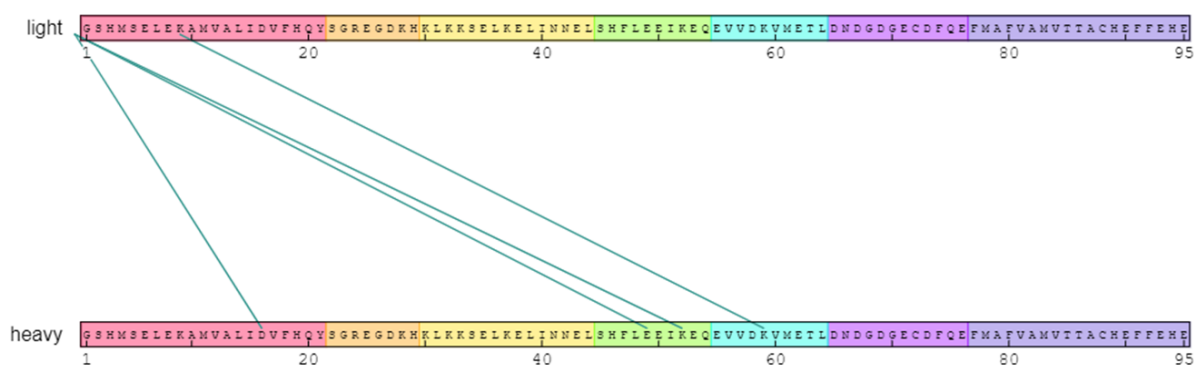


Figure 59. Summary of identified inter-protein crosslinks of the calcium-bound state of ^{14}N and ^{15}N S100 β with DSBU and sulfo-SDA mapped with xiNET. For the calcium-depleted state, only two cross-links (*N*term-K52 and *N*term-E49) were identified.

Table 7. Summary of inter-protein cross-links percentage of ¹⁴N and ¹⁵N S100β. Spectrum that was used for cross-link quantitation with the COMPASS approach can be found in Figure 58 and Supplementary Figures 10, 11, and 12.

DSBU		
	+ Ca2+	- Ca2+
K9-K59	97.70 %	Not detected
	97.44 %	
	-	
Nterm-K52	6.59 %	13.60 %
	12.93 %	14.40 %
	12.16 %	24.00 %
Sulfo-SDA		
	+ Ca2+	- Ca2+
Nterm-D16	31.80 %	Not detected
	32.30 %	
	35.90 %	
Nterm-E49	3.95 %	2.98 %
	10.10 %	3.04 %
	-	4.97 %

3.6. Surface Plasmon Resonance

Prior to ligand immobilization, the optimal pH of sodium acetate for S100β was determined. Three pH values (pH=3.6, 3.9, and 4.2), between the pI value of the carboxymethyl dextran surface (pI=3.5) and the pI value of S100β (pI=4.5), were examined (Figure 60). The optimal pH of S100β immobilization was selected from the pH value with the highest response during the pH scouting step. The highest response was at pH=3.9, therefore, it was used as the pH for immobilization.

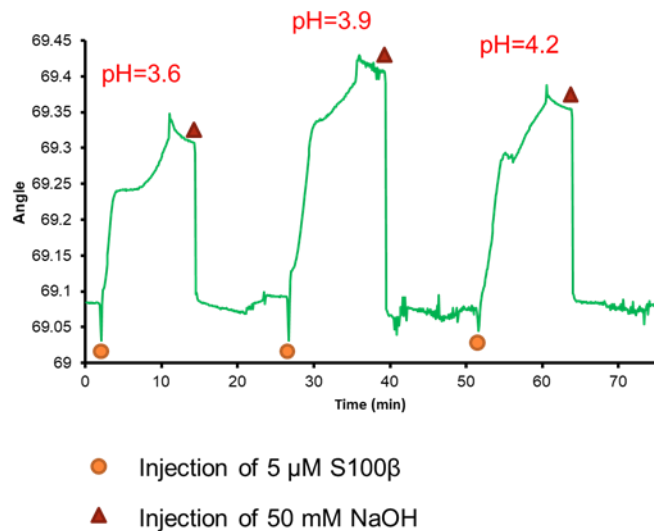


Figure 60. Ligand preparation – pH scouting with 10 mM sodium acetate. Orange circles indicate the injection of the ligand S100 β and the red triangles indicate the sodium hydroxide rinse to remove the bound S100 β . As observed from the response, the optimal pH for ligand immobilization is at pH = 3.9.

Activation of the sensor surface was performed by three successive rounds of injection of NHS/EDC. After the activation, the sensor surface was prepared for amine-coupling with the injected ligand. The ligand, S100 β , was injected in five rounds to ensure maximum immobilization. Deactivation of the unreacted NHS-esters was performed with an injection of 1 M ethanolamine. The difference in the SPR angle, indicated by the y-axis, implicated the successful immobilization of S100 β (Figure 61).

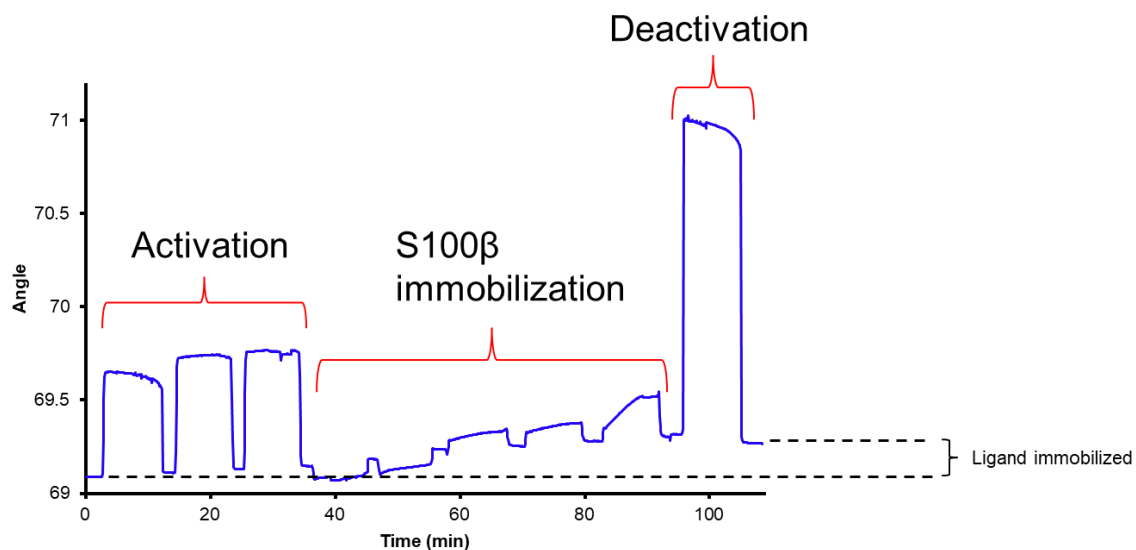


Figure 61. Ligand immobilization of S100 β . showing the response of the activation step with EDC/sulfo-NHS, three rounds of consecutive ligand injection, and deactivation with ethanolamine. The difference of response after deactivation showed the amount of ligand successfully immobilized.

Analyte Injection

After ligand immobilization, analyte injection was performed individually for each p53 variant. For the wild-type p53 (Figure 62a), a binding affinity of 41 μM was measured. For the dimeric L344A p53 variant (Figure 62b) the binding affinity was 60 μM and for the monomeric p53 variant (Figure 62c) a binding affinity of 40 μM was measured. A summary of the binding response according to the concentration of analyte injected was plotted (Figure 62d).

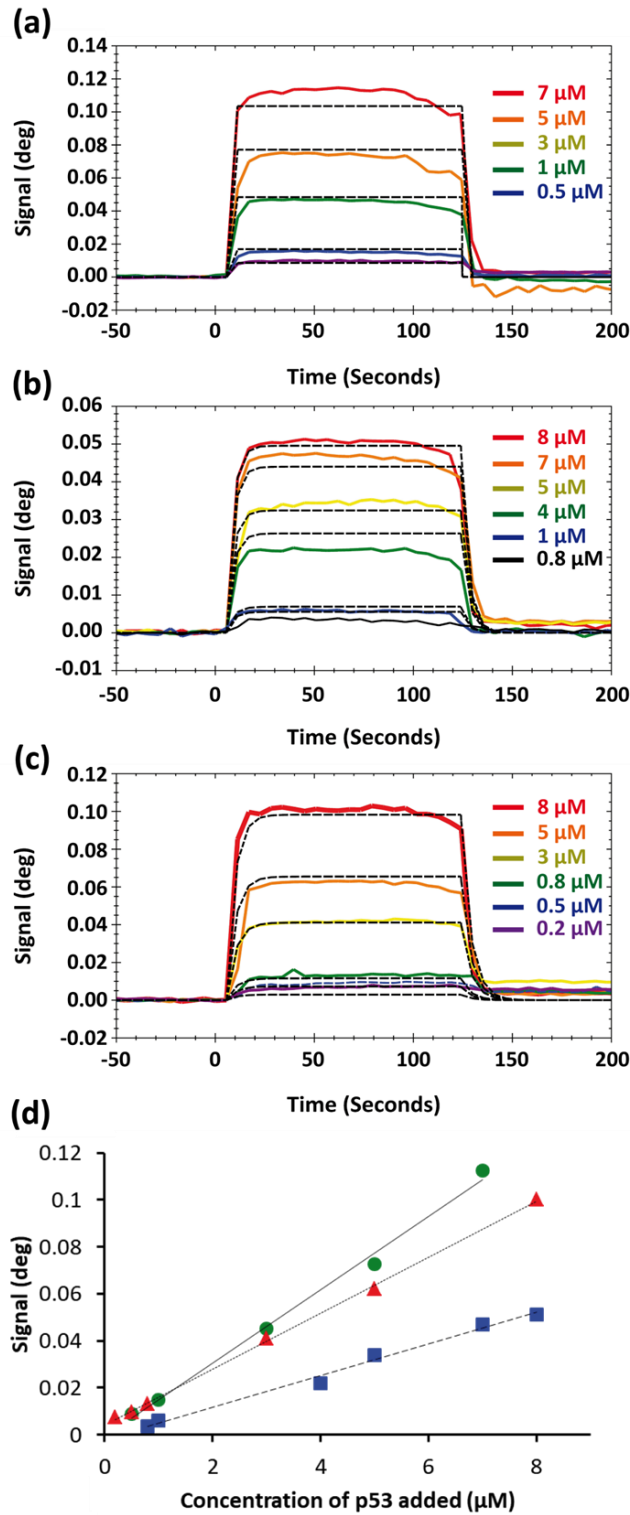


Figure 62. Analyte injection of p53. (a) wild-type tetrameric p53 (b) L344A dimeric p53 (c) L344P monomeric p53 (d) summary of responses of all three variants (green circles – wild-type, blue squares – L344A p53, red triangles, L344P p53). Fitting was performed with TraceDrawer and binding affinity was analyzed.

4. Discussion

4.1. Expression and purification of full-length p53

The tumor suppressor protein p53 is active as a homotetramer. Mutations in p53 often lead to the loss of the tetrameric state and can significantly decrease p53's activity as a tumor suppressor. While it has been well-established that a much higher frequency of mutations happens in the DBD of p53^{[208][209]}, mutations occurring in the TD are observed and mutations have a great influence on the tetramer formation of p53^[210]. Reported thermal denaturation studies have shown that some mutants are thermally unstable near body temperature. Some mutants exist as unfolded monomers (L330R, L344P, and R342P) and some can only form dimers at physiological conditions (F341C, L344A, and A347T)^{[45][211]}. For the p53 variants possessing different oligomeric states, two variants were selected for this study. The dimeric L344A and the monomeric L344P variants show biological significance, as they have been observed from patients with the Li-Fraumeni Syndrome (LFS)^[203]. The protocol for protein purification of the wild-type tetrameric, L344A dimeric, and L344P monomeric p53 has been previously established and was reproduced for this study (chapter 2.2.1.1.)^[202,203]. Native MS was used to confirm the protein activity and the oligomeric states of each p53 variant. Wild-type p53 confirmed the presence of tetrameric, dimeric, and monomeric p53 species. Native MS revealed that the L344A variant contained both the dimeric and monomeric p53, while the L344P variant contained only the monomeric p53.

It has been previously reported that for IDPs, different structures would be observed in-solution compared to gas-phase. For p53, the gas phase collapse phenomenon explained the observation of a smaller cross-collision section (CCS) in ion mobility mass spectrometry (IM-MS) compared to the CCS calculated in a small-angle x-ray scattering (SAXS) model^[212]. Owing to the higher flexibility of IDPs, the four p53 monomers could be closer than the predicted SAXS model^[171]. Furthermore, previous in-solution cross-link results have demonstrated that the p53 tetramer was more compact than calculated. While the effect of the gas phase collapse is not to be ruled out completely, there is a correlation between the in-solution and the gas phase p53 structures^[213].

In addition to p53, presence of the chaperone protein DnaK was observed. Despite multiple purification steps, DnaK remained in the sample. The chaperone protein DnaK assists in protein folding of p53. To completely remove DnaK from the sample, an ATP-agarose column could be used to remove traces of DnaK, however, this could further complicate the experimental approach as the complete removal of the chaperone protein could lead to protein aggregation^[214].

4.2. Expression and purification of S100 β

Purification of the light ^{14}N S100 β followed the previously established protocol (chapter 2.2.1.2.)^[204]. The S100 β protein forms a homodimer in physiological conditions. Despite many structures of the S100 β homodimer have been reported, site-specific interactions within the homodimer has yet to be reported. The S100 β homodimer undergoes a structural rearrangement upon calcium addition^[65]. Structural investigation into the conformational change within the homodimer is crucial as it plays an essential role on the p53-S100 β interaction^[63]. To that, the integrative COMPASS approach was performed (chapter 4.4.). The ^{14}N and the isotope-labeled ^{15}N S100 β were both expressed and purified in order to observe the cross-links connecting the homodimer. The purification of the isotope-labeled ^{15}N S100 β was purified in minimal medium and the purification steps followed the same protocol. Native MS was performed to verify the mass difference, 124 Da, between the light (^{14}N) and the “heavy” isotope-labeled (^{15}N) S100 β .

4.3. Protein-protein interaction of p53 and S100 β

Reported p53-S100 β protein-protein interaction studies have so far been limited to only p53 peptide studies^{[78][67]}. While using peptides or truncated versions of p53 shed a light on the p53-S100 β protein-protein interaction, it lacks complete structural insights. Binding events reflect the protein interaction with the given domains but they do not reflect the binding mechanism of the full-length protein. This could result in a wrong conclusion regarding the structural characterization. To overcome this challenge, full-length constructs of p53 were used and the interaction with S100 β was characterized by covalently modified ESI-MS, XL-MS, and SPR. ESI-MS with the introduction of covalent modification observes the stoichiometry and the protein subunit composition of the p53-S100 β complexes. XL-MS provides structural insights regarding the site-specific interaction between p53 and S100 β . SPR determines the binding affinity between the two proteins.

Investigation into the binding events of different p53 oligomers were performed. The p53 oligomeric state depends on the protein concentration^[49]. One approach for tuning the p53 oligomeric state is to lower the concentration so that monomeric, dimeric, tetrameric p53 could be investigated individually based on the p53 protein concentration. The dimerization constant of p53 is 1 nM and the equilibrium constant for p53 tetramer formation is approximately 100 nM^[215]. Experiments with lower concentrations of p53 could be possibly challenging, as the reported *in vivo* dimerization constants might not reflect on the *in vitro* studies. Additionally, a sufficient amount of protein was required to perform MS and SPR experiments. Hence, as described in chapter 2.2.1.1., three p53 variants each representing a different oligomeric state were produced. With the successful purification of the three oligomeric states of p53, there was less limitations on the p53 protein concentration. The wild-type

tetrameric, the L344A dimeric, and the L344P monomeric, were each investigated individually regarding the binding interaction with S100 β .

4.3.1. Covalently modified ESI-MS of p53-S100 β protein complexes

A contradicting binding interaction has been observed between p53 and S100 β according to the reported analytical ultracentrifugation (AUC) and fluorescence anisotropy experiments (chapter 1.4. and Table 1)^[78]. Biological studies revealed two regulatory pathways of S100 β on p53^{[73][72]}. It has been shown that when bound to monomeric p53, S100 β exerts an inhibitory function. This pathway inhibits the tetramer formation of p53 and causes p53 to lose its tumor suppressor activity^[77]. On the other hand, it has been observed that S100 β binds strictly to tetrameric p53, stabilizing the p53 tetramer, and enables its activity as a tumor suppressor protein^[216]. In this study, the stoichiometry of the p53-S100 β complexes was investigated^[206].

Here, the stoichiometry of the p53-S100 β protein complexes from each of the three variants representing the tetrameric (wild-type), dimeric (L344A), and monomeric (L344P) p53, were investigated individually with covalently modified ESI-MS by injecting the protein samples into the High Mass Q-TOF II mass spectrometer. For many protein complexes, native mass spectrometry (native MS) allows visualization of the composition of the protein subunits and the stoichiometry of the protein complexes. For this study, due to the reported low binding affinities, observing the complex formation without a mild covalent modification (native MS) to fix the protein complexes was challenging and no conclusion could be made (attempt with native MS shown in Supplementary Figure 5). To overcome the challenge, a covalent fixation, a very mild-cross-linking reaction was performed using the DSBU cross-linker. A mild covalent modification with DSBU showed that p53 and S100 β were still at their native states, as the charge states were retained. This showed the DSBU-modified p53-S100 β complexes were measured at their native states and no bias was introduced for the conducted experiments (Figures 35-37, the charge states of p53 remained the same in comparison to Figures 40, 42, and 44). The addition of calcium exposes the hydrophobic pocket of S100 β and binding interaction took place (Figures 40a, 42a, and 44a). A negative control was performed with the addition of a chelating agent, EGTA. Upon addition of EGTA, a calcium-depleted environment was created, and no p53-S100 β complexes was observed (Figure 40b, 42b, and 44b). From the calcium-loaded samples, the stoichiometries of the p53-S100 β complexes were revealed. From the calculated MW from the full scan MS, the tetrameric wild-type p53, upon binding to S100 β , showed that the tetrameric p53 binds a dimeric S100 β (protein complex with a MW of 200 kDa). The dimeric L344A variant showed the binding interaction of a p53 dimer to an S100 β dimer (protein complex with a MW of 111 kDa). The monomeric L344P variant showed the binding interaction between a p53 monomer and an S100 β dimer (protein complex with a MW of 66 kDa). This observation contradicts the reported p53 peptide

studies. It was shown that the tetrameric, dimeric, and the monomeric p53 all bind to an S100 β dimer^[206].

Full scan MS provided the structural insights regarding the stoichiometry of protein complexes. Another technique to verify the composition of the p53-S100 β complexes was performed with tandem MS (CID-MS/MS). By applying a higher collisional energy and a higher pressure in the collision cell, the dissociation of the p53-S100 β complexes have all shown a S100 β monomer being ejected from the complexes (tetrameric wild-type shown in Figure 41, dimeric L344A shown in Figure 43, and the monomeric L344P variant shown in Figure 45). This further confirms the composition of the complexes and the identities of the subunits.

Experimental results from the covalently modified ESI-MS contrasts the existing knowledge of the protein interaction of p53-S100 β . Investigation into the stoichiometry has shown that the binding interaction is independent from the oligomeric state of p53. Contradicting the reported binding mechanism, this study showed that the dimeric p53 also interacts with S100 β in the identical fashion as tetrameric and monomeric p53 (Figure 63).

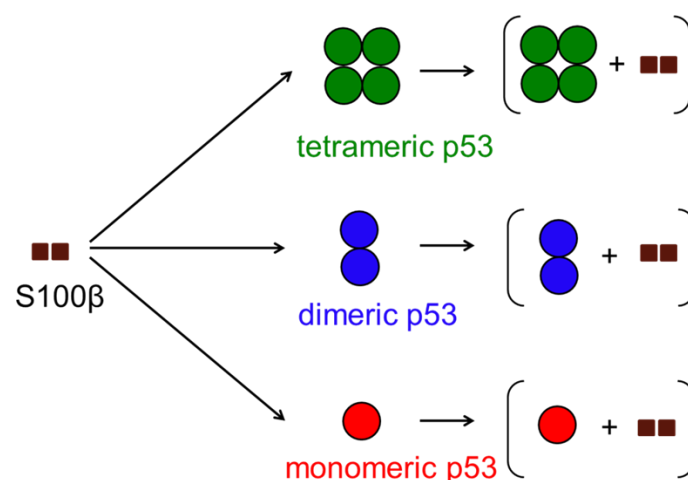


Figure 63. Schematic representation of the stoichiometry of p53-S100 β complexes upon investigation with covalently modified ESI-MS. The stoichiometry of the protein complexes is independent of the protein-protein interaction. All p53 oligomeric states interact with S100 β in the same manner.

4.3.2. Cross-linking mass spectrometry

Cross-linking mass spectrometry (XL-MS) connects proximal, surface-exposed amino acid residues, yielding structural insights by targeting various chemical reactivities (amine- or thiol-reactive cross-linkers). Advantages of XL-MS include the requirement of a very low protein amount and a high chemical specificity to reveal the site-specific interaction among different proteins^[217]. However, there

are some limitations to be taken into consideration when performing cross-linking experiments. While using the amine-reactive cross-linkers, the distance constraints obtained from the data have to be analyzed carefully. The lysine side chains are highly flexible, therefore, complementary cross-linking data using other cross-linkers should be compared in order to come to an accurate conclusion regarding the protein structure^[82]. Excessive cross-linking leads to distortion of the protein structures, which results in an inaccurate depiction of the structural data of the protein-protein interaction^[218]. Covalently modified ESI-MS can be used as a step to verify the optimal conditions of protein cross-linking. Previous studies have reported the binding interaction of S100 β to the C-terminus of p53, though based solely on p53 peptide studies. To investigate the site-specific interaction between p53-S100 β , XL-MS was performed using different cross-linkers that target different chemical reactivities in order to ensure complete coverage of the proteins.

The cross-linker DSBU connects mainly the amine-reactive side chain of the lysine residues and the N-termini of proteins^[173]. Experimental results have shown that for the wild-type tetrameric p53, cross-links were observed between the C-terminus of p53 and the N-terminus of S100 β . The same behavior was observed for the dimeric L344A and the monomeric L344P variants, as cross-links connecting the C-terminus of p53 and the N-terminus of S100 β were also observed (Table 3). The cross-linker sulfo-SDA targets unspecific amino acid residues^[195]. Result has shown one cross-link between E354 of p53 and S45 of S100 β (Table 4) for only the wild-type tetrameric p53. No cross-link was observed between S100 β and the dimeric L344A or the monomeric L344P variant. With the combination of EDC/sulfo-NHS, cross-links connecting the N- and the C-termini of wild-type tetrameric p53 to the N-terminus of S100 β were identified (Table 5). The same behavior was observed for the dimeric and the monomeric variants of p53.

Binding of S100 β to the N-terminus of p53 contradicts the current existing knowledge, as the reported protein-protein interaction has only shown binding interaction to the C-terminus of p53. Structural studies so far have agreed to one solution-state NMR structure (PDB: 1DT7), which showed the regulatory domain (amino acid residues 352 to 393) of p53 as the binding interface when bound to S100 β . Structural modeling has been reported and there have been multiple proposed models of the full-length p53. Structural modeling from Okorokov^[219] and Fersht^[220] have indicated the proximity between the N- and the C-terminus of the full-length p53, which could explain the observation of cross-links to the acidic N-terminus using the EDC/sulfo-NHS cross-linker.

The construct of the self-produced recombinant S100 β protein contains three residual amino acid residues on the N-terminus (glycine-serine-histidine) after thrombin cleavage. As the results indicated the majority of cross-links to the N-terminus of S100 β , validation of the observed cross-links was necessary to avoid a false structural characterization. Thus, commercially available S100 β , without the

artificial *N*-terminus, was purchased. XL-MS results with the commercial S100 β have also shown cross-links connecting the *N*-terminus of S100 β . This further verified the results obtained earlier with XL-MS. The results also showed the cross-links to the *N*-terminus of the commercial S100 β with no artificial residues. It could therefore be concluded that the cross-links generated were valid from this control experiment (Figure 52, Table 6).

XL-MS results obtained from three different cross-linkers showed an identical binding behavior of p53 towards S100 β (Figure 64), similar to the results acquired with ESI-MS. The p53 oligomeric state did not influence the outcome of the protein-protein interaction. One notable observation from ESI-MS was the presence of the chaperone protein DnaK. Results from XL-MS however, did not identify any cross-links between DnaK and p53 or to S100 β . This indicated that despite having DnaK in the sample, it did not interfere with the structural investigation of the binding interaction of p53 and S100 β . Removal of the chaperon protein was proved unnecessary with the experimental results as there has been no interference with the p53-S100 β study.

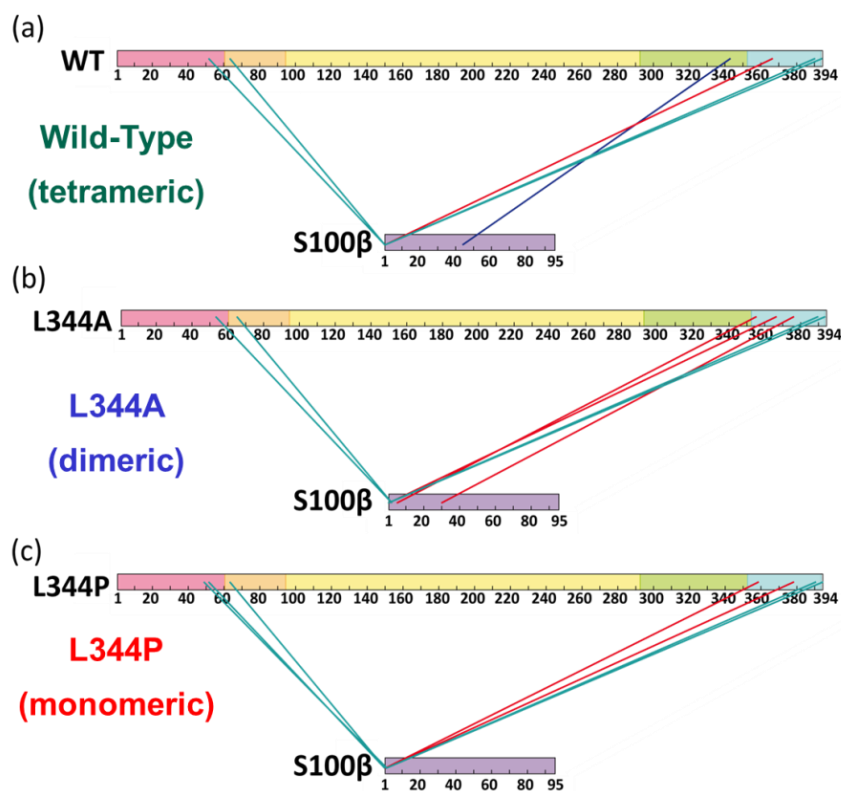


Figure 64. Summary of XL-MS of S100 β and different p53 variants. (a) Tetrameric wild-type (b) dimeric L344A variant (c) monomeric L344P variant with red line showing cross-links with DSBU, cyan line showing cross-links with EDC/sulfo-NHS, and the blue line showing cross-links with sulfo-SDA.

4.3.3. Surface plasmon resonance

Various p53 peptides were used in the reported studies of the binding affinity (K_D) (Table 1). Measurements of the binding affinity have previously been performed with analytical ultracentrifugation (AUC), fluorescence anisotropy, isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) with peptides or truncated versions of different domains of p53^[78]. Due to the great differences among the binding affinities obtained from different regions of the C-terminus of p53, characterizing the real binding affinity using full-length p53 has become essential in order to gain insights of the accurate binding mechanism in this work. SPR was chosen to examine the thermodynamic parameters. Immobilization of S100 β was performed and p53 was injected in increasing concentrations. The binding affinity was calculated and compared.

The three p53 variants, representing the full-length versions of the monomeric, dimeric, and tetrameric p53 were investigated individually. The binding affinity of wild-type p53 to S100 β has a K_D value of 40 μ M, the L344A dimeric p53 has a binding affinity of 60 μ M, and the monomeric L344P variant has a binding affinity of 41 μ M. The binding affinity of the three p53 variants showed a similar binding behavior. In combination with acquired mass spectrometric experimental results, p53 oligomeric state does not influence its binding behavior towards S100 β .

Measuring the binding affinity (K_D) with SPR comes with several advantages. First, measuring time for SPR is relatively fast. Second, the sensor surface with the immobilized ligand is reusable and multiple rounds of analyte injection is performed upon regeneration^[221]. However, there are certain disadvantages that could hinder the success of an SPR experiment and the accurate determination of the binding affinity between proteins. Protein stability at a higher temperature and the possibility of protein aggregation at higher concentrations could lead to inaccurate conclusions of the K_D values^[222]. Other experiments such as ITC^[223] and microscale thermophoresis (MST)^[224] could be performed to compare the experimental results, however, the two techniques also have the above-mentioned shortcomings that need to be addressed. ITC introduced a higher temperature to the protein that could lead to protein aggregation and the loss of protein stability^[225]. MST requires high protein concentrations in order to generate a calibration curve to determine the binding affinity. Moreover, the design of MST requires one protein binding partner to not have any aromatic residues^[226]. Taken into consideration of the different limitations, SPR was the most suitable system for this study, as the high protein concentration or a higher temperature were not required.

4.4. COMPetitive PAiring Statistics

Various methods of the quantitation of protein cross-links have been reported, for example, isobaric tags for relative and absolute quantitation (iTRAQ) and tandem mass tag (TMT) labeling. In order to

structurally characterize cross-links between homodimers and to perform cross-link quantitation, the innovative approach Competitive Pairing StatisticsS (COMPASS) was applied to characterize the structural investigations of the S100 β homodimer.

With the isotopically-labeled version (heavy nitrogen, ^{15}N) of the same protein, cross-links between two different monomers (inter-protein cross-link, represented by light/heavy: $^{14}\text{N}/^{15}\text{N}$ or heavy/light: $^{15}\text{N}/^{14}\text{N}$) or within the same monomer (intra-protein cross-link, represented by light/light: $^{14}\text{N}/^{14}\text{N}$ or heavy/heavy: $^{15}\text{N}/^{15}\text{N}$) could be easily differentiated. By calculating the relative intensity of the intra-protein cross-link and the inter-protein cross-links, the technique sheds a light on the structural conformation of the protein. For studying IDPs, the change in the cross-link abundance can also indicate a change in protein conformation, for example, induced conformational changes or observing the disorder-to-order transition, as observed in structural studies of alpha-synuclein on the molecular mechanism for liquid droplet formation depending on the intra/inter-protein percentage^[198]. For characterizing globular proteins, like S100 β , cross-link quantitation offers a tool to compare the conformational change in various conditions. For this study, the calcium-induced conformational change was investigated.

The experimental results from the SEC-enriched COMPASS-analysis showed a high similarity between the calcium-loaded and the calcium-depleted conditions (Table 7). From the structural perspective, the major difference between the two structures was helix 3 (amino acid residues 55-64). XL-MS results and quantitation with COMPASS showed a high similarity of the site-specific interaction and very similar results of the percentage of inter-protein cross-link. The S100 β protein undergoes a conformational change upon the addition of calcium ions^[65] (chapter 1.3.). Specifically, with the addition of calcium, the helix from residue 55 – 64 (helix III) undergoes a conformational change that exposes the binding pocket. One cross-link, K9-K59, which connected helix I and helix III, was only found in the calcium-bound state. In the calcium-depleted conditions, the particular cross-link (K9-K59) was not identified (Table 7). This conformational change, evidenced by the cross-link results, showed the calcium-induced change of the S100 β homodimer.

5. Conclusions and Outlook

The integrated approach, combining covalently modified ESI-MS, XL-MS, and SPR provides a solid basis for understanding the binding mechanism between p53 and S100 β . ESI-MS showed that the binding mechanism of p53 and S100 β is identical among all three p53 variants with regards to the stoichiometry, as all three oligomeric states of p53 bind to one S100 β dimer. XL-MS showed that the site-specific interaction occurred in a similar fashion, regardless of the oligomeric state of p53, as identified cross-links suggested binding occurred on both the *N*- and *C*-terminus of p53 and the *N*-

terminus of S100 β . With regards to the binding affinities, all three p53 oligomeric states showed comparable binding affinities towards S100 β . In summary, the molecular mechanism and the protein-protein interaction between full-length p53 and S100 β have been well-characterized in this study. The novelty of structural investigation of p53-S100 β using three variants of full-length p53 brings new insight into the protein-protein interaction of p53-S100 β . The integrative approach used in this study contrasts the current understanding of the p53-S100 β interaction. The identical behavior of the tetrameric, dimeric, and monomeric p53 brings new insights into the biological knowledge. Additionally, quantitation of intra/inter-protein cross-links was performed via the innovative COMPASS method. Structural insights into the homodimer S100 β was gained, and the differentiation of cross-links within the same monomer and between two different monomers was achieved.

For future work, more protein-protein interactions with full-length p53 can be investigated. Due to many modifications of p53 occurring *in vivo* and their reported different binding mechanisms, studies detailing the effect of p53 stabilization and whether the binding affinity would be greatly influenced by the PTMs can be further characterized. A thorough study into the p53-S100 β studies *in vivo* in comparison to the *in vitro* studies reported in this thesis is to be conducted.

References

- [1] P. E. Wright, H. J. Dyson, *J. Mol. Biol.* **1999**, *293*, 321–331.
- [2] F. Sanger, *Science (80-.)*. **1959**, *129*, 1340–1344.
- [3] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662–666.
- [4] C. B. Anfinsen, *Science (80-.)*. **1973**, *181*, 223–230.
- [5] V. N. Uversky, *Chem. Rev.* **2014**, *114*, 6557–6560.
- [6] F. Eisenhaber, B. Persson, P. Argos, *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 1–94.
- [7] A. Ben-Naim, *Open J. Biophys.* **2012**, *02*, 1–3.
- [8] B. Wallner, A. Elofsson, *Protein Sci.* **2003**, *12*, 1073–1086.
- [9] C. Levinthal, *J. Chim. Phys.* **1968**, *65*, 44–45.
- [10] R. Zwanzig, A. Szabo, B. Bagchi, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 20–22.
- [11] R. Huber, *Trends Biochem. Sci.* **1979**, *4*, 271–276.
- [12] Z. Peng, M. J. Mizianty, B. Xue, L. Kurgan, V. N. Uversky, *Mol. Biosyst.* **2012**, *8*, 1886–1901.
- [13] L. Piersimoni, M. Abd el Malek, T. Bhatia, J. Bender, C. Brankatschk, J. Calvo Sánchez, G. W. Dayhoff, A. Di Ianni, J. O. Figueroa Parra, D. Garcia-Martinez, J. Hesselbarth, J. Köppen, L. M. Lauth, L. Lippik, L. Machner, S. Sachan, L. Schmidt, R. Selle, I. Skalidis, O. Sorokin, D. Ubbiali, B. Voigt, A. Wedler, A. A. J. Wei, P. Zorn, A. K. Dunker, M. Köhn, A. Sinz, V. N. Uversky, *Cell. Mol. Life Sci.* **2022**, *79*, 449.
- [14] V. N. Uversky, *Protein Sci.* **2002**, *11*, 739–756.
- [15] F. Meng, V. N. Uversky, L. Kurgan, *Cell. Mol. Life Sci.* **2017**, *74*, 3069–3090.
- [16] B. He, K. Wang, Y. Liu, B. Xue, V. N. Uversky, A. K. Dunker, *Cell Res.* **2009**, *19*, 929–949.
- [17] P. R. Romero, S. Zaidi, Y. Y. Fang, V. N. Uversky, P. Radivojac, C. J. Oldfield, M. S. Cortese, M. Sickmeier, T. LeGall, Z. Obradovic, A. K. Dunker, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8390–8395.
- [18] H. Xie, S. Vucetic, L. M. Iakoucheva, C. J. Oldfield, A. K. Dunker, Z. Obradovic, V. N. Uversky, *J. Proteome Res.* **2007**, *6*, 1917–1932.
- [19] P. Tompa, *Trends Biochem. Sci.* **2012**, *37*, 509–516.
- [20] R. Linding, L. J. Jensen, F. Diella, P. Bork, T. J. Gibson, R. B. Russell, *Structure* **2003**, *11*, 1453–1459.
- [21] K. M. Ruff, R. V. Pappu, *J. Mol. Biol.* **2021**, *433*, DOI 10.1016/j.jmb.2021.167208.
- [22] F. Ferron, S. Longhi, B. Canard, D. Karlin, *Proteins Struct. Funct. Genet.* **2006**, *65*, 1–14.
- [23] A. Sinz, C. Arlt, D. Chorev, M. Sharon, *Protein Sci.* **2015**, *24*, 1193–1209.

- [24] M. Kachala, E. Valentini, D. I. Svergun, *Adv. Exp. Med. Biol.* **2015**, *870*, 261–289.
- [25] G. Mitra, *Proteomics* **2021**, *21*, DOI 10.1002/pmic.202000011.
- [26] P. A. Faull, H. V. Florance, C. Q. Schmidt, N. Tomczyk, P. N. Barlow, T. R. Hupp, P. V. Nikolova, P. E. Barran, *Int. J. Mass Spectrom.* **2010**, *298*, 99–110.
- [27] M. A. Ziemer, A. Mason, D. M. Carlson, *J. Biol. Chem.* **1982**, *257*, 11176–11180.
- [28] S. J. Baker, E. R. Fearon, J. M. Nigro, S. R. Hamilton, A. C. Preisinger, J. M. Jessup, P. Vantuinen, D. H. Ledbetter, D. F. Barker, Y. Nakamura, R. White, B. Vogelstein, *Science (80-)*. **1989**, *244*, DOI 10.1126/science.2649981.
- [29] J. Bargonetti, C. Prives, *J. Mol. Cell Biol.* **2019**, *11*, 605–609.
- [30] R. Kamada, Y. Toguchi, T. Nomura, T. Imagawa, K. Sakaguchi, *Biopolymers* **2016**, 598–612.
- [31] T. Ozaki, A. Nakagawara, *Cancers (Basel)*. **2011**, *3*, 994–1013.
- [32] L. Carlsen, W. S. El-Deiry, *Int. J. Mol. Sci.* **2021**, *22*, DOI 10.3390/ijms22211828.
- [33] C. Chao, S. Saito, C. W. Anderson, E. Appella, Y. Xu, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 11936–11941.
- [34] E. De Stanchina, M. E. McCurrach, F. Zindy, S. Y. Shieh, G. Ferbeyre, A. V. Samuelson, C. Prives, M. F. Roussel, C. J. Sherr, S. W. Lowe, *Genes Dev.* **1998**, *12*, 2434–2442.
- [35] Y. Liu, O. Tavana, W. Gu, *J. Mol. Cell Biol.* **2019**, *11*, 564–577.
- [36] V. N. Uversky, *Int. J. Mol. Sci.* **2016**, *17*, 1874–1910.
- [37] L. M. M. Jenkins, S. R. Durell, S. J. Mazur, E. Appella, *Carcinogenesis* **2012**, *33*, 1441–1449.
- [38] L. Smeenk, S. J. van Heeringen, M. Koepfel, B. Gilbert, E. Janssen-Megens, H. G. Stunnenberg, M. Lohrum, *PLoS One* **2011**, *6*, DOI 10.1371/journal.pone.0017574.
- [39] A. C. Joerger, A. R. Fersht, *Annu. Rev. Biochem.* **2008**, *77*, 557–582.
- [40] D. P. Teufel, S. M. Freund, M. Bycroft, A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7009–7014.
- [41] C. W. Lee, J. C. Ferreon, A. C. M. Ferreon, M. Arai, P. E. Wright, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 19290–19295.
- [42] E. A. Slee, D. J. O'Connor, X. Lu, *Oncogene* **2004**, *23*, 2809–2818.
- [43] K. K. Walker, A. J. Levine, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15335–15340.
- [44] W. S. El-Deiry, S. E. Kern, J. A. Pietenpol, K. W. Kinzler, B. Vogelstein, *Nat. Genet.* **1992**, *1*, 42–49.
- [45] M. Olivier, M. Hollstein, P. Hainaut, *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001008.
- [46] A. C. Joerger, A. R. Fersht, *Oncogene* **2007**, *26*, 2226–2242.
- [47] A. C. Joerger, H. C. Ang, A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15056–15061.
- [48] N. P. Pavletich, K. A. Chambers, C. O. Pabo, *Genes Dev.* **1993**, *7*, 2556–2564.

- [49] S. Rajagopalan, F. Huang, A. R. Fersht, *Nucleic Acids Res.* **2011**, *39*, 2294–2303.
- [50] J. Gencel-Augusto, G. Lozano, *Genes Dev.* **2020**, *34*, 1128–1146.
- [51] J. Ahn, C. Prives, *Nat. Struct. Biol.* **2001**, *8*, 730–732.
- [52] H. Kim, K. Kim, J. Choi, K. Heo, H. J. Baek, R. G. Roeder, W. An, *J. Mol. Biol.* **2012**, *415*, 843–854.
- [53] M. E. Anderson, B. Woelker, M. Reed, P. Wang, P. Tegtmeyer, *Mol. Cell. Biol.* **1997**, DOI 10.1128/mcb.17.11.6255.
- [54] A. K. Dunker, M. S. Cortese, P. Romero, L. M. Iakoucheva, V. N. Uversky, *FEBS J.* **2005**, *272*, 5129–5148.
- [55] C. J. Oldfield, J. Meng, J. Y. Yang, M. Q. Qu, V. N. Uversky, A. K. Dunker, *BMC Genomics* **2008**, *9*, DOI 10.1186/1471-2164-9-S1-S1.
- [56] B. W. Moore, *Biochem. Biophys. Res. Commun.* **1965**, *19*, DOI 10.1016/0006-291X(65)90320-7.
- [57] D. B. Zimmer, E. H. Cornwall, A. Landar, W. Song, *Brain Res. Bull.* **1995**, *37*, 417–429.
- [58] F. Michetti, V. Corvino, M. C. Geloso, W. Lattanzi, C. Bernardini, L. Serpero, D. Gazzolo, *J. Neurochem.* **2012**, *120*, DOI 10.1111/j.1471-4159.2011.07612.x.
- [59] R. Donato, B. R. Cannon, G. Sorci, F. Riuzzi, K. Hsu, D. J. Weber, C. L. Geczy, *Curr. Mol. Med.* **2012**, *13*, 24–57.
- [60] B. C. M. Potts, J. Smith, M. Akkeu, T. J. Macke, K. Okazaki, H. Hidaka, D. A. Case, W. J. Chazin, *Nat. Struct. Biol.* **1995**, *2*, 790–796.
- [61] P. M. Kilby, L. J. Van Eldik, G. C. K. Roberts, *Structure* **1996**, *4*, 1041–1052.
- [62] S. P. Smith, G. S. Shaw, *Structure* **1998**, *6*, 211–222.
- [63] H. Chen, C. Xu, Q. Jin, Z. Liu, *Am. J. Cancer Res.* **2014**, *4*, 89–115.
- [64] C. W. Heizmann, G. Fritz, B. W. Schäfer, *Front. Biosci.* **2002**, *7*, DOI 10.2741/a846.
- [65] K. G. Hartman, P. T. Wilder, K. Varney, A. D. MacKerell, A. Coop, D. Zimmer, R. Lapidus, D. J. Weber, in *Melanoma - From Early Detect. to Treat.*, **2013**.
- [66] T. BECKER, V. GERKE, E. KUBE, K. WEBER, *Eur. J. Biochem.* **1992**, *207*, 541–547.
- [67] R. R. Rustandi, D. M. Baldisseri, D. J. Weber, *Nat. Struct. Biol.* **2000**, *7*, 570–574.
- [68] D. B. Zimmer, J. G. Dubuisson, *Cell Calcium* **1993**, *14*, 323–332.
- [69] R. Donato, *Int. J. Biochem. Cell Biol.* **2001**, *33*, 637–668.
- [70] D. Foell, S. Seeliger, T. Vogl, H. G. Koch, H. Maschek, E. Harms, C. Sorg, J. Roth, *Thorax* **2003**, *58*, 613–617.
- [71] J. Markowitz, I. Chen, R. Gitti, D. M. Baldisseri, Y. Pan, R. Udan, F. Carrier, A. D. MacKerell, D. J. Weber, *J. Med. Chem.* **2004**, *47*, 5085–5093.
- [72] C. Scotto, J. C. Deloulme, D. Rousseau, E. Chambaz, J. Baudier, *Mol. Cell. Biol.* **1998**, *18*, 4272–4281.

- [73] J. Lin, M. Blake, C. Tang, D. Zimmer, R. R. Rustandi, D. J. Weber, F. Carrier, *J. Biol. Chem.* **2001**, *276*, 35037–35041.
- [74] M. R. Fernandez-Fernandez, T. J. Rutherford, A. R. Fersht, *Protein Sci.* **2008**, DOI 10.1110/ps.035527.108.
- [75] J. van Dieck, M. R. Fernandez-Fernandez, D. B. Veprintsev, A. R. Fersht, *J. Biol. Chem.* **2009**, *284*, 13804–13811.
- [76] C. Scotto, C. Delphin, J. C. Deloulme, J. Baudier, *Mol. Cell. Biol.* **1999**, DOI 10.1128/mcb.19.10.7168.
- [77] J. Lin, M. Blake, C. Tang, D. Zimmer, R. R. Rustandi, D. J. Weber, F. Carrier, *J. Biol. Chem.* **2001**, *276*, 35037–35041.
- [78] M. R. Fernandez-Fernandez, D. B. Veprintsev, A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4735–4740.
- [79] M. S. Cosgrove, K. Bever, J. L. Avalos, S. Muhammad, X. Zhang, C. Wolberger, *Biochemistry* **2006**, *45*, 7511–7521.
- [80] D. B. Kristensen, K. Imamura, Y. Miyamoto, K. Yoshizato, *Electrophoresis* **2000**, *21*, 430–439.
- [81] K. Biemann, *Annu. Rev. Biochem.* **1992**, *61*, 977–1010.
- [82] L. Piersimoni, P. L. Kastritis, C. Arlt, A. Sinz, *Chem. Rev.* **2022**, *122*, 7500–7531.
- [83] J. R. Yates 3rd, *Trends Genet.* **2000**, *16*.
- [84] B. Domon, R. Aebersold, *Science (80-.)*. **2006**, *312*, 212–217.
- [85] K. L. Vikse, J. Scott McIndoe, *J. Mass Spectrom.* **2018**, *53*, 1026–1034.
- [86] M. S. B. Munson, F. H. Field, *J. Am. Chem. Soc.* **1966**, *88*, 2621–2630.
- [87] A. C. Gingras, M. Gstaiger, B. Raught, R. Aebersold, *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 645–654.
- [88] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science (80-.)*. **1989**, *246*, 64–71.
- [89] M. Karas, F. Hillenkamp, *Anal. Chem.* **1988**, *60*, 2299–2301.
- [90] M. Wilm, *Mol. Cell. Proteomics* **2011**, *10*, DOI 10.1074/mcp.M111.009407.
- [91] J. H. Gross, *Mass Spectrometry: A Textbook*, Springer Cham, **2004**.
- [92] M. Dole, R. L. Hines, L. L. Mack, R. C. Mobley, L. D. Ferguson, M. B. Alice, *Macromolecules* **1968**, *1*, 96–97.
- [93] J. N. Smith, R. C. Flagan, J. L. Beauchamp, *J. Phys. Chem. A* **2002**, *106*, 9957–9967.
- [94] R. L. Grimm, J. L. Beauchamp, *J. Phys. Chem. A* **2010**, *114*, DOI 10.1021/jp907162w.
- [95] E. Aliyari, L. Konermann, *Anal. Chem.* **2020**, *92*, 10807–10814.
- [96] L. Konermann, *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 496–506.
- [97] P. Kebarle, U. H. Verkcerk, *Mass Spectrom. Rev.* **2009**, *28*, 898–917.

- [98] J. A. Loo, *Mass Spectrom. Rev.* **1997**, *16*, 1–23.
- [99] L. Konermann, E. Ahadi, A. D. Rodriguez, S. Vahidi, *Anal. Chem.* **2013**, *85*, 2–9.
- [100] A. Schmidt, M. Karas, T. Dülcks, *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 492–500.
- [101] M. Wilm, M. Mann, *Anal. Chem.* **1996**, *68*, 1–8.
- [102] J. L. P. Benesch, B. T. Ruotolo, D. A. Simmons, C. V. Robinsons, *Chem. Rev.* **2007**, *107*, 3544–3567.
- [103] R. Juraschek, T. Dülcks, M. Karas, *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300–308.
- [104] K. H. Kingdon, *Phys. Rev.* **1923**, *21*, 408–418.
- [105] A. Makarov, *Anal. Chem.* **2000**, *72*, 1156–1162.
- [106] M. Scigelova, A. Makarov, in *Proteomics*, **2006**, pp. 16–21.
- [107] S. Eliuk, A. Makarov, *Annu. Rev. Anal. Chem.* **2015**, *8*, 61–80.
- [108] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, *Anal. Chem.* **2006**, *78*, 2113–2120.
- [109] M. Hardman, A. A. Makarov, *Anal. Chem.* **2003**, *75*, 1699–1705.
- [110] T. Dienes, S. J. Pastor, S. Schürch, J. R. Scott, J. Yao, S. Cui, C. L. Wilkins, *Mass Spectrom. Rev.* **1996**, *15*, 163–211.
- [111] Y. Zhang, S. B. Ficarro, S. Li, J. A. Marto, *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1425–1434.
- [112] T. Geiger, J. Cox, M. Mann, *Mol. Cell. Proteomics* **2010**, *9*, 2252–2261.
- [113] W. Paul, H. Steinwedel, *Zeitschrift für Naturforsch. - Sect. A J. Phys. Sci.* **1953**, *8*, 448–450.
- [114] J. H. Batey, *Vacuum* **2014**, *101*, 410–415.
- [115] D. J. Douglas, *Mass Spectrom. Rev.* **2009**, *28*, 937–960.
- [116] H. Budzikiewicz, *Anal. Bioanal. Chem.* **2005**, *381*, 1319–1320.
- [117] F. A. Mellon, in *Encycl. Food Sci. Nutr.*, **2003**.
- [118] K. Giles, S. D. Pringle, K. R. Worthington, D. Little, J. L. Wildgoose, R. H. Bateman, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401–2414.
- [119] L. Cheng, R. A. Fenton, T. Pisitkun, in *Recent Adv. Syst. Biol. Res.*, **2014**, pp. 89–107.
- [120] D. R. Allen, B. C. McWhinney, *Clin. Biochem. Rev.* **2019**, *40*, 135–146.
- [121] F. Lanucara, S. W. Holman, C. J. Gray, C. E. Eyers, *Nat. Chem.* **2014**, *6*, 281–294.
- [122] B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S. J. Hyung, C. V. Robinson, *Nat. Protoc.* **2008**, *3*, 1139–1152.
- [123] A. Yousef, S. Shrestha, L. A. Viehland, E. P. F. Lee, B. R. Gray, V. L. Ayles, T. G. Wright, W. H. Breckenridge, *J. Chem. Phys.* **2007**, *127*, DOI 10.1063/1.2774977.
- [124] T. Wyttenbach, M. T. Bowers, *Annu. Rev. Phys. Chem.* **2007**, *58*, 511–533.

- [125] H. W. Ellis, *Planet. Space Sci.* **1989**, *37*, 1023.
- [126] J. N. Dodds, E. S. Baker, *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2185–2195.
- [127] R. Cumeras, E. Figueras, C. E. Davis, J. I. Baumbach, I. Gràcia, *Analyst* **2015**, *140*, 1376–1390.
- [128] G. A. Harris, S. Graf, R. Knochenmuss, F. M. Fernández, *Analyst* **2012**, *137*, 3039–3044.
- [129] A. A. Shvartsburg, R. D. Smith, *Anal. Chem.* **2008**, *80*, 9689–9699.
- [130] I. A. Buryakov, E. V. Krylov, E. G. Nazarov, U. K. Rasulev, *Int. J. Mass Spectrom. Ion Process.* **1993**, *128*, 143–148.
- [131] F. A. Fernandez-Lima, D. A. Kaplan, M. A. Park, *Rev. Sci. Instrum.* **2011**, *82*, DOI 10.1063/1.3665933.
- [132] P. Dwivedi, P. Wu, S. J. Klopsch, G. J. Puzon, L. Xun, H. H. Hill, *Metabolomics* **2008**, *4*, 63–80.
- [133] C. Wu, W. F. Siems, G. R. Asbury, H. H. Hill, *Anal. Chem.* **1998**, *70*, 4929–4938.
- [134] L. Sleno, D. A. Volmer, *J. Mass Spectrom.* **2004**, *39*, DOI 10.1002/jms.703.
- [135] A. M. Frank, *J. Proteome Res.* **2009**, *8*, 2226–2240.
- [136] K. F. Medzihradszky, R. J. Chalkley, *Mass Spectrom. Rev.* **2015**, *34*, 43–63.
- [137] Y. Shen, N. Tolić, F. Xie, R. Zhao, S. O. Purvine, A. A. Schepmoes, R. J. Moore, G. A. Anderson, R. D. Smith, *J. Proteome Res.* **2011**, *10*, DOI 10.1021/pr200052c.
- [138] A. R. Johnson, E. E. Carlson, *Anal. Chem.* **2015**, *87*, 10668–10678.
- [139] C. Tu, J. Li, S. Shen, Q. Sheng, Y. Shyr, J. Qu, *PLoS One* **2016**, *11*, DOI 10.1371/journal.pone.0160160.
- [140] J. S. Brodbelt, *Anal. Chem.* **2016**, *88*, 30–51.
- [141] J. S. Beckman, V. G. Voinov, M. Hare, D. Sturgeon, Y. Vasil'Ev, D. Oppenheimer, J. B. Shaw, S. Wu, R. Glaskin, C. Klein, C. Schwarzer, G. Stafford, *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 2081–2091.
- [142] Y. Qi, D. A. Volmer, *Mass Spectrom. Rev.* **2017**, *36*, 4–15.
- [143] R. A. Zubarev, A. R. Zubarev, M. M. Savitski, *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 753–761.
- [144] P. Roepstorff, J. Fohlman, *Biol. Mass Spectrom.* **1984**, *11*, 601.
- [145] H. R. Morris, T. Paxton, A. Dell, J. Langhorne, M. Berg, R. S. Bordoli, J. Hoyes, R. H. Bateman, *Rapid Commun. Mass Spectrom.* **1996**, *10*, 889–896.
- [146] S. Tamara, M. A. Den Boer, A. J. R. Heck, *Chem. Rev.* **2022**, *122*, 7269–7326.
- [147] D. P. Demarque, A. E. M. Crotti, R. Vessecchi, J. L. C. Lopes, N. P. Lopes, *Nat. Prod. Rep.* **2016**, *33*, 432–455.
- [148] V. Popa, D. A. Trecroce, R. G. McAllister, L. Konermann, *J. Phys. Chem. B* **2016**, *120*, 5114–5124.
- [149] R. H. H. Van Den Heuvel, E. Van Duijn, H. Mazon, S. A. Synowsky, K. Lorenzen, C. Versluis, S. J.

- J. Brouns, D. Langridge, J. Van Der Oost, J. Hoyes, A. J. R. Heck, *Anal. Chem.* **2006**, *78*, DOI 10.1021/ac061039a.
- [150] J. C. Williamson, A. V. G. Edwards, T. Verano-Braga, V. Schwämmle, F. Kjeldsen, O. N. Jensen, M. R. Larsen, *Proteomics* **2016**, *16*, 907–914.
- [151] S. W. Wu, T. H. Pu, R. Viner, K. H. Khoo, *Anal. Chem.* **2014**, *86*, 5478–5486.
- [152] J. R. Yates, *F1000Research* **2019**, *8*, DOI 10.12688/f1000research.16987.1.
- [153] C. Ihling, D. Tänzler, S. Hagemann, A. Kehlen, S. Hüttelmaier, C. Arlt, A. Sinz, *J. Proteome Res.* **2020**, *19*, 4389–4392.
- [154] M. W. Senko, P. M. Remes, J. D. Canterbury, R. Mathur, Q. Song, S. M. Eliuk, C. Mullen, L. Earley, M. Hardman, J. D. Blethrow, H. Bui, A. Specht, O. Lange, E. Denisov, A. Makarov, S. Horning, V. Zabrouskov, *Anal. Chem.* **2013**, *85*, DOI 10.1021/ac403115c.
- [155] A. M. Haag, in *Adv. Exp. Med. Biol.*, **2016**, pp. 157–169.
- [156] E. de Hoffmann, *J. Mass Spectrom.* **1996**, *31*.
- [157] L. Anderson, C. L. Hunter, *Mol. Cell. Proteomics* **2006**, *5*, 573–588.
- [158] J. A. Silveira, M. E. Ridgeway, F. H. Laukien, M. Mann, M. A. Park, *Int. J. Mass Spectrom.* **2017**, *413*, 168–175.
- [159] J. A. Silveira, K. Michelmann, M. E. Ridgeway, M. A. Park, *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 585–595.
- [160] F. Meier, M. A. Park, M. Mann, *Mol. Cell. Proteomics* **2021**, *20*, DOI 10.1016/j.mcpro.2021.100138.
- [161] F. Meier, S. Beck, N. Grassl, M. Lubeck, M. A. Park, O. Raether, M. Mann, *J. Proteome Res.* **2015**, *14*, 5378–5387.
- [162] F. Meier, A. D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J. Decker, T. Kosinski, M. A. Park, N. Bache, O. Hoerning, J. Cox, O. Räther, M. Mann, *Mol. Cell. Proteomics* **2018**, *17*, 2534–2545.
- [163] A. C. Leney, A. J. R. Heck, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5–13.
- [164] F. Chen, B. Gülbakan, S. Weidmann, S. R. Fagerer, A. J. Ibáñez, R. Zenobi, *Mass Spectrom. Rev.* **2016**, *35*, 48–70.
- [165] N. P. Barrera, C. V. Robinson, *Annu. Rev. Biochem.* **2011**, *80*, 247–271.
- [166] F. Sobott, H. Hernández, M. G. McCammon, M. A. Tito, C. V. Robinson, *Anal. Chem.* **2002**, *74*, 1402–1407.
- [167] A. N. Verentchikov, W. Ens, K. G. Standing, *Anal. Chem.* **1994**, *66*, 126–133.
- [168] R. H. H. Van Den Heuvel, A. J. R. Heck, *Curr. Opin. Chem. Biol.* **2004**, *8*, 519–526.
- [169] J. Snijder, R. J. Rose, D. Veessler, J. E. Johnson, A. J. R. Heck, *Angew. Chemie - Int. Ed.* **2013**, *52*, 4020–4023.
- [170] A. Dyachenko, G. Wang, M. Belov, A. Makarov, R. N. De Jong, E. T. J. Van Den Bremer, P. W. H. I. Parren, A. J. R. Heck, *Anal. Chem.* **2015**, *87*, 6095–6102.

- [171] C. Arlt, V. Flegler, C. H. Ihling, M. Schäfer, I. Thondorf, A. Sinz, *Angew. Chemie - Int. Ed.* **2017**, *56*, 275–279.
- [172] F. J. O'Reilly, J. Rappsilber, *Nat. Struct. Mol. Biol.* **2018**, *25*, 1000–1008.
- [173] A. Sinz, *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
- [174] P. Novak, G. H. Kruppa, M. M. Young, J. Schoeniger, *J. Mass Spectrom.* **2004**, *39*, 322–328.
- [175] M. Faini, F. Stengel, R. Aebersold, *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 966–974.
- [176] R. M. Kaake, X. Wang, A. Burke, C. Yu, W. Kandur, Y. Yang, E. J. Novtisky, T. Second, J. Duan, A. Kao, S. Guan, D. Vellucci, S. D. Rychnovsky, L. Huang, *Mol. Cell. Proteomics* **2014**, *13*, 3533–3543.
- [177] C. Yu, A. Huszagh, R. Viner, E. J. Novitsky, S. D. Rychnovsky, L. Huang, *Anal. Chem.* **2016**, *88*, 10301–10308.
- [178] B. Arora, R. Tandon, P. Attri, R. Bhatia, *Curr. Protein Pept. Sci.* **2016**, *18*, DOI 10.2174/1389203717666160724202806.
- [179] P. K. Mishra, C. M. Yoo, E. Hong, H. W. Rhee, *ChemBioChem* **2020**, *21*, 924–932.
- [180] M. Stadlmeier, L. S. Runtsch, F. Streshnev, M. Wühr, T. Carell, *ChemBioChem* **2020**, *21*, 103–107.
- [181] L. Piersimoni, A. Sinz, *Anal. Bioanal. Chem.* **2020**, *412*, 5981–5987.
- [182] S. E. Hagen, K. Liu, Y. Jin, L. Piersimoni, P. C. Andrews, H. D. Showalter, *Org. Biomol. Chem.* **2018**, *16*, 6867–6870.
- [183] X. Tang, J. E. Bruce, *Mol. Biosyst.* **2010**, *6*, 939–947.
- [184] A. Sinz, *Anal. Bioanal. Chem.* **2017**, *409*, 33–44.
- [185] M. Matzinger, K. Mechtler, *J. Proteome Res.* **2021**, *20*, 78–93.
- [186] C. Iacobucci, M. Götze, C. H. Ihling, C. Piotrowski, C. Arlt, M. Schäfer, C. Hage, R. Schmidt, A. Sinz, *Nat. Protoc.* **2018**, *13*, 2864–2889.
- [187] M. Q. Müller, F. Dreiocker, C. H. Ihling, M. Schäfer, A. Sinz, *Anal. Chem.* **2010**, *82*, 6958–6968.
- [188] M. Götze, J. Pettelkau, S. Schaks, K. Bosse, C. H. Ihling, F. Krauth, R. Fritzsche, U. Kühn, A. Sinz, *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 76–87.
- [189] M. P. Wickramathilaka, B. Y. Tao, *J. Biol. Eng.* **2019**, *13*, DOI 10.1186/s13036-019-0191-2.
- [190] J. V. Staros, R. W. Wright, D. M. Swingle, *Anal. Biochem.* **1986**, *156*, 220–222.
- [191] Z. Grabarek, J. Gergely, *Anal. Biochem.* **1990**, *185*, 131–135.
- [192] A. V. West, G. Muncipinto, H. Y. Wu, A. C. Huang, M. T. Labenski, L. H. Jones, C. M. Woo, *J. Am. Chem. Soc.* **2021**, *143*, 6691–6700.
- [193] J. Brunner, H. Senn, F. M. Richards, *J. Biol. Chem.* **1980**, *255*, 3313–3318.
- [194] P. J. A. Weber, A. G. Beck-Sickinger, *J. Pept. Res.* **1997**, *49*, 375–383.

- [195] C. Iacobucci, M. Götze, C. Piotrowski, C. Arlt, A. Rehkamp, C. Ihling, C. Hage, A. Sinz, *Anal. Chem.* **2018**, *90*, 2805–2809.
- [196] M. Götze, J. Pettelkau, R. Fritzsche, C. H. Ihling, M. Schäfer, A. Sinz, *J. Am. Soc. Mass Spectrom.* **2014**, *26*, 83–97.
- [197] Z. Kukačka, M. Rosůlek, J. Jelínek, L. Slavata, D. Kavan, P. Novák, *J. Proteome Res.* **2021**, *20*, 2021–2027.
- [198] D. Ubbiali, M. Fratini, L. Piersimoni, C. H. Ihling, M. Kipping, I. Heilmann, C. Iacobucci, A. Sinz, *Angew. Chemie - Int. Ed.* **2022**, *61*, e202205726.
- [199] R. B. M. Schasfoort, *Handbook of Surface Plasmon Resonance*, **2008**.
- [200] C. M. Miyazaki, F. M. Shimizu, M. Ferreira, in *Nanocharacterization Tech.*, **2017**, pp. 183–200.
- [201] F. E. Ahmed, J. E. Wiley, D. A. Weidner, C. Bonnerup, H. Mota, *Cancer Genomics and Proteomics* **2010**, *7*, 303–310.
- [202] C. Arlt, Massenspektrometrische Untersuchungen Am Tumorsuppressorprotein P53, Martin-Luther-Universität Halle-Wittenberg, **2018**.
- [203] W. Schultze, Tetramerisierungs-Mutaten: P53 in Monomere Und Dimere Zwingen, Martin-Luther-Universität Halle-Wittenberg, **2021**.
- [204] M. H. Clare, Rekombinante Expression von S100 β Und Interaktionsstudien Mit Dem Tumorsuppressor P53, **2018**.
- [205] F. Otto, Optimierung Der Probenvorbereitung Für Die Native Massenspektrometrie von Proteinen, **2022**.
- [206] A. A. J. Wei, C. Iacobucci, W. Schultze, C. H. Ihling, C. Arlt, A. Sinz, *ChemBioChem* **2022**, *23*, DOI 10.1002/cbic.202100665.
- [207] R. R. Rustandi, A. C. Drohat, D. M. Baldisseri, P. T. Wilder, D. J. Weber, *Biochemistry* **1998**, DOI 10.1021/bi972701n.
- [208] A. N. Bullock, A. R. Fersht, *Nat. Rev. Cancer* **2001**, *1*, 68–76.
- [209] W. M. Chan, W. Y. Siu, A. Lau, R. Y. C. Poon, *Mol. Cell. Biol.* **2004**, *24*, 3536–3551.
- [210] R. Kamada, Y. Toguchi, T. Nomura, T. Imagawa, K. Sakaguchi, *Biopolymers* **2016**, *106*, 598–612.
- [211] J. Kamihara, H. Q. Rana, J. E. Garber, *Hum. Mutat.* **2014**, *35*, 654–662.
- [212] C. Arlt, C. H. Ihling, A. Sinz, *Proteomics* **2015**, *15*, 2746–2755.
- [213] A. Di Ianni, C. Tütung, M. Kipping, C. H. Ihling, J. Köppen, C. Iacobucci, C. Arlt, P. L. Kastiris, A. Sinz, *Sci. Rep.* **2023**, *13*.
- [214] J. Demand, J. Lüders, J. Höfeld, *Mol. Cell. Biol.* **1998**, *18*, 2023–2028.
- [215] G. Gaglia, Y. Guan, J. V. Shah, G. Lahav, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 15497–15501.
- [216] C. Scotto, C. Delphin, J. C. Deloulme, J. Baudier, *Mol. Cell. Biol.* **1999**, *19*, 7168–7180.
- [217] K. Lee, F. J. O'Reilly, *Essays Biochem.* **2023**, *67*, 215–228.

- [218] E. D. Merkley, J. R. Cort, J. N. Adkins, *J. Struct. Funct. Genomics* **2013**, *14*, 77–90.
- [219] A. L. Okorokov, M. B. Sherman, C. Plisson, V. Grinkevich, K. Sigmundsson, G. Selivanova, J. Milner, E. V. Orlova, *EMBO J.* **2006**, *25*, 5191–5200.
- [220] A. C. Joerger, A. R. Fersht, *Annu. Rev. Biochem.* **2008**, DOI 10.1146/annurev.biochem.77.060806.091238.
- [221] H. H. Nguyen, J. Park, S. Kang, M. Kim, *Sensors (Switzerland)* **2015**, *15*, 10481–10510.
- [222] E. Helmerhorst, D. J. Chandler, M. Nussio, C. D. Mamotte, *Clin. Biochem. Rev.* **2012**, *33*, 161–173.
- [223] A. Saponaro, *BIO-PROTOCOL* **2018**, *8*, e2957.
- [224] M. Plach, K. Grasser, T. Schubert, *BIO-PROTOCOL* **2017**, *7*, e2632.
- [225] P.-L. Hagedoorn, *Front. Catal.* **2022**, *2*, DOI 10.3389/fctls.2022.906668.
- [226] R. P. Sparks, R. Fratti, in *Methods Mol. Biol.*, **2019**, pp. 191–198.

Appendix

Amino Acid Sequences

Amino acid sequence of human p53

Wild-Type (tetrameric p53 variant)

MEEPQSDPSV EPPLSQETFSDLWKLLPENN VLSPLPSQAM DDLMLSPDDI
EQWFTEDPGP DEAPRMPEEAPPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ
KTYQGSYGFR LGFLHSGTAK SVTCTYSPAL NKMFCQLAKT CPVQLWVDST
PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE RCSDSDGLAP PQHLIRVEGN
LRVEYLDDRN TFRHSVVVPY EPPEVGS DCT TIHYNMCNS SCMGGMNRFP
ILTIITLEDSSGNLLGRNSF EVRVCACPGR DRRTEEENLR KKGEPPHHELP
PGSTKRALPN NTSSSPQPKK KPLDGEYFTL QIRGRERFEM FRELNEALEL
KDAQAGKEPG GSRAHSSHLK SKKGQSTSRH KKL MFKTEGP DSD

L344A (dimeric p53 variant)

MEEPQSDPSV EPPLSQETFSDLWKLLPENN VLSPLPSQAM DDLMLSPDDI
EQWFTEDPGP DEAPRMPEEAPPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ
KTYQGSYGFR LGFLHSGTAK SVTCTYSPAL NKMFCQLAKT CPVQLWVDST
PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE RCSDSDGLAP PQHLIRVEGN
LRVEYLDDRN TFRHSVVVPY EPPEVGS DCT TIHYNMCNS SCMGGMNRFP
ILTIITLEDSSGNLLGRNSF EVRVCACPGR DRRTEEENLR KKGEPPHHELP
PGSTKRALPN NTSSSPQPKK KPLDGEYFTL QIRGRERFEM FREANEALEL
KDAQAGKEPG GSRAHSSHLK SKKGQSTSRH KKL MFKTEGP DSD

L344P (monomeric p53 variant)

MEEPQSDPSV EPPLSQETFSDLWKLLPENN VLSPLPSQAM DDLMLSPDDI
EQWFTEDPGP DEAPRMPEEAPPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ
KTYQGSYGFR LGFLHSGTAK SVTCTYSPAL NKMFCQLAKT CPVQLWVDST

PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE RCSDSDGLAP PQHLIRVEGN
LRVEYLDDRN TFRHSVVVPY EPPEVGS DCT TIHYNMCNS SCMGGMNR RP
ILTIITLEDS SGNLLGRNSF EVRVCACPGR DRRTEENLR KKGEPPHHELP
PGSTKRALPN NTSSSPQPKK KPLDGEYFTL QIRGRERFEM FREPNEALEL
KDAQAGKEPG GSRAHSSHLK SKKGQSTSRH KKL MFKTEGP DSD

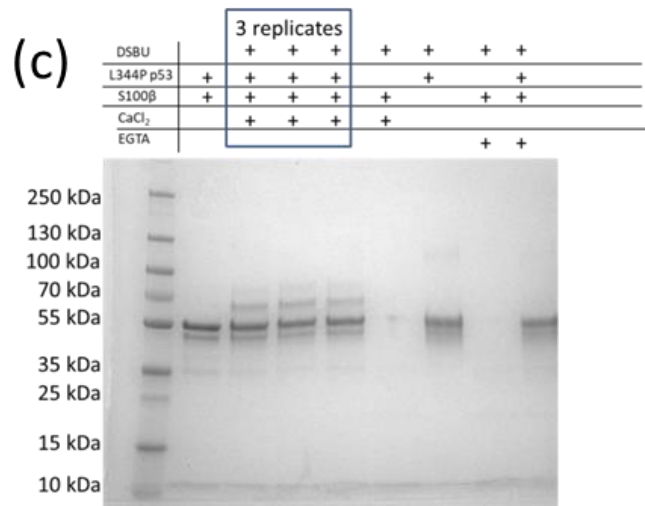
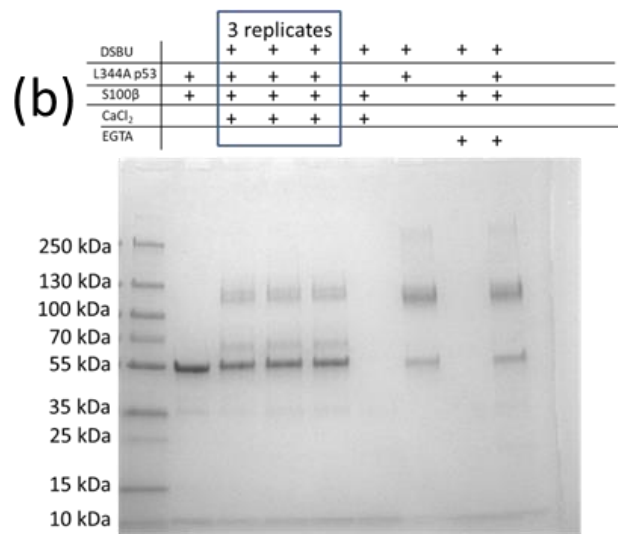
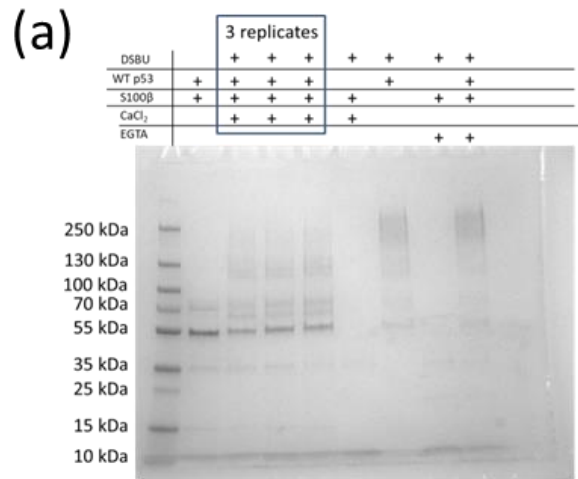
Amino acid sequence of S100 β

S100 β (self-produced)

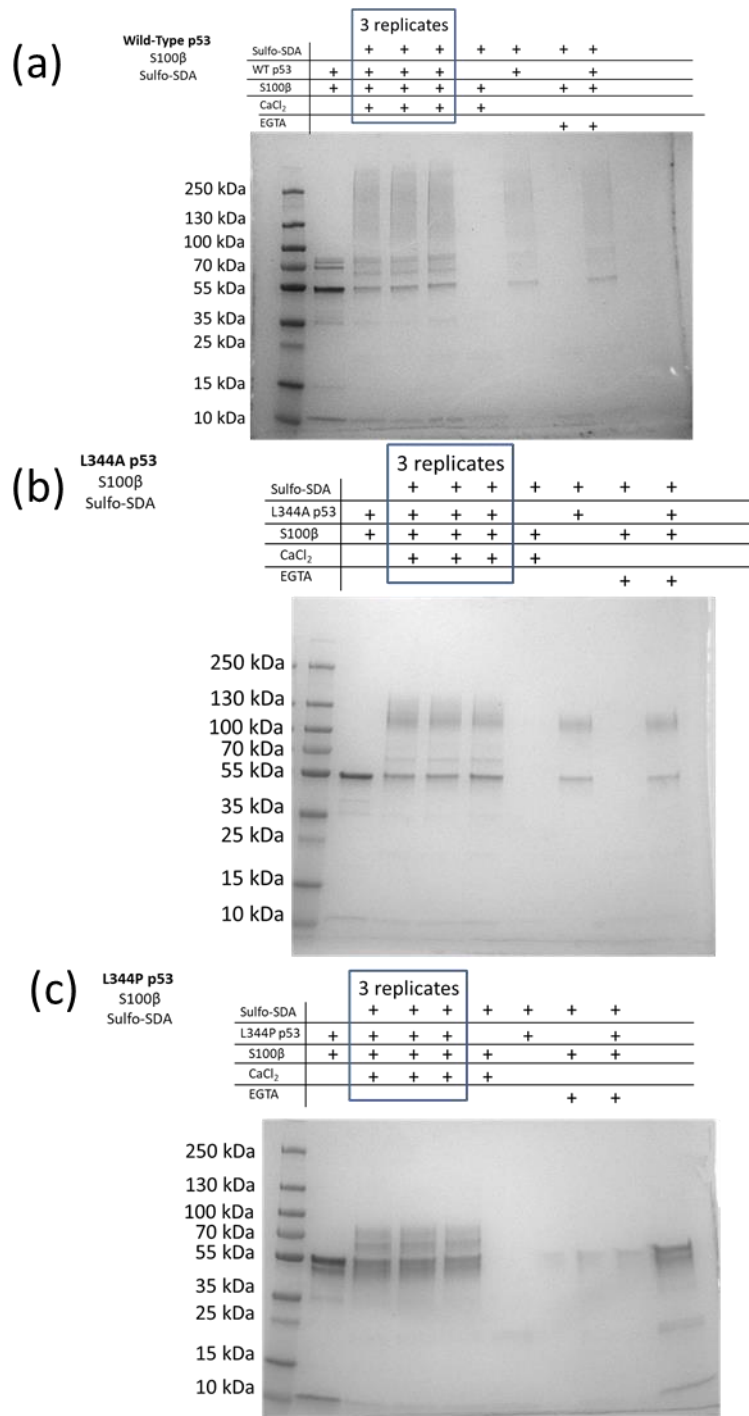
GS~~H~~MSELEKA MVALIDVFHQ YSGREGDKHK LKKSELKELI NNELSHFLEE
IKEQEVVDKV METLDNDGDG ECDFQEFMAF VAMVTTACHE FFEHE

S100 β (commercial)

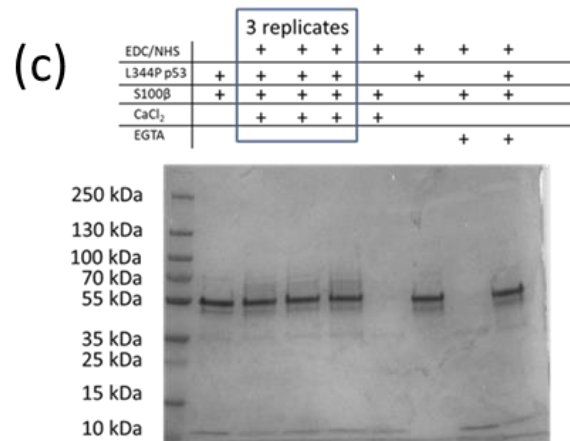
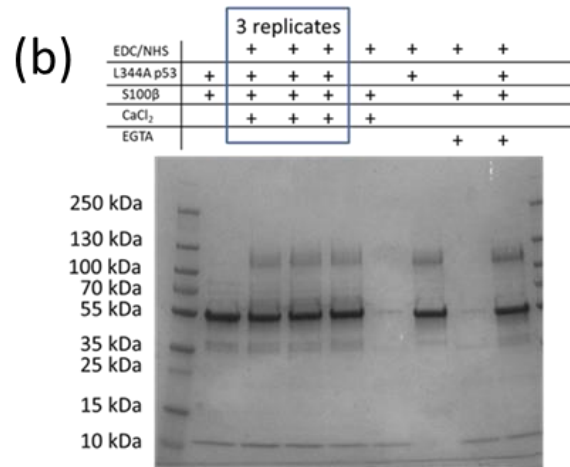
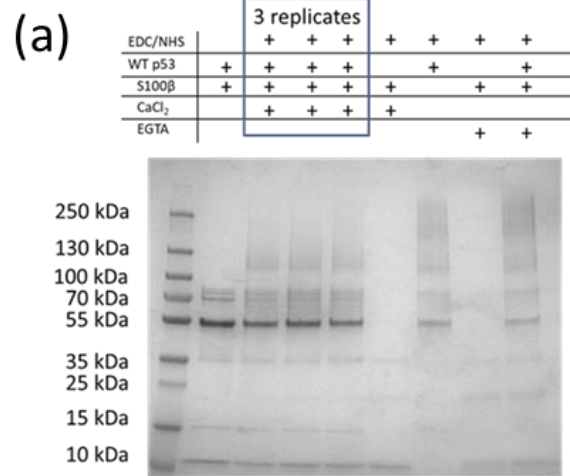
MSELEKAMVA LIDVFHQYSG REGDKHKLKK SELKELINNE LSHFLEEIKE
QEVVDKVMET LDNDGDGECDFQEFMAFVAM VTTACHEFFE HE



Supplementary Figure 1. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100β with DSBU (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.

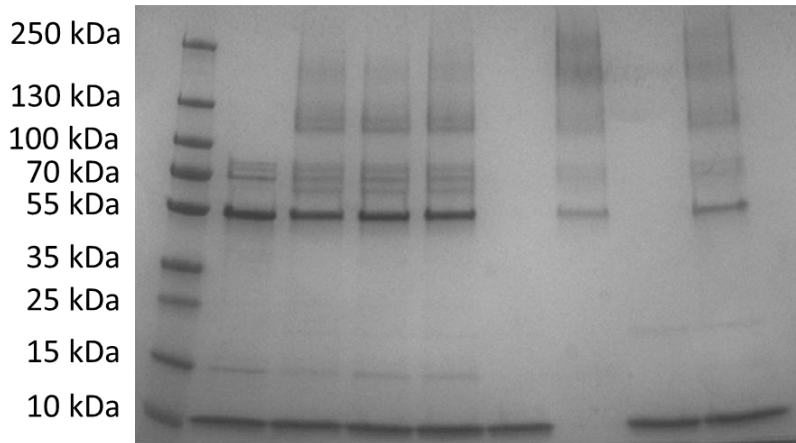


Supplementary Figure 2. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100β with sulfo-SDA (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.

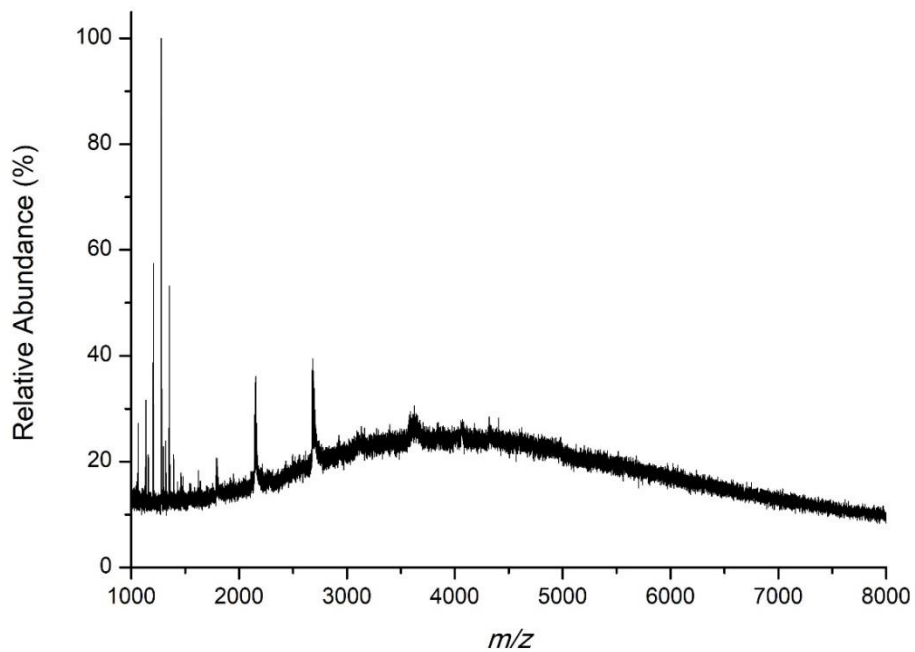


Supplementary Figure 3. SDS-PAGE (4-20% gradient) for cross-linking of p53-S100β with EDC/sulfo-NHS (a) wild-type p53 (b) dimeric L344A p53 (c) monomeric L344P p53.

		3 replicates					
DSBU		+	+	+	+	+	+
WT p53	+	+	+	+	+	+	+
S100 β	+	+	+	+	+	+	+
CaCl ₂		+	+	+	+		
EGTA						+	+



Supplementary Figure 4. SDS-PAGE (4-20% gradient) for cross-linking wild-type p53 and the commercial S100 β without the artificial residues (Gly-His-Ser).



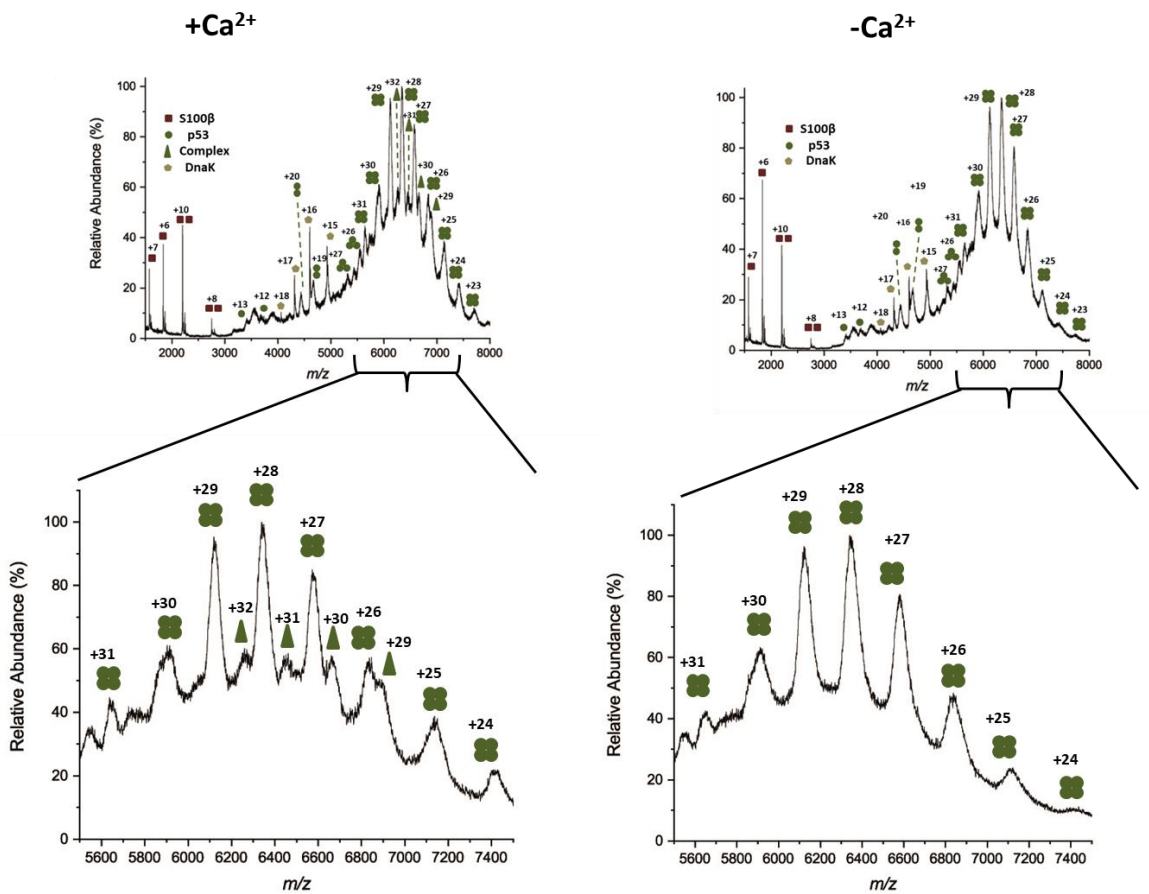
Supplementary Figure 5. Native mass spectrum of wild-type p53 and S100 β with no covalent fixation (no cross-linker added).

Wild-Type (tetrameric)

	m/z	
+32 ▲	6260.024	(200288 Da)
+31 ▲	6461.591	(200278 Da)
+30 ▲	6659.522	(199755 Da)
+29 ▲	6881.975	(199548 Da)

▲ ~200 kDa

(One p53 tetramer + one S100β dimer)



Supplementary Figure 6. ESI mass spectra showing the complex formation of S100β and wild-type tetrameric p53. Peak assignments of calcium-added and -depleted samples are presented.

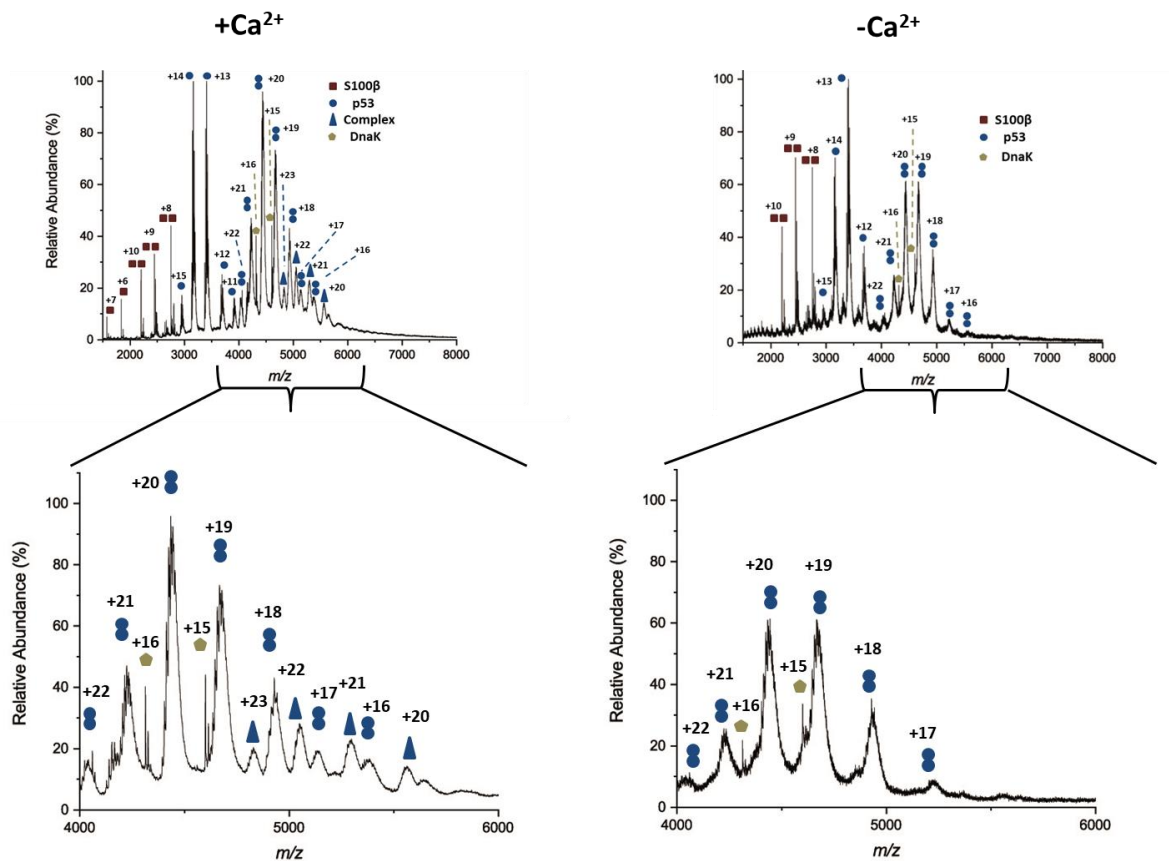
L344A (dimeric)

m/z

+23 ▲	4829.25	(111049 Da)
+22 ▲	5049.26	(111061 Da)
+21 ▲	5289.60	(111061 Da)
+20 ▲	5562.60	(111232 Da)

▲ (111 kDa)

(one p53 dimer + one S100B dimer)

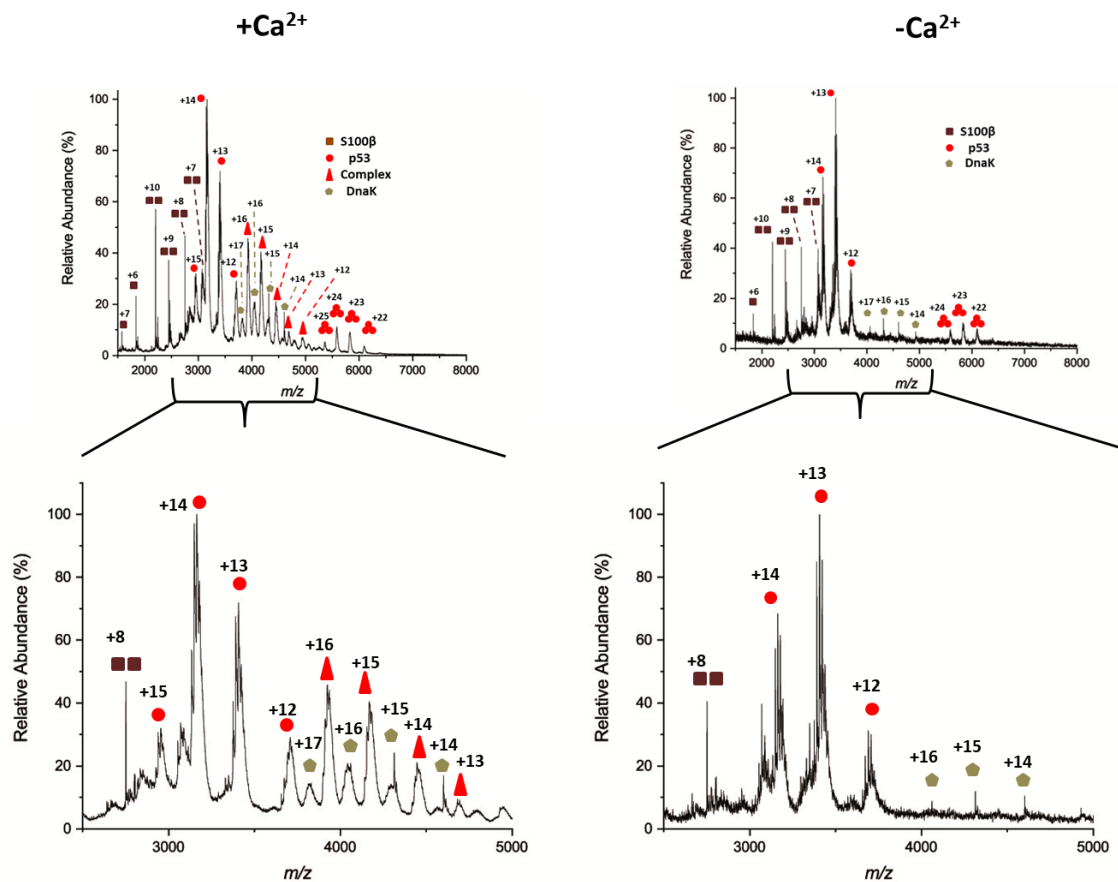


Supplementary Figure 7. ESI mass spectra showing the complex formation of S100β and L344A dimeric p53. Peak assignments of calcium-added and -depleted samples are presented.

L344P (monomeric)

	<i>m/z</i>	
+16	▲ 3922.58	(66666 Da)
+15	▲ 4169.30	(66692 Da)
+14	▲ 4446.41	(66681 Da)
+13	▲ 4693.71	(65697 Da)

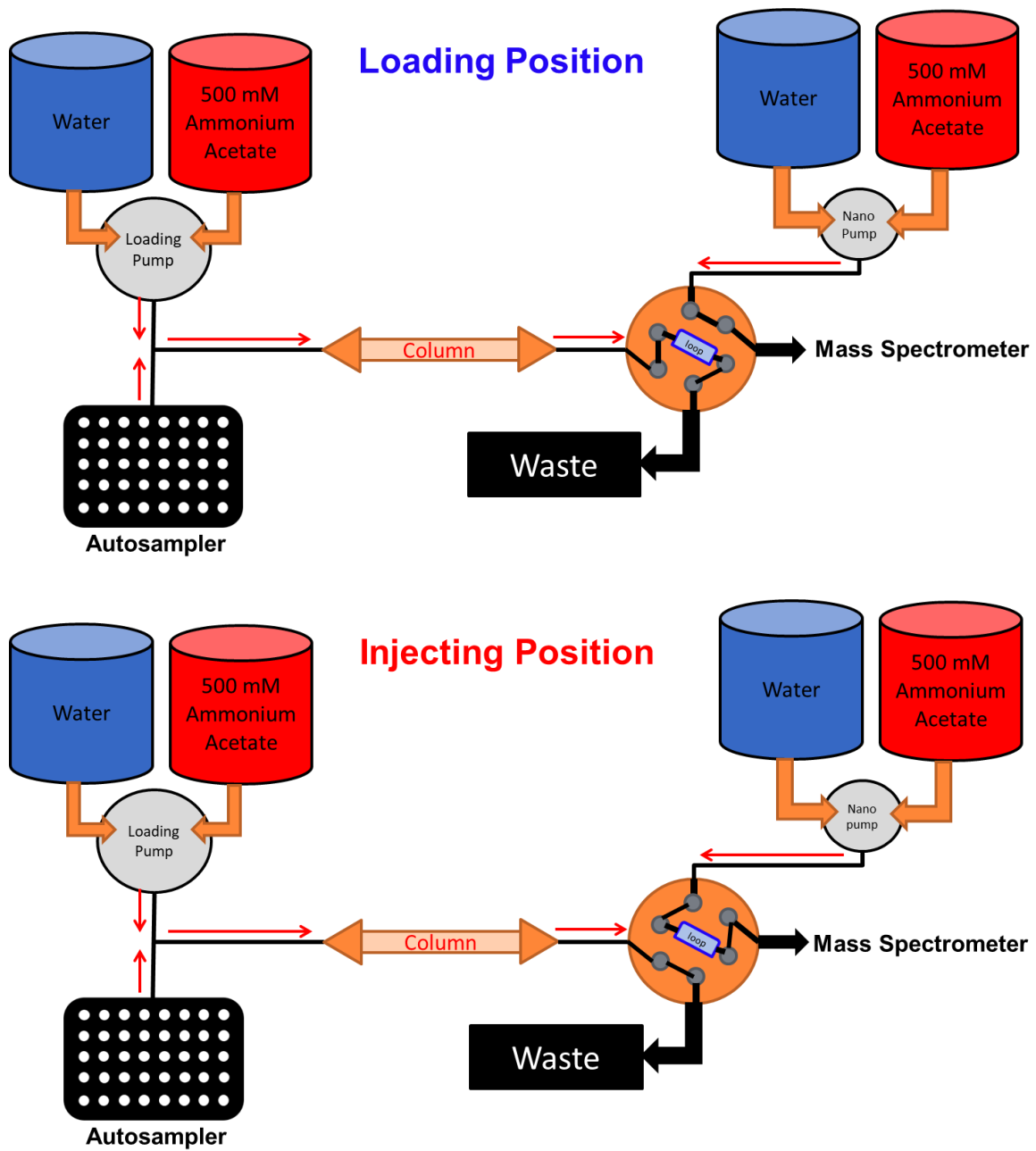
▲ 66 kDa
(One p53 monomer + one S100β dimer)



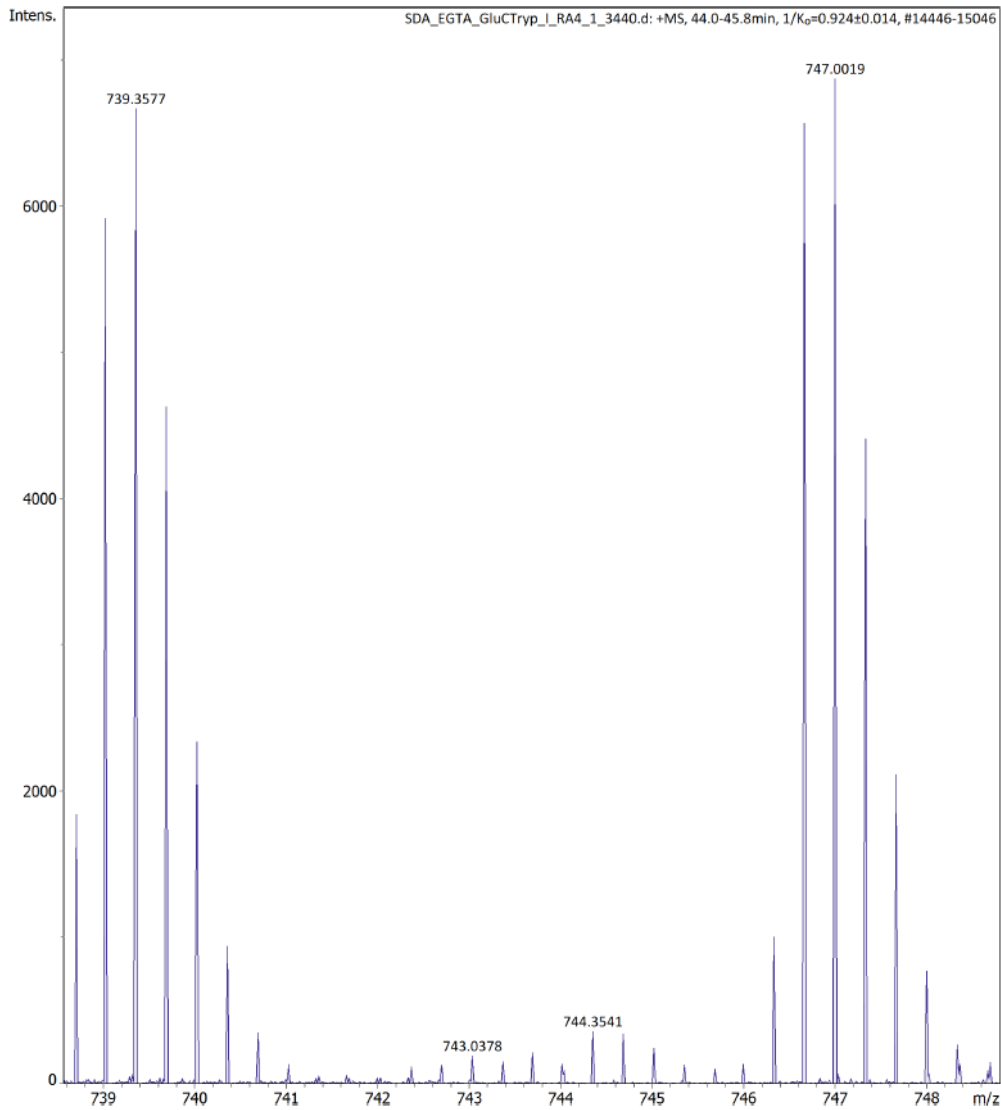
Supplementary Figure 8. ESI mass spectra showing the complex formation of S100β and L344P monomeric p53. Peak assignments of calcium-added and -depleted samples are presented.

Supplementary Table 1. Theoretical and experimental masses of p53 and complexes.

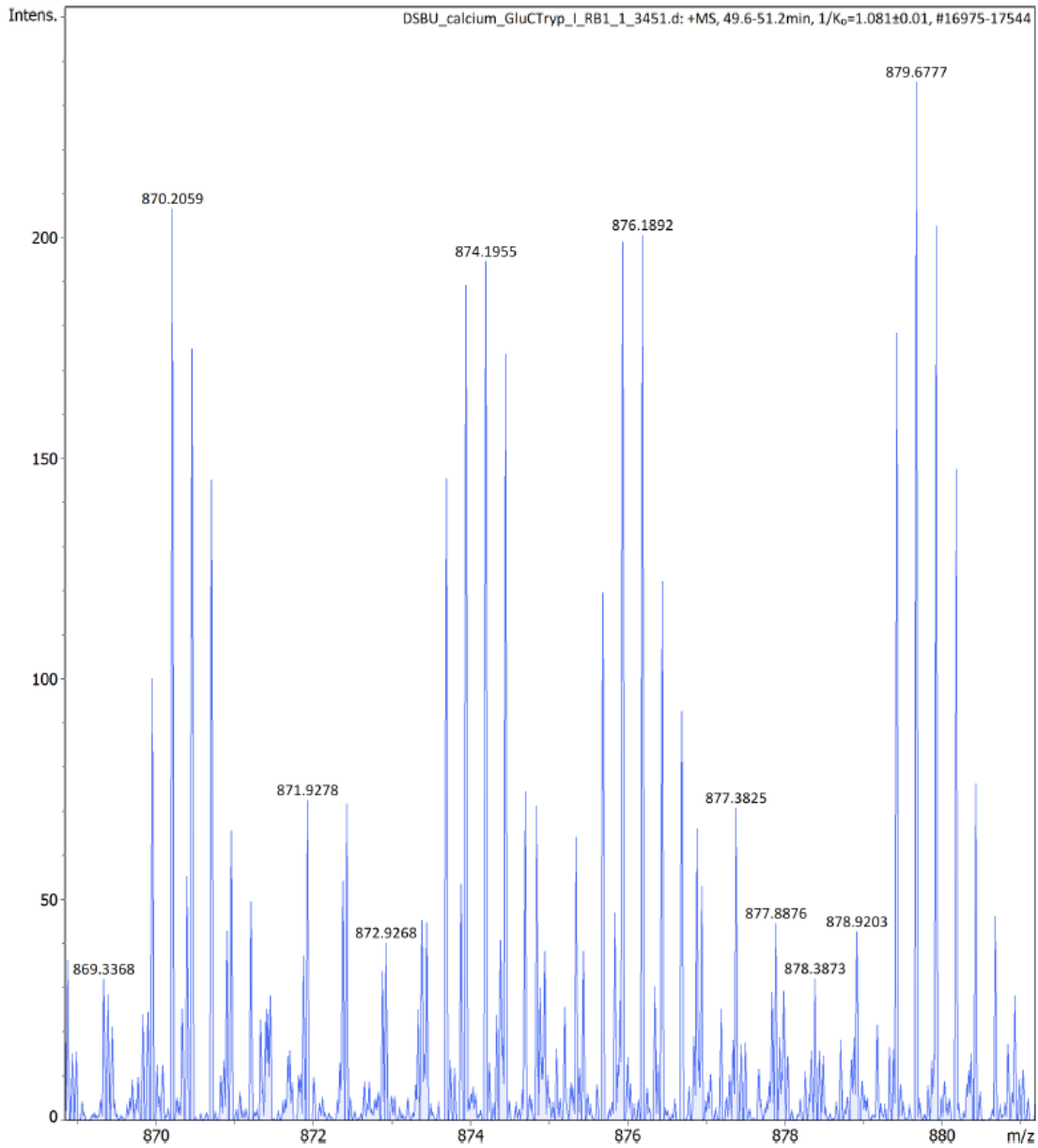
Species	Theoretical Mass (kDa)	Experimental Mass (kDa)
S100β monomer	10.994	11.062
S100β dimer	21.988	22.010
Tetrameric Wild-Type p53/S100β complex	196.948	199.967
Tetrameric Wild-Type p53	174.960	177.432
Dimeric L344A p53/S100β complex	109.468	111.100
Dimeric L344A p53 variant	87.480	88.215
Monomeric L344P p53/S100β complex	65.728	66.430
Monomeric L344P p53 variant	43.740	44.063



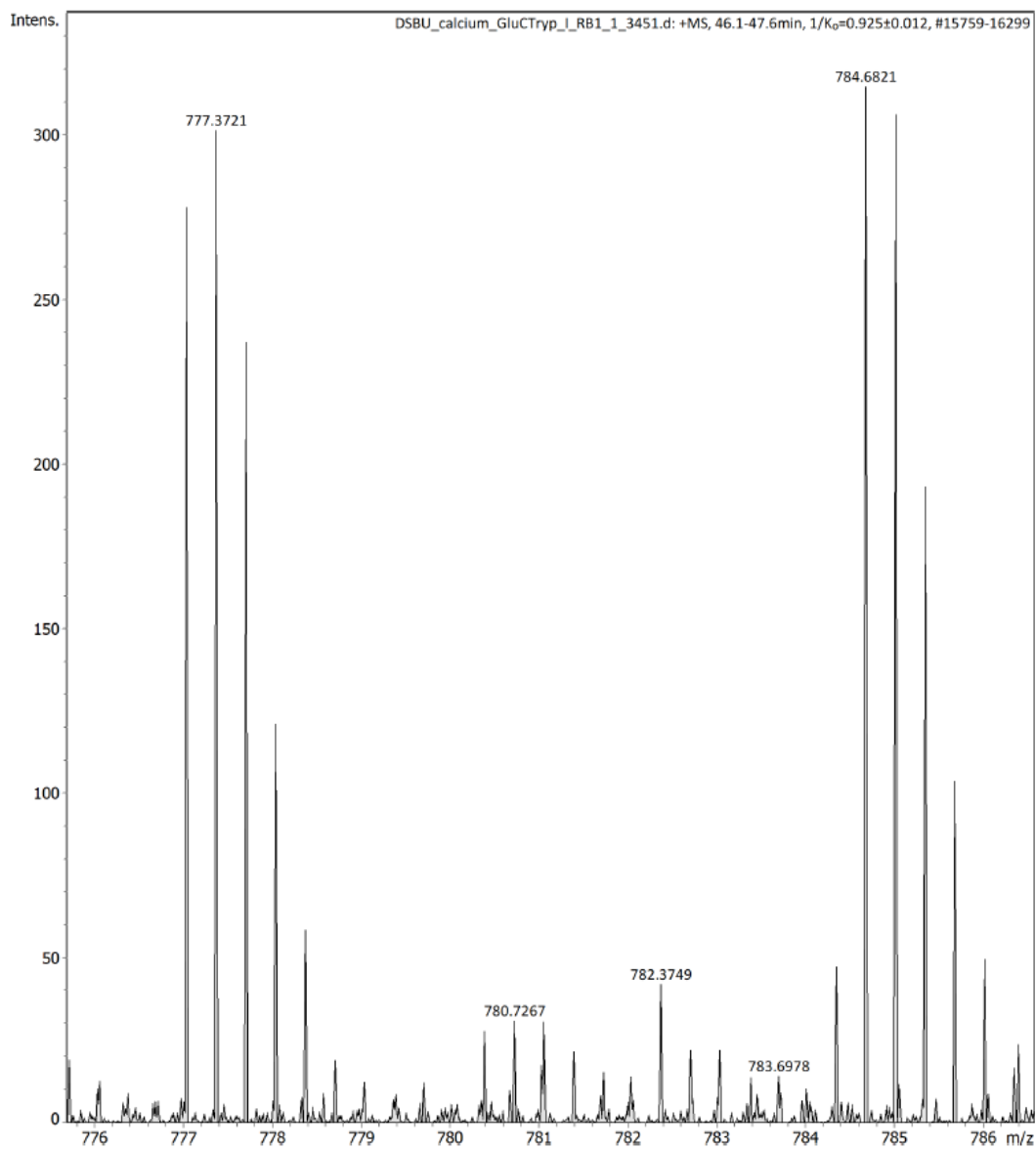
Supplementary Figure 9. Instrumental setup of the online buffer exchange (OBE) at load (top) and inject (bottom) position.



Supplementary Figure 10. Full-scan mass spectrum showing signals of the SDA cross-linked peptide between Nterm and E49 of ^{14}N and ^{15}N S100 β with a retention time 44.0-45.8 min, scan number 14446-15046.



Supplementary Figure 11. Full-scan mass spectrum showing signals of the DSBU cross-linked peptide between K9 and K59 of ¹⁴N and ¹⁵N S100β with a retention time 49.6-51.2 min, scan number 16975-17544.



Supplementary Figure 12. Full-scan mass spectrum showing signals of the DSBU cross-linked peptide between *N*-terminus and K52 of ¹⁴N and ¹⁵N S100β with a retention time 46.1-47.6 min, scan number 15759-16299.

Acknowledgement

Coming from a chemistry background, it took a bit of time for me to get used to working with proteins. There are so many people that I would like to express my gratitude for their patience, their constant guidance, and their emotional support.

I would like to thank my supervisor, Prof. Dr. Andrea Sinz, for giving me a chance to do my PhD thesis in her group. I am grateful for all the support and all the knowledge she has taught me throughout the years. I would like to thank the Research Training School GRK2467 for giving me a chance to learn a great deal about Intrinsically Disordered Proteins. I thank Jun. Prof. Dr. Carla Schmidt, Dr. Maria Ott, and Dr. Lolita Piersimoni, for giving me constructive feedback regarding my research. Thank you Claudia Spielmann for always keeping the RTG administratively-coordinated.

This thesis would not be possible without all the help from the group members. I would like to thank Dr. Christian Ihling and Dr. Marc Kipping for all the tutorials and knowledge on operating different mass spectrometers. I would like to thank Dr. Christian Arlt for all the knowledge about p53, from purification to how to perform experiments and proper handling. Thank you Dirk Tänzler for always being so patient and showing me how to optimize protein purification and how to handle things in the lab. I would like to thank Dr. Claudio Iacobucci for developing the COMPASS approach and his help on cross-linking analyses, moving the project forward, and really guiding me through the p53 project. I thank Dr. Vishnupriya Pandey and Frank Hause for carefully reading my thesis and giving me constructive critics on how I can improve. Thank you Wiebke Schultze, for preparing the monomeric and the dimeric p53 variants.

Very special thanks to my co-workers and my good friends, Alessio Di Ianni and Wiebke Günther. For the days where things get tough and I could not even crack a smile, thank you for being there for me. Thank you both for the huge amount of support during my PhD. Grazie mille fratellino. Ich möchte mich recht herzlich bedanken. Juan Camilo Rojas Echeverri, Lydia Kollhoff, Janett Köppen, Wiebke Schultze, and Florian Wolfgang Otto, thank you for making the lab a pleasant environment to work at. I love how we always crack different jokes and laugh at the same thing no matter how many times we have repeated, thank you for making the PhD experience joyful.

I would like to thank my family for their immense amount of support.

A-má, kám-siā lí tsū guá sè-hàn kàu tuā-hàn lóng it-tit kā guá kóo-lē. Guá it-tit lóng tsiann phah-piann. Guá tsai-iánn lí tī thinn-tíng it-tíng kám-kak tsiok huann-hí mā tsiok kong-îng ê. Tsiit phinn lūn-bûn sī kim-sun beh hiàn hōo lí ê.

謝謝爸爸和媽媽一路以來的支持。謝謝爸爸從博士一開始就給我許多建議，謝謝爸爸一直以來給我的鼓勵。謝謝媽媽一直以來的支持，鼓勵我。謝謝你們一直讓我知道你們一直以我為榮。

Thank you my dear sister Annie, for always listening to my complaints, always encouraging me, and always telling me how I am capable of getting my PhD.

Rafał, dziękuję bardzo. Dziękuję za to, że zawsze mnie wspierasz i motywujesz. Podobają mi się wszystkie nasze dyskusje na temat nauki i badań.

Publikationen und Tagungsbeiträge

Originalarbeiten

Piersimoni, L., Abd el Malek, M., Bhatia, T., Bender, J., Brankatschk, C., Calvo Sánchez, J., Dayhoff, GW., Di Ianni, A., Figueroa Parra, JO., Garcia-Martinez, D., Hesselbarth, J., Köppen, J., Lauth, LM., Lippik, L., Machner, L., Sachan, S., Schmidt, L., Selle, R., Skolidis, I., Sorokin, O., Ubbiali, D., Voigt, B., Wedler, A., **Wei, AAJ.**, Zorn, P., Dunker., AK., Köhn, M., Sinz, A., Uversky, NV. Lighting up Nobel Prize-winning studies with protein intrinsic disorder. *Cellular and Molecular Life Sciences* (2022)

Wei, AAJ., Iacobucci, I., Schultze, W., Ihling, CH., Arlt, C., Sinz, A. Different Oligomeric States of the Tumor Suppressor p53 Show Identical Binding Behavior towards the S100 β Homodimer. *ChemBioChem* (2022)

Wei, AAJ., Joshi, A., Chen, Y., A., McIndoe, JS. Strategies for avoiding saturation effects in ESI-MS. *International Journal of Mass Spectrometry* (2020)

Vorträge

“Structural Characterization of the Protein Interactions of Full-length Tumor Suppressor p53”

GRK2467 Retreat, Nimbschen, 2022

“Different Oligomeric States of the Tumor Suppressor p53 Show Identical Binding Behavior towards the S100 β Homodimer”

DGMS Young Scientists 2nd Fall Meeting, Hünfeld, 2022

“Structural Characterization of the Protein Interactions of Full-length Tumor Suppressor p53”

GRK2467 Retreat, Wittenberg, 2019

Posterpräsentationen

“Different Oligomeric States of the Tumor Suppressor p53 Show Identical Binding Behavior towards the S100 β Homodimer”

Intrinsically Disordered Proteins Gordon Research Conference, Les Diablerets, CH, 2022

“Structural Characterization of the Protein Interactions of Full-length Tumor Suppressor p53”

53. Jahrestagung der Deutschen Gesellschaft für Massenspektrometrie, Münster, 2020

“Structural Characterization of the Protein Interactions of Full-length Tumor Suppressor p53”

9th Symposium on Structural Proteomics, Göttingen, 2019

Lebenslauf

Persönliche Daten

Geburtsdatum, -ort 01.11.1995, Taoyuan (Taiwan)

Staatsangehörigkeit kanadisch, taiwanesisch

Schulbildung

09/2009 – 06/2013 **Elgin Park Secondary School**
Surrey, British Columbia, Canada

Hochschulausbildung

09/2013 – 06/2017 **Bachelor of Science, Chemistry**
McGill University, Montréal, Quebec, Canada

09/2017 – 04/2019 **Master of Science, Chemistry**
Thema: Mass Spectrometric Analysis of the Reactivity of the Trityl Cation (Prof. Dr. J. Scott McIndoe)
University of Victoria, Victoria, British Columbia, Canada

Promotion

09/2019 – 01/2023 **Wissenschaftlicher Mitarbeiter/Doktorand**
Thema: Structural Characterization of the Protein-Protein Interaction between the Tumor Suppressor p53 and S100 β
Abteilung für Pharmazeutische Chemie und Bioanalytik, Center for Structural Mass Spectrometry, Institut für Pharmazie, Martin-Luther-Universität Halle-Wittenberg (Prof. Dr. Andrea Sinz)

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Dissertation selbständig und ohne unerlaubte fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen genutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet. Bei der Auswahl und Auswertung des Materials, bei der Herstellung des Manuskripts, sowie bei der geistigen Herstellung der vorgelegten Arbeit waren keine anderen Personen beteiligt. Insbesondere wurde weder die Hilfe eines Promotionsberaters in Anspruch genommen, noch haben Dritte von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die vorgelegte Arbeit ist weder im Inland noch in Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt worden. Ich habe keine früheren erfolglosen Promotionsversuche unternommen.