universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Investigating the binding-site and possible long-range effects of the interaction between BRCA1 and MAX-MAX using NMR spectroscopy“

verfasst von / submitted by

Christian Manuel Kitzler BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019/ Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 834
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Molekulare Biologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Robert Konrat

Abstract

For a long time, protein function was associated with a single defined structure. As more and more proteins with important cell functions were discovered, lacking a single fold, the so-called intrinsically disordered proteins (IDPs) became an important field of study. These disordered proteins have in common to be flexible and to be able to adopt different conformations. Today, the main method for studying these proteins is by solution state NMR spectroscopy.

In this study, we are analysing the local geometry of the binding site of the heterodimer v-Myc-MAX to BRCA1, the product of the *breast cancer susceptibility gene 1*. To investigate this binding site, a small construct of BRCA1 is measured by NMR and a titration of a MAX-MAX homodimer to BRCA1 is performed, resulting in chemical shift changes in the BRCA1-spectrum. The use of the MAX-MAX homodimer is possible because it has been shown that v-Myc-MAX binds to BRCA1 in the same manner as MAX-MAX. We further want to find out if this binding has some influence on other parts of BRCA1. To study this, a large construct of BRCA1 was made. Due to the fact that BRCA1 is disordered in this region, making it prone to degradation, ligation of two smaller constructs was performed. For this, the enzyme Sortase A was used to fuse these two constructs together.

In our experiments we could narrow down the v-Myc-MAX binding site to only 20 amino acids. Additionally, ^{15}N-^{1}H correlation spectra of the two large constructs were made. Moreover, we were able to create a large construct of BRCA1 by the use of Sortase A.

Zusammenfassung

Lange Zeit wurde die Funktion von Proteinen primär durch eine definierte Struktur erklärt. Da immer mehr Proteine bekannt wurden, die wichtige Funktionen in Zellen haben, jedoch keine eindeutige Faltung besitzen, wurden die sogenannten intrinsisch ungeordneten Proteine(IDPs) zu einem wichtigen Forschungsfeld. Diese Proteine sind sehr flexibel und besitzen die Fähigkeit verschiedene Konformationen einzunehmen. Die wichtigste Methode um solche Proteine zu untersuchen ist die Kernspinresonanzspektroskopie (NMR).

In dieser Arbeit analysieren wir die lokale Geometrie der Bindungstelle des Heterodimers v-Myc-MAX zu BRCA1, dem Genprodukt des *breast cancer susceptibility gene 1*. Um diese Bindungsstelle zu untersuchen, wird ein kleines Konstrukt von BRCA1 mittels NMR gemessen und eine Titration des Homodimers MAX-MAX zu BRCA1 wird durchgeführt, was zu Änderungen der chemischen Verschiebung beim BRCA1 Spektrum führt. Die Verwendung des Homodimers MAX-MAX ist möglich, da gezeigt wurde, dass das Homodimer auf die gleiche Art bindet wie das v-Myc-MAX Heterodimer. Weiters wollten wir herausfinden, ob das Binden von MAX-MAX Auswirkungen auf andere Teile von BRCA1 hat. Dafür haben wir ein größeres BRCA1 Fragment hergestellt. Da dieser Teil von BRCA1 unstrukturiert ist, wodurch er anfällig für Abbau ist, wurde eine Ligation von zwei kleineren Konstrukten durchgeführt. Für die Fusion dieser zwei Teile wurde das Enzym Sortase A verwendet.

In unseren Experimenten konnten wir die v-Myc-MAX Bindungsstelle bis auf ca 20 Aminosäuren eingrenzen. Weiters konnten wir ^{15}N-^{1}H Korrelationsspektren von den Konstrukten aufnehmen. Außerdem war es uns möglich, durch die Verwendung von Sortase A, ein großes Konstrukt von BRCA1 herzustellen.

Contents

List of Figures

1 Introduction

1.1 Intrinsically disordered proteins

For a long time it was assumed that a defined structure is required for all proteins to carry out certain functions. Regions without a defined structure were believed to exist only in loops or to function as linkers. But already in 1978, proteins were discovered, whose functional (RNA-binding) domain could not be resolved with X-ray crystallography [1]. In the same year, it was shown by nuclear magnetic resonance spectroscopy (NMR) that glucagon, a small hormone, was partly unstructured in solution [2]. In the 1990s several disordered proteins, that folded upon binding, like p21 [3] and FlgM [4] were discovered. This showed that not all proteins needed a solid structure to be functional. Romero et. al. developed a tool to predict disorder in proteins [5], which showed that these proteins were a lot more common than previously thought. Finally, as more examples of unstructured proteins were discovered, the "Protein Trinity" was introduced in 2001. This model suggests three native states for proteins (structured, molten globule and random coil) [6]. Later, the terms intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) were commonly used, as they describe this state more precisely than the term random coil.

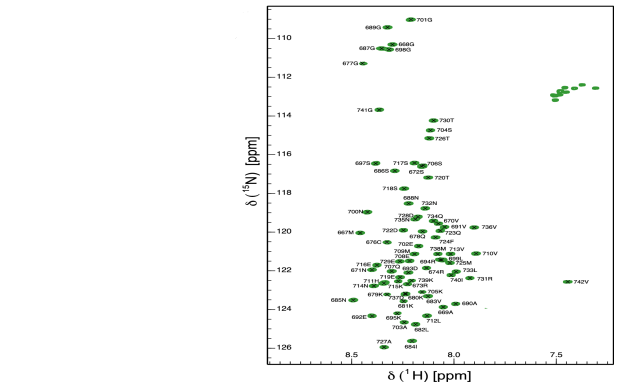
Disordered regions often have high variabilities in regions that are not involved in binding, but they are conserved at binding sites [7]. They possess a low average hydrophobicity, a high net charge [8] and they often have many binding partners. Additionally, it was shown by predictions as well as by experiments that phosphorylation of proteins occurs predominantly in IDRs [9] [10]. These features indicate their important role in many functions, such as cell signalling [11], DNA binding [12] and complex formation [13]. In eukaryotes, partly or completely unstructured proteins make up about 33 percent of proteins [14]. Despite their frequency of occurrence, today still relatively little is known about them.

1.2 NMR

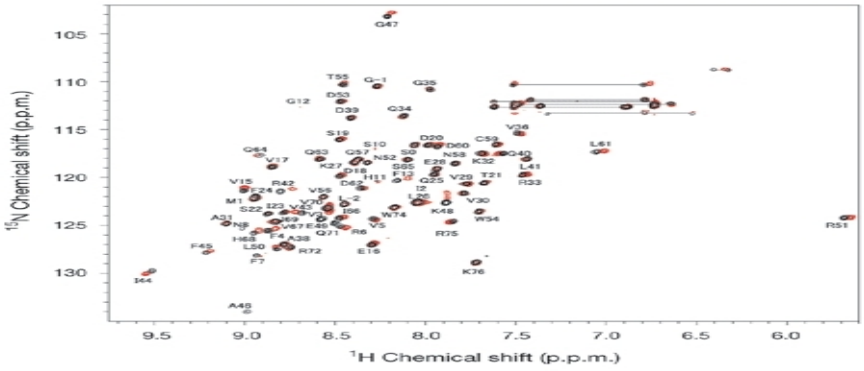
Nuclear magnetic resonance spectroscopy is a powerful tool to study molecules in solution. To be able to measure molecules with this method, their nuclei need to possess a magnetic moment. In organic molecules the most used isotopes possessing this feature are ^{1}H, ^{13}C and ^{15}N. When put into a strong magnetic field, the magnetic moment of nuclei partially orients itself along the magnetic field. Although the thermal motion of the molecules changes the orientation of the magnetic moments randomly, there is still a small net alignment of the spins resulting in a macroscopic z-magnetization. By the use of radio frequency (RF) pulses, the orientation of this bulk z-magnetization can be rotated by 90°or 180°. After a 90°pulse, the magnetization precesses around the z-direction of the magnetic field in the transverse (x-y)-plane, inducing a current in the detection coil. To change this signal, called the free induction decay (FID), from a function of time into a function of frequency, the Fourier transformation is used. The frequency of this precession depends on the Larmor frequency of a single spin, and therefore on the sort of nuclei and on the strength of the magnetic field. Consequently, a molecule measured on spectrometers of different field strength would produce different spectra. To standardize NMR spectra, the chemical shift scale is used. In this scale, the peak of an isotope is compared to a reference compound which is defined as zero. This chemical shift δ is given by $\delta(ppm) = 10^6 * \frac{v - v_{ref}}{v_{ref}}$ with v and v_{ref} being the frequencies of the measured isotope and the reference compound.

If the Larmor frequency depends only on the type of the isotope and the magnetic field strength, how can we distinguish between isotopes of the same type that are in different positions in a molecule? The solution to this are small local magnetic fields in the molecule that slightly change the total magnetic field experienced by the spin. Consequently, the chemical shifts of the same isotopes in different positions in a molecule differ from each other. The local fields are primarily produced by the local chemical environment, for example by electrons in the molecule and by neighbouring spins. In proteins, the chemical shifts are thus influenced by the type of amino acids,

the secondary structure, the tertiary structure and also by hydrogen-bonds between two binding partners. This leads to broader spectra in globular proteins than in IDPs as seen in Fig. 1.1.



(a)



(b)

Figure 1.1: Spectra of an unstructured vs. a structured protein: **(a):** ^{1}H-^{15}N spectrum of a part of CD44. The backbone amide peaks have ^{1}H chemical shifts ranging from 7.8 to 8.6 [16]. **(b):** ^{1}H-^{15}N spectrum of the ubiquitin like domain of parkin. The backbone amide peaks have ^{1}H chemical shifts ranging from 7.0 to 9.5 [17].

1.2.1 Relaxation and size dependence in NMR

If the bulk magnetization is in the transverse plane after an RF pulse, this transverse magnetization decreases over time. This happens both by

transversal and longitudinal relaxation. Longitudinal relaxation re-establishes the equilibrium z-magnetization in the same way as an RF pulse. In this mechanism, local magnetic fields oscillating at the Larmor frequency are responsible for changing the orientation of spins. Over time this leads to a return to the equilibrium state where there is no transverse magnetization, and the original z-magnetization is restored. [18]

As we can see, the transverse magnetization decreases by longitudinal relaxation. However, this can also happen without re-establishing the z-magnetization. If the transverse magnetization gets distributed equally in the x-y plane, there is no more bulk magnetization left and the signal is lost. This can be caused by the anisotropy of chemical shifts, meaning that the local field for each spin of the ensemble has a different orientation respective to the external magnetic field. Spins experiencing different magnetic fields have altered Larmor frequencies and therefore gradually get out of phase. This transverse relaxation process, also called dephasing, gets smaller for molecules that rotate faster, as the local fields get averaged. [18]

Both the transversal and the longitudinal relaxation, depend on the motion of the molecules. To describe the transversal relaxation it is sufficient to know the correlation time (τ_c), which is defined as the average time it takes for a molecule to change its orientation by one radians. The shorter the correlation time, the slower the transveral relaxation happens. [18]

The longitudinal relaxation occurs only if the frequency of the motion is at the Larmor frequency. To quantify the rotational motion, we can use the reduced correlation function c(t) as shown in Eq.1.1. Figure 1.2 (a) shows how fast the correlation decreases over time for molecules of different size. To get the amount of motion present at a given frequency, the reduced spectral density (Eq.1.2), which is the Fourier transform of the reduced correlation function, is used. The spectral density at the Larmor frequency is highest at a correlation time of $\frac{1}{\omega_0}$, which means that longitudinal relaxation is fastest for molecules with this frequency of motion. In Figure 1.2(b) it can be seen that for larger molecules the spectral density decreases with frequency leading to less longitudinal relaxation. [18]

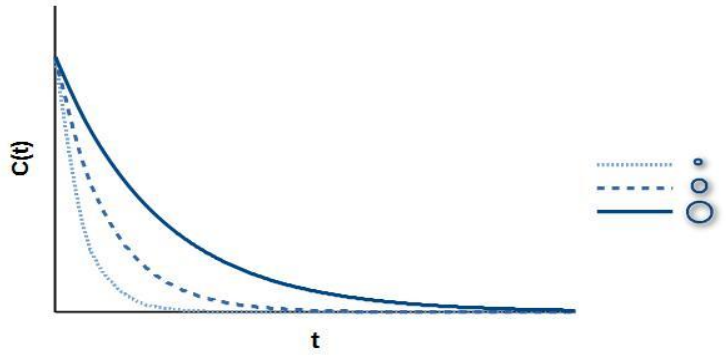
For proteins the transverse relaxation, which increases with size, is the

dominant relaxation mechanism. However, it has been shown that this doesn't apply to IDPs in the same way [19]. Their high flexibility allows for internal motion, resulting in correlation times of just a few ns [19]. That's why the signal isn't lost as fast as for globular proteins, allowing us to measure IDPs with much more residues.

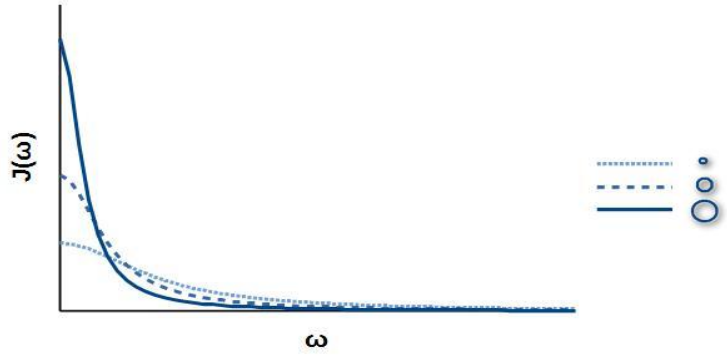
$$c(t) = \exp \frac{-t}{\tau_c} \tag{1.1}$$

$$j(\omega) = \frac{2\tau_c}{1 + \omega^2 \tau_c^2} \tag{1.2}$$

τ_c = correlation time
t = time
ω = frequency
c(t) = reduced correlation function
j(ω) = reduced spectral density



(a)



(b)

Figure 1.2: Correlation function and spectral density: **(a):** Correlation function for proteins of different size.**(b):** Spectral density for proteins of different size.

1.2.2 HSQC

Due to the numerous residues in proteins, most peaks cannot be linked to residues in the protein solely by their chemical shifts. To overcome this difficulty, the fact that magnetization can be transferred over scalar coupled (covalently bonded) atoms is used. This results in a series of signals that are modulated by the Larmor frequency of neighbouring atoms [15]. By connecting the atoms of the protein backbone, an assignment of the peaks becomes possible.

In the Heteronuclear Single Quantum Coherence, HSQC, this magnetization transfer can be used to create 2D ^{15}N-^{1}H spectra. During this HSQC, the magnetization is at first transferred from ^{1}H to ^{15}N. Then the N spin is allowed to evolve for a time t_1, obtaining another time domain. After that, the magnetization is transferred back to the ^{1}H. Now the FID is recorded and decoupling of ^{15}N is performed at the same time. Decoupling is used to suppress Scalar coupling, thus preventing peak doublets (or multiplets). If the peaks are already assigned, this pulse sequence is well suited to check for binding partners by observing the changes in chemical shifts for certain residues. For IDPs, that have narrow spectra and thus more overlap, as can be seen in Fig.3.5, higher dimensionality experiments have to be performed in order to resolve all the peaks. [18]

1.2.3 NOESY

The bulk magnetization can also be described by the population difference between the α and the β state. In equilibrium the α state is more populated, leading to net magnetization along the external magnetic field. Through longitudinal relaxation, the β state is switched back into the α state. For two spins that are close together in space, thus being in a two-spin system, 4 energy levels ($\alpha\alpha, \alpha\beta, \beta\alpha, \beta\beta$) exist. In equilibrium (Fig. 1.3 (a)) the most populated energy level is where both spins are in the α-state ($\alpha\alpha$). If now only spin 1 gets excited (Fig. 1.3 (b)), the system is out of balance, and tries to go back to equilibrium state by relaxation. This can happen by self relaxation (W_1), with spin 1 going back to the α-state, or by cross-relaxation, with both spins changing their state. The two possibilities for cross-relaxation are a double-quantum transition (W_2) or a zero-quantum transition (W_0). The cross-relaxation rate constant (σ_{12}), that describes this magnetization transfer is defined as $\sigma_{12} = W_2 - W_0$. In this process, the population of spin 2 is changed, although it wasn't excited. This effect is called the nuclear Overhauser effect (NOE). In the NOESY experiment, this effect is used to create H-H cross-peaks that show if hydrogen atoms are close to each other. For proton-proton interactions this effect can be seen up to a

distance of approximately 5Å . [18]

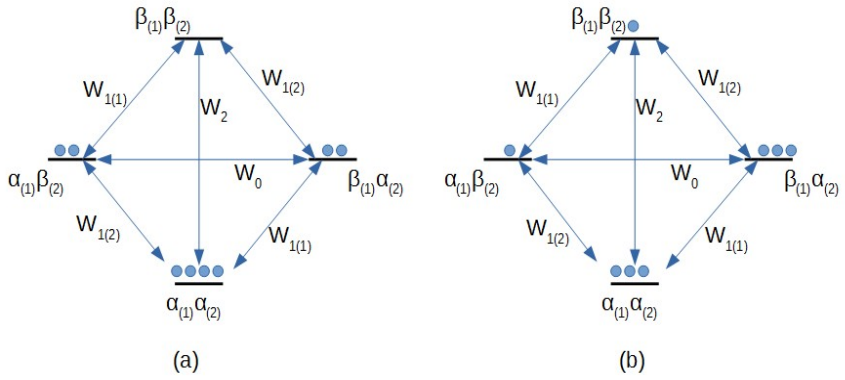


Figure 1.3: Energy levels of a two-spin system: **(a):** equilibrium state. **(b):** spin 1 got excited. $W_{1(1)}$ are single-quantum transitions of spin 1, W_0 is the zero-quantum transition and W_2 is a double quantum transition

1.3 BRCA1

BRCA1 is a large protein, consisting of 1863 amino acids. The gene for this protein (*breast cancer susceptibility gene 1*) lies on chromosome 17q21 and consists of 24 exons which span a region of 81kb [20] [21] [22]. The large number of exons results in many different isoforms of BRCA1 mRNA [23]. Together with its many binding partners, BRCA1 functions as an important tumor supressor, since a large percentage of hereditary breast and ovarian cancers are linked to mutations in BRCA1 [24]. BRCA1 also seems to play an important role in embryonic cells, as its mRNA levels show a clear transcription pattern in these cells. In addition, mRNA is also present in adult epithelial cells with an increased level of transcription in mammary cells during pregnancy [25]. Transcription and expression of BRCA1 are cell cycle dependent with a maximum abundance at the G1/S boundary [26] [27] indicating a major role in cell cycle control.

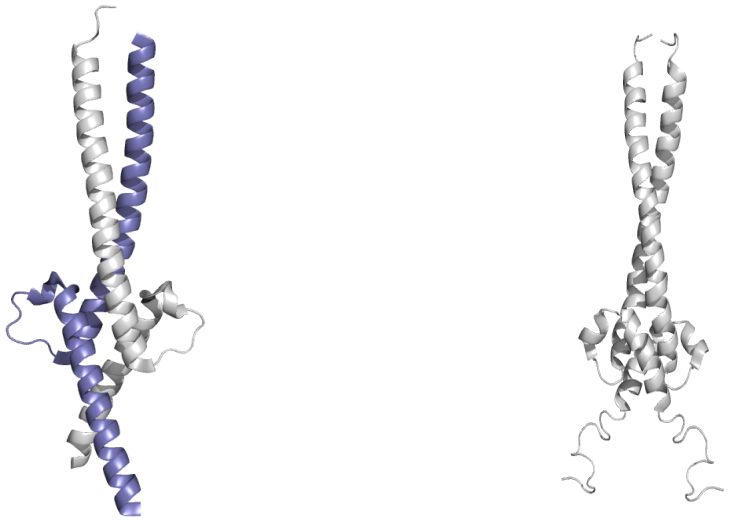
At the N-terminus, BRCA1 has a structured RING-domain, which interacts with BARD1 to mediate nuclear import [28] [29]. On the C-terminus there are two consecutive BRCT (BRCA1 C- terminal) domains. They form a 3-D structure, and bind many proteins, including histone deacetylases HDAC1 and HDAC2, and CtIP [30] [31]. Both structured regions have already been solved by X-ray crystallography [32] [33]. The unstructured part of BRCA1 ranges from position 104 to 1645. Despite the lack of structure and the low percentage of conserved residues, several protein binding sites in this region have been discovered [34]. Among those important binding partners are c-Myc [35], p53 [36], RB [37], Rad50 [38], Rad51 [39], FANCA [40], JunB [41] and BRCA2 [42]. Furthermore, many mutations in this region are linked to hereditary breast and ovarian cancer [34].

In the cell, BRCA1 has several different functions, which classify it as a tumor suppressor. One of these functions is binding to p53 and activation of c-Jun, triggering apoptosis [43] [44]. BRCA1 gets recruited to sites of double-strand breaks by NBA1 [45], and interacts with CCDC98 which acts as G2/M damage induced checkpoint control [46]. There it forms a large complex named BASC which may act as a DNA-damage sensor [47]. BRCA1 also interacts with γ-tubulin, mediating chromosome segregation [48]. These numerous functions demonstrate the importance of BRCA1 in cell cycle control, genome integrity, and DNA-damage repair. It has also been shown that BRCA1 regulates transcription by histone modification through interaction with histone deacetylases [30] and by binding to c-Myc, which is a transcription activator [35].

1.4 Myc-MAX

c-Myc is a proto-oncogene that forms a heterodimer with MAX [49]. The binding occurs at the basic helix-loop-helix leucine-zipper of both, MAX and c-Myc [50]. If present in too high quantities, for example by reduced turnover [51], cells may transform into tumor cells. c-Myc-MAX does so by acting as a transcription activator, when binding to DNA [52]. This DNA-

bound state has already been shown by X-ray structures of c-Myc-MAX [53]. There is also a viral version of c-Myc, which is called v-Myc. V-Myc is an oncogene, that is used by several retro viruses for transcription activation of their own proteins [54] [55] [56] [57] [58]. When it is inserted into the genome of the host, it leads to transformation into cancer cells [59]. NMR and circular dichroism studies of v-Myc show, that it stays a monomer if no MAX is present. [60].



(a) (b)

Figure 1.4: (a): Myc-MAX heterodimer bHLH motive (PDB structure 1NKP [53]; processed with pymol [61]) (b): MAX-MAX homodimer bHLH motive (PDB structure 1r05 [62]; processed with pymol [61]).

1.5 Protein ligation

Gram-positive bacteria use a mechanism called sorting to anchor proteins to the cell wall. The protein catalysing this process in Staphylococcus aureus is the transpeptidase Sortase A [63]. It has been shown that the recognition sequence for anchoring by Sortase A is LPXTG [64]. At the -COOH of

this threonine, a thioester-sortase intermediate is built, which releases the C-terminus of the protein [65] [66]. The protein is now transferred to a pentaglycine, which is attached to a peptidoglycan of the cell wall [67]. This mechanism has become a well-established tool to promote protein-protein fusion [68].

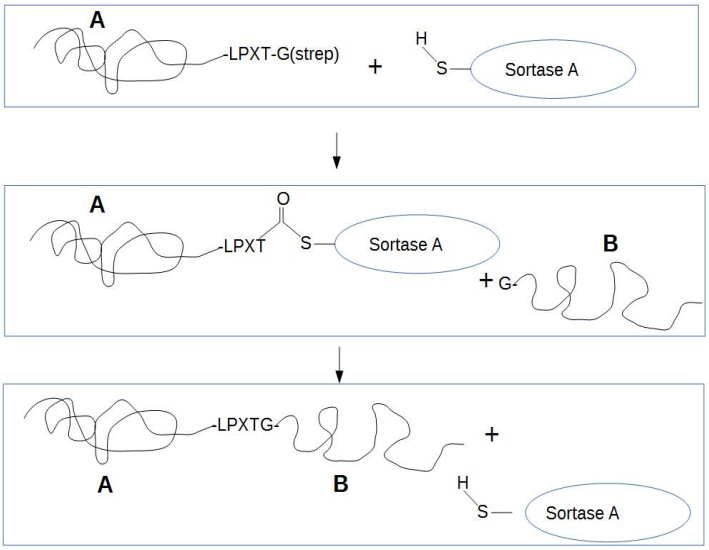


Figure 1.5: Sortase A reaction mechanism: At first Sortase A makes a nucleophilic attack on the threonine of the LPXT-G sequence leading to a thioester bond between the Sortase and the protein, and releasing the C-terminal part of the protein. Then a protein with an N-terminal glycine gets fused to the threonine thereby regenerating Sortase A.

1.6 Aim of the study

The aim of our study is to get a better picture of the binding site of v-Myc-MAX to BRCA1. We also want to investigate if there are long-range effects of this binding. A part of this is the production and measurement of a disordered region of BRCA1 of more than 1000 amino acids in length. This would also be the largest IDP measured by NMR to date.

2 Materials and methods

2.1 Buffers, media and solutions

Lysis-buffer:

25mM Tris
300mM NaCl
6M Guanidinium chloride
10mM β-mercaptoethanol
adjust to pH = 8.0 with HCl and NaOH

HisTrap elution buffer:

50mM Tris
300mM NaCl
5% Glycerol
1mM β-mercaptoethanol
200mM Imidazole
adjust to pH = 8.0

HisTrap binding buffer:

25mM Tris
300mM NaCl
4M Guanidinium chloride
1mM β-mercaptoethanol
adjust to pH = 8.0

Tris buffer:

50mM Tris
300mM
1mM β-mercaptoethanol
50 μl protease inhibitors (for lysis)
adjust to pH = 8.0

HisTrap washing buffer:

50mM Tris
150mM NaCl
5% Glycerol
1mM β-mercaptoethanol
adjust to pH = 8.0

High imidazole tris buffer:

50mM Tris
300mM NaCl
1mM β-mercaptoethanol
500mM Imidazole
adjust to pH = 8.0

Strep binding buffer:

50mM Tris
150mM NaCl
5% Glycerol
1mM EDTA
1mM Dithiothreitol(DTT)
adjust to pH = 8.0

Strep elution buffer:

50mM Tris
150mM NaCl
5% Glycerol
1mM EDTA
1mM DTT
50mM d-Biotin
adjust to pH = 8.0

Sortase reaction buffer:

50mM Tris
300mM NaCl
5% Glycerol
10mM $CaCl_2$
1mM DTT

LB-medium:

dissolve 20 g of LB-Broth in 1l H_2O
autoclave at 120°C
Add 1ml Kanamycin (1M) and 1ml Chloramphenicol (1M) to the medium before use

NMR measurement buffer:

20mM MES
20mM NaCl
5% Glycerol (only for Constructs A and B)
100mM ArgCl
1mM TCEP (fresh)
adjust to pH = 5.5

10xM9-minimal Stock Solution:

30 g KH_2PO_4
67.8 g Na_2PO_4 x $2H_2O$
5 g NaCl
1 l H_2O
autoclave at 120°C

100x Trace Elements:

830 mg $FeCl_3$ x $2H_2O$
84 g $ZnCl_2$
13 g $CuCl_2$ x $2H_2O$
10 g $CoCl_2$ x $2H_2O$
10 g H_3BO_3
1.6 g $MnCl_2$ x $4H_2O$
5 g EDTA
1 l H_2O

autoclave at 120°C

Coomassie Staining Solution:

5 g Serva Blue G-250
500 ml Methanol
100 ml Acetic acid

400 ml H_2O

LB-Agar:

dissolve 40 g of LB-Agar in 1 l H_2O
autoclave at 120°C
let it cool to 50°C in a water bath
1 ml Kanamycin (1M) and Chloramphenicol (1M) added before pouring 10-20 ml into petri dishes

M9-minimal medium (^{15}N):

100 ml 10x M9 Stock Solution
10 ml Trace Elements
1 g $^{15}NH_4Cl$
880 ml H_2O
4 g Glucose
2 ml $MgSO_4$ 1M
300 μl $CaCl_2$ 1M

1 mM Kanamycin (1M) and Chloramphenicol (1M)

Laemmli Buffer:

20 g Sodium dodecyl sulfate (SDS)
60 g Tris HCl
286 g Glycin

Dissolve in 2 l H_2O

SDS-PAGE 15% Separating Gel:

2.67 ml Tris pH=9
8 ml 40% Polyacrylamid (PAA) 37.5:1
10.6 ml H_2O
108 μl 20% SDS
16 μl TEMED
44 μl 10% APS

SDS-PAGE 10% Separating Gel:

2.67 ml Tris pH=9
5.2 ml 40% Polyacrylamid (PAA)37.5:1
13.4 ml H_2O
108 μl 20% SDS
16 μl TEMED
44 μl 10% APS

SDS-PAGE Stacking Gel:

1.17 ml Tris pH=6.8
1.17 ml 40% Polyacrylamid(PAA)19:1
7 ml H_2O
23.2 μl 20% SDS
16 μl TEMED
44 μl 10% APS

2.2 Gene constructs and cloning

2.2.1 Used constructs

Many constructs of different length and fusion site were tested by our group using SLIC-cloning. Those SLIC-reactions that worked, and had the least mutations due to creation of the Sortase A recognition site, were used for the experiment. These were the constructs A and B(old). Fig. 2.1 shows which parts of BRCA1 these constructs represent.
Construct MMB represents a smaller version of a previously tested construct with 56 amino acids, which also showed binding to MAX-MAX.

Construct A:
Homo Sapiens BRCA1 His-(3C)-199-676-LPATG-strep

Construct B(old):
Homo Sapiens BRCA1 His-(3C)-676-1357

Construct B:
Homo Sapiens BRCA1 His-(TEV)-676-1357-strep

Construct MMB (MAX-MAX binding):
Homo Sapiens BRCA1 His-MBP-(3C)-374-409

MAX:
Homo Sapiens His-(3C)-MAX

Legend:
His = 6xHistine-Tag
strep = Strep-Tactin tag
3C = 3C protease cleavage site
TEV = Tobacco Etch Virus protease cleavage site
MBP = Maltose binding protein

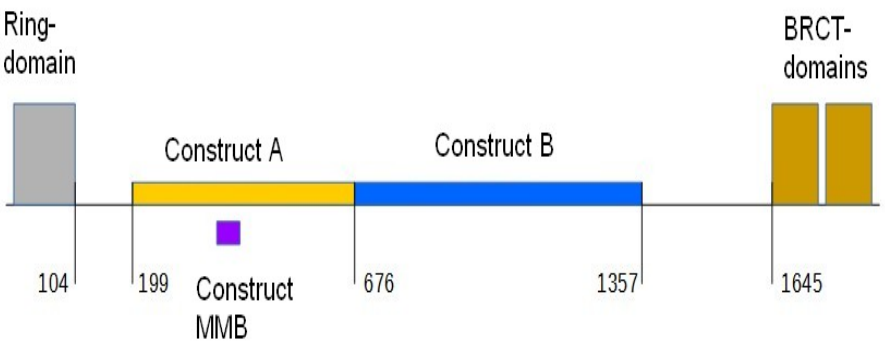


Figure 2.1: Portrayal of the Constructs: The used constructs, in comparison to the whole protein sequence, are shown.

2.2.2 Cloning

- Linear Construct B
  Starting from the full length BRCA1.

  PCR mix:

  1 μl forward 676 primer

  1 μl reverse complement 1357 primer

  5 μl dNTP(2mM Stock)

  5 μl Buffer Pfu+$MgCl_2$

  0.5 μl template (full length BRCA1 60-120ng/μl)

  36.5 μl ddH_2O

  0.5 μl Pfu Polymerase

  PCR setup:

  98°C 5 min

  -begin 30 cycles:

98°C 1 min

72°C 5 min

-end cycles:

72°C 10 min

4°C until removal

clean the PCR reaction product with a QIAGEN QIAquick PCR-purification kit (50)

- Linear pCoofy32

  PCR mix:

  1 μl forward primer (3C-F)

  1 μl reverse complement primer (strep-R)

  5 μl dNTP(2mM Stock)

  5 μl Buffer Pfu+$MgCl_2$

  0.5 μl template pCoofy32 (60-120ng/μl) 36.5 μl ddH_2O

  0.5 μl Pfu polymerase

  PCR setup:

  98°C 5 min

  -begin 30 cycles:

  98°C 1 min

  65°C 1 min

  72°C 11 min

  -end cycles:

  72°C 10 min

  4°C until removal

  clean the PCR reaction product with the QIAGEN QIAquick PCR Purification Kit (50)

- SLIC reaction 2.44 μl of Insert (0.12pmol)

  1.15 μl of Vector (0.027pmol)

  1 μl of RecA reaction buffer (NEB)

  1 μl of RecA protein (NEB, M0249 S)

  4.4 μl H_2O

  incubate for 30 min at 35°C

- Transformation:

  Put 10 μl of plasmid on a glycerol stock of TOP10 competent cells and leave it on ice for 30 min. Heat up to 42°C for 45 sec and put it back on ice for 3 min. Now add 300 μl of LB-medium and let the cells grow for 1 h at 37°C. Spin the cells down for 20 sec at 8000rpm, discard 340 μl of the supernatant and resuspend the cells in the remaining 70 μl. Take the 70 μl of the cells, spread them on an LB-Agar plate(Kan) and let them grow over night (o/n) at 37°C.

- Plasmid extraction:

  After an over night culture of 20 ml of the TOP10 cells, use the QIAGEN QIAprep Spin Miniprep Kit (250) to obtain the DNA.

- Mutation of the cleavage site (3C $\Rightarrow$ TEV)

  This was done by a single mutation event: The primers used were 60 nucleotides (nt) long, had a 22 nt overhang on one side and a 18 nt on the other site with 2 times 2 consecutive point mutations and an insert of 3 nt in between.

  PCR, DNA-cleaning and transformation were done the same way as above.

  All the PCR results were checked by DNA sequencing.

2.3 Protein expression and purification

BRCA1 Construct A and Construct B

- Transformation and Preculture

  Put 1 μl of plasmid on a glycerol stock of Rosetta phage resistant competent cells and leave the cells for 30 min on ice. Heat up to 42°C for 45 sec and put it back on ice for 3 min. Now add 300 μl of LB-medium and let the cells grow for 1 h at 37°C. Take 70 μl of the cells, spread them on an LB-Agar plate (Kan/CoA) and let them grow over night (o/n) at 37°C.
  For the preculture prepare 200 ml of LB-medium in a shaking flask, take several colonies from the LB-agar plate and add them to the LB-medium. Let the cells grow o/n at 37°C.

- Expression

  Prepare 4 x 1 l of LB-medium in shaking flasks and add 20-40 ml of preculture. Let the cells shake at 37°C until they reach an optical density (OD) of 0.6-0.8. Then centrifuge the cultures for 15 min at 3000rpm, discard the supernatant and resuspend the pellets in 1 l of M9-medium. Now put them back in the shaker and wait for 45-60 min before inducing the expression with 0.8mM IPTG.

- Harvesting

  After 2 h for Construct A or 2.5 h for Construct B centrifuge the culture at 5000rpm for 30 min. Resuspend the pellet in 50 ml of Lysis buffer, transfer to a sonication flask and sonicate on ice at 50% for 2 x 3 min. Centrifuge the lysed cells for 25 min at 18000rpm and put the supernatant into a Falcon tube.

- Purification

  Equilibrate the HisTrap with the HisTrap binding buffer. Dilute the sample with Tris buffer to get a 4M guanidinium chloride concentration before loading it on the column with 2-3 ml/min. Wash the column with the HisTrap washing buffer and elute with the HisTrap elution buffer. To get rid of the imidazole, dialyse the sample o/n at 4°C in StrepTactinXT binding buffer. For TEV-cleavage of Construct B add TEV-protease after one hour of dialysis. After that equilibrate the StrepTactinXT column with StrepTactinXT binding buffer and load the sample onto the column. Wash the column with the same buffer before elution with the StrepTactinXT elution buffer. Concentrate the sample on a Amicon Ultra15 Centrifugal Filter Device with a 30 kDa cut-off.

Construct MMB and MAX

- Transformation and preculture

  The transformations and precultures of these plasmids are done in the same way as for Constructs A and B.

- Expression

  Prepare 4 x 1 l LB in shaking flasks, add 20-40 ml of preculture and let the cells grow at 37°C to an OD of 0.6-0.8.

  For the unlabeled MAX, induce the protein production by adding 0.4mM IPTG and let the cells grow o/n at 28°C.

  For the ^{15}N labeled BRCA1 MAX-binding site, centrifuge the cells for 15 min at 3000rpm, discard the supernatant and resuspend the pellets in 1 l of M9-medium. Afterwards wait for 45 min, induce with 0.4mM IPTG and express o/n at 28°C.

- Harvesting

  After one night, centrifuge the cultures at 5000rpm for 30 min. Resuspend the pellet in 50 ml of Tris buffer, transfer to a sonication flask and sonicate on ice at 50% for 2 x 3 min. Centrifuge the lysed cells for 25 min at 18000rpm and put the supernatant into a Falcon tube.

- Purification

  Equilibrate the HisTrap with the Tris buffer. Load the supernatant on the column with 2-3 ml/min, wash with 30 ml of Tris buffer and then elute with a gradient of high imidazole buffer, increasing to 100% in 60 min. To get rid of the imidazole, dialyse the sample o/n at 4°C in Tris buffer. Then add 3C protease in a ratio 1:100 (w/w) and let the reaction run for 1-2 h. After that, run the HisTrap again with the same buffers and collect the flow-through. Concentrate the sample on an Amicon Ultra15 Centrifugal Filter Device with a 3 kDa cut-off.

2.4 Protein ligation

Blend Constructs A and B at a ratio of 1:1 to get a concentration in the low μM range (approx. 20μM). Add Sortase A at a ratio of 1:50 to the sample. To ensure proper function of Sortase A, 10mM $CaCl_2$ is added. After 1 h at 40°C the reaction is stopped by adding 15mM EDTA.

2.5 Separation of the fusion construct

As the fusion by Sortase A only works to about 30%, the fusion product still needs to be separated from the two constructs. This is done by size exclusion chromatography with a Superdex S200 column in the Sortase reaction buffer. The product of the first run needs to be run again on the same column to increase the purity.

2.6 NMR experiments and analysis

All the experiments were measured on a Bruker Avance III HD spectrometer which operates at 800 MHz or 18.79 tesla. Spectra processing was done with NMRpipe [69]. Analysing and assignment completion were done using CcpNmr [70].

2.6.1 Constructs A and B

For all experiments the NMR measurement buffer was used. The experiments were done at 293K(Constructs A and B) and at 295K(Construct MMB).

- Construct A ^{1}H-^{15}N-HSQC
  ^{15}N-labeled Construct A had a concentration of ∼80μM in 400 μl containing 10% of D_2O.

- Construct B ^{1}H-^{15}N-HSQC
  ^{15}N-labeled Construct had a concentration of ∼50μM in 400 μl containing 10% of D_2O.

- Construct MMB ^{1}H-^{15}N-HSQC-NOESY
  ^{15}N-labeled Construct MMB for the HSQC-NOESY experiment had a concentration of ∼300μM in 400 μl containing 10% of D_2O.

- Construct MMB ^{1}H-^{15}N-HSQC MAX titration
  ^{15}N-labeled Construct MMB for the MAX titration experiment had a concentration of ∼48μM in 400 μl containing 10% of D_2O.

3 Results and discussion

3.1 MAX-MAX binding domain

3.1.1 Construct MMB and MAX

The buffers for both proteins were changed to the NMR measurement buffer. The operating temperature was 295K.

- MAX titration to Construct MMB

  The ^{15}N-labelled Construct MMB was diluted to a concentration of 50μM with 10% of D_2O to 400 μl. After an HSQC measurement, the sample was taken out and ^{14}N unlabeled MAX was added. In the first step of titration, MAX was added to a concentration of 30μM, which diluted Construct MMB to 48μM. The second titration had 95μM MAX and 45μM of Construct MMB and the last titration had 135μM MAX and 42μM of Construct MMB. Between every titration step an HSQC was recorded.

- HSQC and NOESY of Construct MMB and MAX

  For this experiment Construct MMB was diluted to a final concentration of 0.3mM with MAX being 0.8mM and with 10% D_2O to a volume of 400 μl. First an HSQC measurement was done, followed by an HSQC-NOESY and afterwards again an HSQC to check if the spectra stayed the same during the long measurement. Due to unpublished data (done by group Konrat) it is known that a MAX-MAX-homodimer binds to BRCA1 in the same manner as the v-Myc-MAX-heterodimer. The shifts upon binding of MAX have already been tracked to residues 374 and 430 of BRCA1. Construct MMB was used to further narrow the binding site down. Most of the peaks could be identified by using the unpublished assignment of a larger construct of BRCA1. The rest of the peaks could be assigned by using the cross-peaks and the knowledge of the primary protein sequence. Cross-peaks between backbone N-H were only seen for neighbouring amino acids. The type of these neighbouring amino acids could then be identified by the use of the side

chain cross-peaks. This was sufficient information to complete the assignment by elimination method. With the assignment completed and with the titration of MAX, the location of the binding site was narrowed down. Fig. 3.3 shows the chemical shift changes of the backbone N-H. These changes were seen between residues 384 and 408 (Fig.3.3), but were largest between residues 390 and 395. So this is likely the region contributing most strongly to the binding.

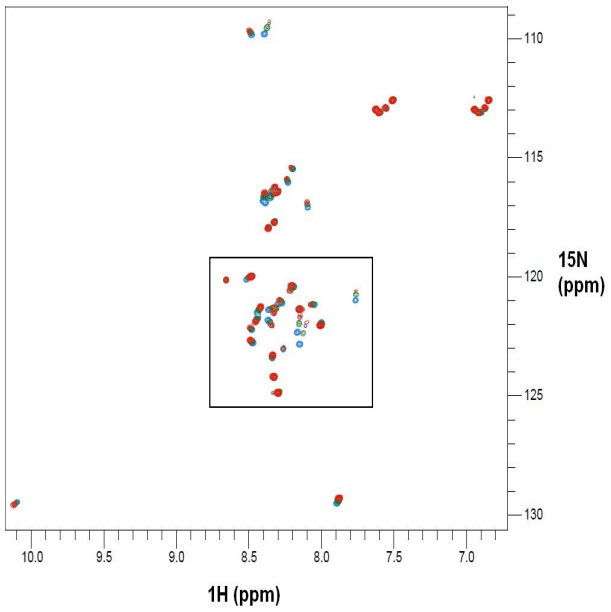


Figure 3.1: ^{1}H-^{15}N-HSQC-Spectra of Construct MMB and MAX: Blue peaks are from Construct MMB without MAX ;red peaks are after saturation with MAX

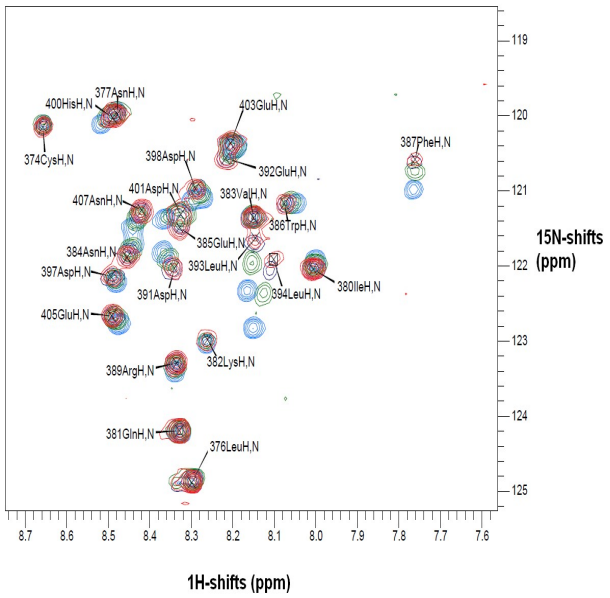


Figure 3.2: ^{1}H-^{15}N-HSQC-Spectra of Construct MMB and MAX (zoom): In this zoom the labelled peaks and the chemical shift changes can be seen.

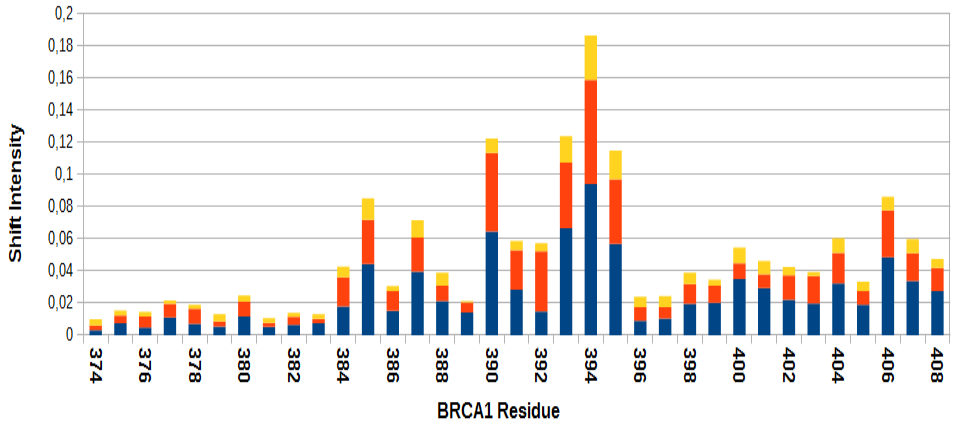
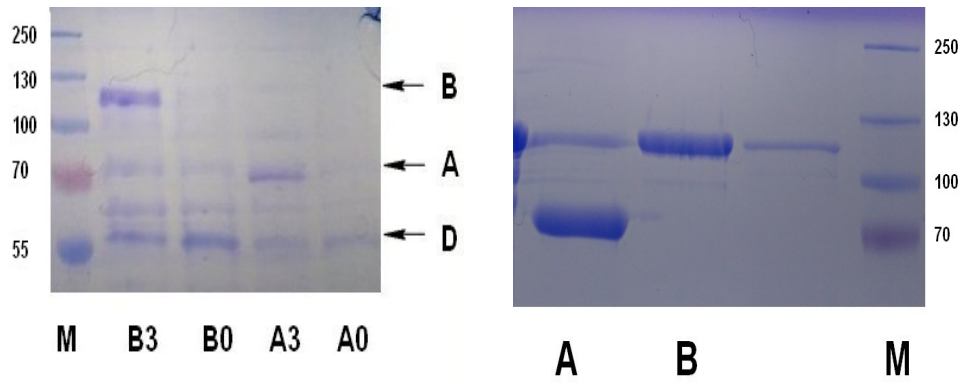


Figure 3.3: Construct MMB + MAX chemical shifts: Sum of the peak-shifts in both dimensions(^{1}H and ^{15}N) added up.

3.2 Production and measurement of Construct A and B

- Expression and purification
  For constructs A and B optimization of expression and purification was one of the most difficult parts. This is because both constructs are highly disordered and therefore degrade within hours in E.coli cells [71]. In an over night expression test all the protein was already degraded. If expressed only for 2 to 3 hours, as shown in Fig. 3.4 (a), the best results could be obtained.



(a) (b)

Figure 3.4: Expression and purification of Constructs A and B: (15% gels) **(a):** Expression Test:Marker in kDa (M); Construct B after 3 h (B3) expression(exp.); B o/n exp. (B0); Construct A 3h exp. (A3); A o/n exp. (A0). Both constructs show degradation when expressed o/n. **(b):** After purification with a His-Column and a Strep-Column high purity of Construct A(band A) and Construct B(band B) can be observed.

To avoid fast degradation during lysis of the cells, denaturing conditions were created using 6M guanidinium chloride, which should unfold all the proteases. Additionally, adding a Strep-Tactin tag on Construct B led to samples of high purity due to highly specific interaction with

the Strep-Tactin column (Fig. 3.4 (b)). To get a high enough concentration of each sample for more complex NMR experiments, much larger expressions have to be made.

- NMR measurements
  With the available 800MHz spectrometer, we were able to obtain ^{1}H-^{15}N HSQC spectra of high signal to noise ratio of Constructs A (Fig.3.5 and of Construct B (Fig. 3.6). The concentrations were 70μM for Construct A and 50μM for Construct B. Due to the massive overlap, only very little information can be extracted from these spectra. However, the N-H of the tryptophane side chains are clearly separated and of the right quantity for each construct. To get more information or to assign the peaks of the spectra 4D or 5D spectra would need to be recorded, and for that a ^{13}C + ^{15}N labelled sample with a concentration of at least 0.4mM would need to be made. For Construct A and Construct B we have already tested to concentrate them. Stable samples with concentrations of about 0.5mM could be made. Furthermore, a spectrometer with a higher field strength and with a Cryo-Cooled detection coil is an advantage for such large constructs as this produces sharper peaks with less thermal noise.

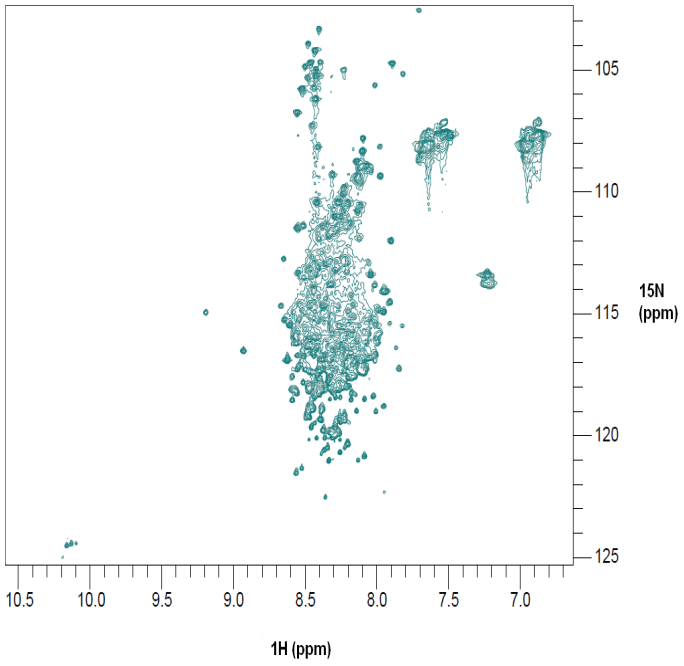


Figure 3.5: Construct A(508aa), ^{1}H-^{15}N-HSQC: Due to the many amino acids there is a huge overlap in the central part of the spectrum. The spectrum shows 4 tryptophane(W)-sidechain peaks at 10.2ppm(H-dimension) and 124ppm(N-dimension).

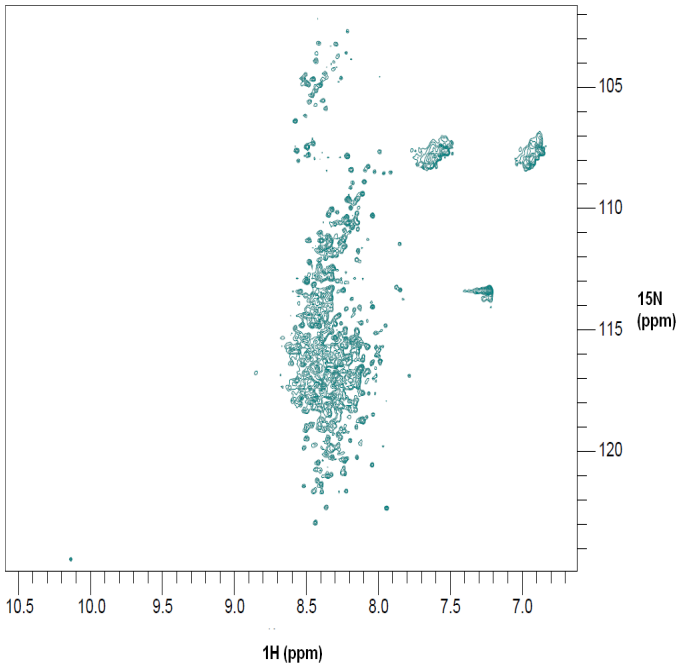
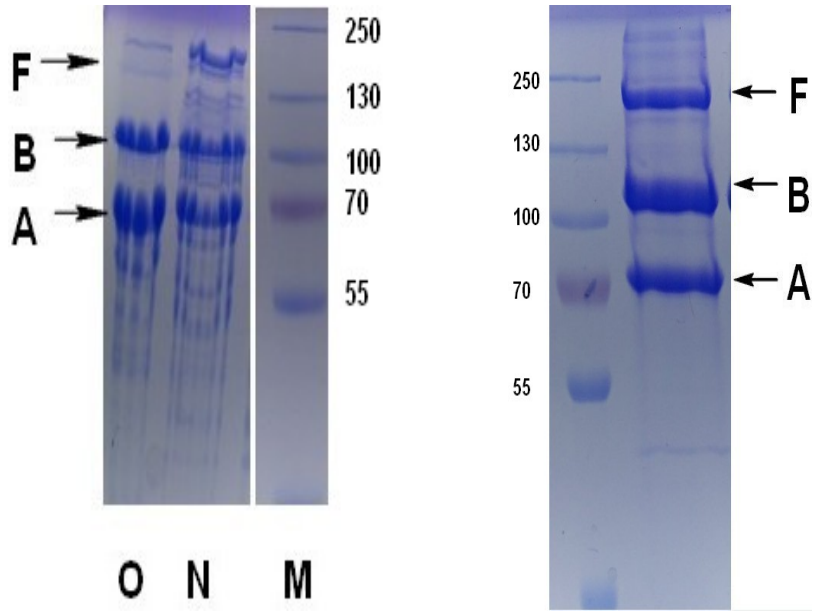


Figure 3.6: Construct B(709aa), ^{1}H-^{15}N-HSQC: One W-sidechain peak at 10.2ppm(H-dimension) and 124 ppm(N-dimension) can be seen.

3.3 Protein ligation results

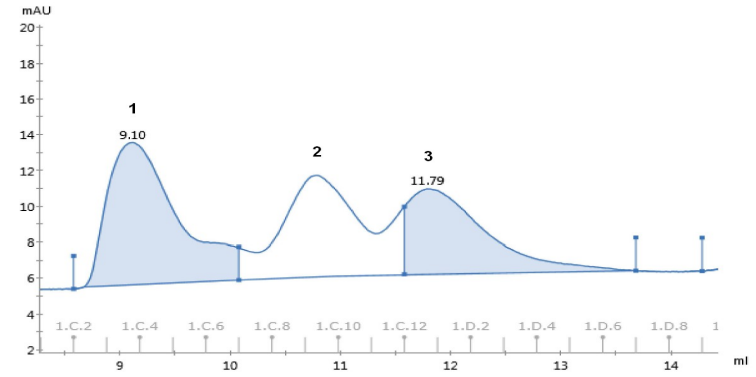
The recognition sequence used on the C-terminus of Construct A was LPATG followed by a serine of the Strep-Tactin tag. At first, a Construct B with a 3C protease cleavage site was used, which started with a GP- after digestion. With that construct, no fusion could be observed. After cloning of Construct B, it started with GG- due to the usage of a TEV protease cleavage site. This led to good results that were optimized further by changing the reaction conditions to the final setup (shown in Fig. 2.4). Further reaction optimisation may be achieved by dialysis during the reaction, which should get rid of the G-Strep-Tactin fragment that was cut off but which may still act as a target for fusion. We haven't tested this approach yet.



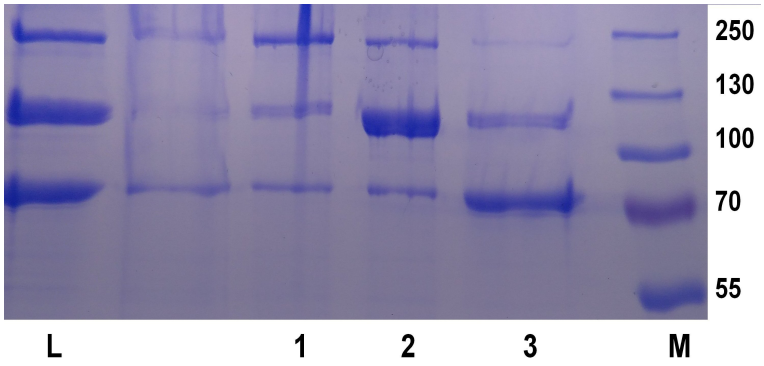
(a) (b)

Figure 3.7: Fusion of the constructs: (10% gels) **(a):** Sortase reaction: Old Construct B (O) with almost no fusion construct; new Construct B (N) with high yield fusion construct(F).**(b):** Best reaction result with pure samples; F being the fusion construct of Construct B (B) and Construct A (A). Marker (M) in kDa.

- Separation of the fusion product
  The enrichment of the fusion construct seems to work (Fig. 3.9), although the dilution, and ultimately the protein loss during each run, is pretty significant since the column needs to be run at least twice. Another promising approach could be to add two different tags on the N-term of Construct A and on the C-terminus of Construct B. This second tag needs to be resistant to denaturation by 6M Guanidinium chloride and it can't be the His-Tag of Construct A, as both constructs bind to the His-Column unspecifically. This would lead to a fusion construct that has two different highly specific tags, as opposed to the fragments, which would still only have one tag each. An overview of possible affinity tags was published by Kimple et. al. [72].

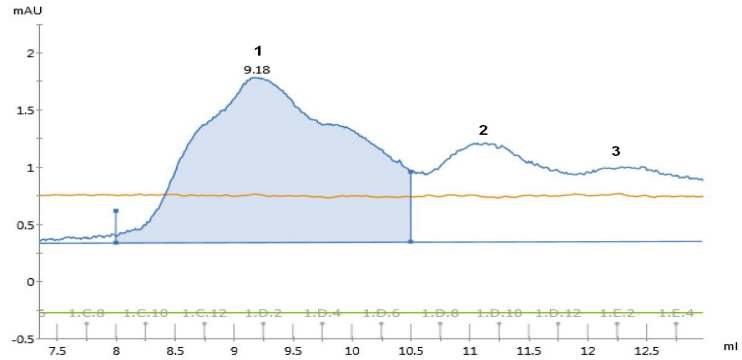


(a)

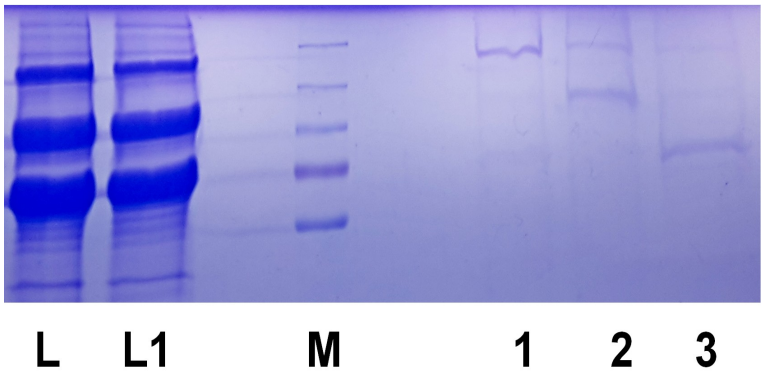


(b)

Figure 3.8: First separation run of the fusion construct: (10% gels) **(a):** First separation: load (L); first peak (1); second peak (2); third peak (3); Marker in kDa (M).**(b):** second separation: load((L)peak 1 of first run); load of first run (L1); marker (M); first peak(1); second peak(2); third peak(3).



(a)



(b)

Figure 3.9: Second separation run of the fusion construct: (10% gels) **(a):** First separation: load (L); first peak (1); second peak (2); third peak (3); Marker in kDa (M).**(b):** second separation: load((L)peak 1 of first run); load of first run (L1); marker (M); first peak(1); second peak(2); third peak(3).

4 Conclusion

In our effort to better understand the binding of the v-Myc-MAX heterodimer to BRCA1, some progress was made. By the use of a MAX-MAX homodimer and a small fragment of BRCA1 (Construct MMB), we could narrow down the binding site on BRCA1. This was done by first assigning the peaks from a ^{1}H-^{15}N-HSQC of the BRCA1 fragment, using the assignment of a larger construct. Then a titration of the MAX-MAX homodimer was made, showing chemical shift changes of the backbone N-H peaks at the binding site.

For identification of possible long-range effects of this binding, two large fragments of BRCA1 (Construct A and B) were successfully expressed. Together these two fragments represent a large part of the intrinsically disordered region of BRCA1. Both constructs could already be measured by NMR and showed promising spectra for further experiments. It was also possible to get stable samples with high enough concentrations for multidimensional NMR experiments.

These two fragments were fused together into one large construct. This was done by Sortase A ligation. Due to the fact that such a ligation is never flawless, separations had to be performed to get rid of the non-ligated constructs. By using size-exclusion chromatography, some accumulation of the fusion construct was achieved.

To further characterize the binding structure, a ^{13}C-^{15}N labelled Construct MMB has to be expressed, purified and measured. This allows to measure the side chains and investigate their involvement in binding. To complete the picture, it would of course be necessary to investigate the binding site from the perspective of the MAX-MAX dimer. This is however a little bit more difficult, because the MAX-MAX dimer is very elongated, resulting in a slow rotational motion. As described in Sec. 1.2.1 this decreases the spectral density leading to decreased intensities. By using higher temperatures while measuring, the tumbling

of the proteins can be increased. The problem here is that the MAX-MAX dimer gets less stable with higher temperature. To overcome this problem, a more stable dimer would have to be used.

For investigating long range effects of MAX-MAX binding to BRCA1, it is necessary to produce large amounts of a ^{13}C-^{15}N labelled fusion construct.

References

[1] A. C. Bloomer, J. N. Champness, G. Bricogne, R. Staden, and A. Klug. Protein disk of tobacco mosaic virus at 2.8 Å resolution showing the interactions within and between subunits. *Nature*, 276(5686):362–368, nov 1978.

[2] Chris BOESCH, Arno BUNDI, Max OPPLIGER, and Kurt WuTHRICH. 1H Nuclear-Magnetic-Resonance Studies of the Molecular Conformation of Monomeric Glucagon in Aqueous Solution. *European Journal of Biochemistry*, 91(1):209–214, nov 1978.

[3] R W Kriwacki, L Hengst, L Tennant, S I Reed, and P E Wright. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 93(21):11504–9, oct 1996.

[4] G W Daughdrill, M S Chadsey, J E Karlinsey, K T Hughes, and F W Dahlquist. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. *Nature structural biology*, 4(4):285–91, apr 1997.

[5] P Romero, Z Obradovic, C R Kissinger, J E Villafranca, E Garner, S Guilliot, and A K Dunker. Thousands of proteins likely to have long disordered regions. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, pages 437–48, 1998.

[6] A K Dunker, J D Lawson, C J Brown, R M Williams, P Romero, J S Oh, C J Oldfield, A M Campen, C M Ratliff, K W Hipps, J Ausio, M S Nissen, R Reeves, C Kang, C R Kissinger, R W Bailey, M D Griswold, W Chiu, E C Garner, and Z Obradovic. Intrinsically disordered protein. *Journal of molecular graphics & modelling*, 19(1):26–59, 2001.

[7] A K Dunker, E Garner, S Guilliot, P Romero, K Albrecht, J Hart, Z Obradovic, C Kissinger, and J E Villafranca. Protein disorder and the evolution of molecular recognition: theory, predictions

and observations. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, pages 473–84, 1998.

[8] V N Uversky, J R Gillespie, and A L Fink. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins*, 41(3):415–27, nov 2000.

[9] L. M. Iakoucheva, Predrag Radivojac, Celeste J Brown, Timothy R O'Connor, Jason G Sikes, Zoran Obradovic, and A Keith Dunker. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Research*, 32(3):1037–1049, feb 2004.

[10] Mark O. Collins, Lu Yu, Iain Campuzano, Seth G. N. Grant, and Jyoti S. Choudhary. Phosphoproteomic Analysis of the Mouse Brain Cytosol Reveals a Predominance of Protein Phosphorylation in Regions of Intrinsic Sequence Disorder. *Molecular & Cellular Proteomics*, 7(7):1331–1348, jul 2008.

[11] Dana Vuzman and Yaakov Levy. Intrinsically disordered regions as affinity tuners in protein–DNA interactions. *Mol. BioSyst.*, 8(1):47–57, 2012.

[12] Peter E. Wright and H. Jane Dyson. Intrinsically Disordered Proteins in Cellular Signaling and Regulation. *Nature reviews. Molecular cell biology*, 16(1):18, 2015.

[13] A. Keith Dunker, Marc S. Cortese, Pedro Romero, Lilia M. Iakoucheva, and Vladimir N. Uversky. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS Journal*, 272(20):5129–5148, oct 2005.

[14] J.J. Ward, J.S. Sodhi, L.J. McGuffin, B.F. Buxton, and D.T. Jones. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology*, 337(3):635–645, mar 2004.

[15] Krzysztof Kazimierczuk, Jan Stanek, Anna Zawadzka-Kazimierczuk, and Wiktor Koźmiński. Random sampling in

multidimensional NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 57(4):420–434, nov 2010.

[16] Benjamin Frühbauer, Borja Mateos, and Robert Konrat. 1H, 15N, 13C resonance assignment of the human CD44 cytoplasmic tail (669-742). *Biomol NMR Assign (2018)*, 2018.

[17] Eri Sakata, Yoshiki Yamaguchi, Eiji Kurimoto, Jun Kikuchi, Shigeyuki Yokoyama, Shingo Yamada, Hiroyuki Kawahara, Hideyoshi Yokosawa, Nobutaka Hattori, Yoshikuni Mizuno, Keiji Tanaka, and Koichi Kato. Parkin binds the Rpn 10 subunit of 26S proteasomes through its ubiquitin-like domain. *EMBO Reports*, 2003.

[18] O. W. Sørensen. James Keeler. Understanding NMR Spectroscopy. *Magnetic Resonance in Chemistry*, 2006.

[19] Shahid N. Khan, Cyril Charlier, Rafal Augustyniak, Nicola Salvi, Victoire Déjean, Geoffrey Bodenhausen, Olivier Lequin, Philippe Pelupessy, and Fabien Ferrage. Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. *Biophysical Journal*, 2015.

[20] J M Hall, M K Lee, B Newman, J E Morrow, L A Anderson, B Huey, and M C King. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science (New York, N.Y.)*, 250(4988):1684–9, dec 1990.

[21] D F Easton, D T Bishop, D Ford, and G P Crockford. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *American journal of human genetics*, 52(4):678–701, apr 1993.

[22] T M Smith, M K Lee, C I Szabo, N Jerome, M McEuen, M Taylor, L Hood, and M C King. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1. *Genome research*, 6(11):1029–49, nov 1996.

[23] Mara Colombo, Marinus J. Blok, Phillip Whiley, Marta Santamariña, Sara Gutiérrez-Enríquez, Atocha Romero, Pilar Garre, Alexandra Becker, Lindsay Denise Smith, Giovanna De Vecchi, Rita D. Brandão, Demis Tserpelis, Melissa Brown, Ana Blanco, Sandra Bonache, Mireia Menéndez, Claude Houdayer, Claudia Foglia, James D. Fackenthal, Diana Baralle, Barbara Wappenschmidt, Eduardo Díaz-Rubio, Trinidad Caldés, Logan Walker, Orland Díez, Ana Vega, Amanda B. Spurdle, Paolo Radice, and Miguel De La Hoya. Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Human Molecular Genetics*, 23(14):3666–3680, jul 2014.

[24] J A Duncan, J R Reeves, and T G Cooke. BRCA1 and BRCA2 proteins: roles in health and disease. *Molecular pathology : MP*, 51(5):237–47, oct 1998.

[25] T F Lane, C Deng, A Elson, M S Lyu, C A Kozak, and P Leder. Expression of Brca1 is associated with terminal differentiation of ectodermally and mesodermally derived tissues in mice. *Genes & development*, 9(21):2712–22, nov 1995.

[26] J V Rajan, M Wang, S T Marquis, and L A Chodosh. Brca2 is coordinately regulated with Brca1 during proliferation and differentiation in mammary epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America*, 93(23):13078–83, nov 1996.

[27] J M Gudas, T Li, H Nguyen, D Jensen, F J Rauscher, and K H Cowan. Cell cycle regulation of BRCA1 messenger RNA in human breast epithelial cells. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 7(6):717–23, jun 1996.

[28] L C Wu, Z W Wang, J T Tsan, M A Spillman, A Phung, X L Xu, M C Yang, L Y Hwang, A M Bowcock, and R Baer. Identification

of a RING protein that can interact in vivo with the BRCA1 gene product. *Nature genetics*, 14(4):430–40, dec 1996.

[29] Megan Fabbro, Jose A Rodriguez, Richard Baer, and Beric R Henderson. BARD1 induces BRCA1 intranuclear foci formation by increasing RING-dependent BRCA1 nuclear import and inhibiting BRCA1 nuclear export. *The Journal of biological chemistry*, 277(24):21315–24, jun 2002.

[30] R I Yarden and L C Brody. BRCA1 interacts with components of the histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America*, 96(9):4983–8, apr 1999.

[31] X Yu, L C Wu, A M Bowcock, A Aronheim, and R Baer. The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *The Journal of biological chemistry*, 273(39):25388–92, sep 1998.

[32] Peter S. Brzovic, Ponni Rajagopal, David W. Hoyt, Mary-Claire King, and Rachel E. Klevit. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nature Structural Biology*, 8(10):833–837, oct 2001.

[33] R. Scott Williams, Ruth Green, and J.N. Mark Glover. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nature Structural Biology*, 8(10):838–842, oct 2001.

[34] Wen-Ying Mark, Jack C C Liao, Ying Lu, Ayeda Ayed, Rob Laister, Blair Szymczyna, Avi Chakrabartty, and Cheryl H Arrowsmith. Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *Journal of molecular biology*, 345(2):275–87, jan 2005.

[35] Q Wang, H Zhang, K Kajino, and M I Greene. BRCA1 binds c-Myc and inhibits its transcriptional and transforming activity in cells. *Oncogene*, 17(15):1939–48, oct 1998.

[36] H Zhang, K Somasundaram, Y Peng, H Tian, H Zhang, D Bi, B L Weber, and W S El-Deiry. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene*, 16(13):1713–21, apr 1998.

[37] O N Aprelikova, B S Fang, E G Meissner, S Cotter, M Campbell, A Kuthiala, M Bessho, R A Jensen, and E T Liu. BRCA1-associated growth arrest is RB-dependent. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21):11866–71, oct 1999.

[38] Q Zhong, C F Chen, S Li, Y Chen, C C Wang, J Xiao, P L Chen, Z D Sharp, and W H Lee. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science (New York, N.Y.)*, 285(5428):747–50, jul 1999.

[39] R Scully, J Chen, A Plug, Y Xiao, D Weaver, J Feunteun, T Ashley, and D M Livingston. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell*, 88(2):265–75, jan 1997.

[40] Alexandra Folias, Mara Matkovic, Donald Bruun, Sonja Reid, James Hejna, Markus Grompe, Alan D'Andrea, and Robb Moses. BRCA1 interacts directly with the Fanconi anemia protein FANCA. *Human molecular genetics*, 11(21):2591–7, oct 2002.

[41] Yan-Fen Hu and Rong Li. JunB potentiates function of BRCA1 activation domain 1 (AD1) through a coiled-coil-mediated interaction. *Genes & development*, 16(12):1509–17, jun 2002.

[42] J Chen, D P Silver, D Walpita, S B Cantor, A F Gazdar, G Tomlinson, F J Couch, B L Weber, T Ashley, D M Livingston, and R Scully. Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells. *Molecular cell*, 2(3):317–28, sep 1998.

[43] X Xu, W Qiao, S P Linke, L Cao, W M Li, P A Furth, C C Harris, and C X Deng. Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis. *Nature genetics*, 28(3):266–71, jul 2001.

[44] Y R Chen, X Wang, D Templeton, R J Davis, and T H Tan. The role of c-Jun N-terminal kinase (JNK) in apoptosis induced by ultraviolet C and gamma radiation. Duration of JNK activation may determine cell death and proliferation. *The Journal of biological chemistry*, 271(50):31929–36, dec 1996.

[45] B. Wang, K. Hurov, K. Hofmann, and S. J. Elledge. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes & Development*, 23(6):729–739, mar 2009.

[46] Hongtae Kim, Jun Huang, and Junjie Chen. CCDC98 is a BRCA1-BRCT domain–binding protein involved in the DNA damage response. *Nature Structural & Molecular Biology*, 14(8):710–715, aug 2007.

[47] Yi Wang, David Cortez, Parvin Yazdi, Norma Neff, Stephen J. Elledge, and Jun Qin. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes and Development*, 2000.

[48] L C Hsu, T P Doan, and R L White. Identification of a gamma-tubulin-binding domain in BRCA1. *Cancer research*, 61(21):7713–8, nov 2001.

[49] T D Littlewood, B Amati, H Land, and G I Evan. Max and c-Myc/Max DNA-binding activities in cell extracts. *Oncogene*, 7(9):1783–92, sep 1992.

[50] E M Blackwood and R N Eisenman. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science (New York, N.Y.)*, 251(4998):1211–7, mar 1991.

[51] F Bahram, N von der Lehr, C Cetinkaya, and L G Larsson. c-Myc hot spot mutations in lymphomas result in inefficient ubiquitination and decreased proteasome-mediated turnover. *Blood*, 95(6):2104–10, mar 2000.

[52] Bruno Amati, Stephen Dalton, Mary W. Brooks, Trevor D. Littlewood, Gerard I. Evan, and Hartmut Land. Transcriptional activation by the human c-Myc oncoprotein in yeast requires interaction with Max. *Nature*, 359(6394):423–426, oct 1992.

[53] Satish K Nair and Stephen K Burley. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell*, 112(2):193–205, jan 2003.

[54] Clement M. Lee and E. Premkumar Reddy. The v-myc oncogene, 1999.

[55] N C Kan, C S Flordellis, G E Mark, P H Duesberg, and T S Papas. Nucleotide sequence of avian carcinoma virus MH2: two potential onc genes, one related to avian virus MC29 and the other related to murine sarcoma virus 3611. *Proceedings of the National Academy of Sciences of the United States of America*, 1984.

[56] N Walther, R Lurz, T Patschinsky, H W Jansen, and K Bister. Molecular cloning of proviral DNA and structural analysis of the transduced myc oncogene of avian oncovirus CMII. *Journal of virology*, 1985.

[57] J Hayflick, P H Seeburg, R Ohlsson, S Pfeifer-Ohlsson, D Watson, T Papas, and P H Duesberg. Nucleotide sequence of two overlapping myc-related genes in avian carcinoma virus OK10 and their relation to the myc genes of other viruses and the cell. *Proceedings of the National Academy of Sciences of the United States of America*, 1985.

[58] D L Doggett, a L Drake, V Hirsch, M E Rowe, V Stallard, and J I Mullins. Structure, origin, and transforming activity of feline

leukemia virus-myc recombinant provirus FTT. *Journal of virology*, 1989.

[59] Clement M Lee and E Premkumar Reddy. The v-myc oncogene. *Oncogene*, 18(19):2997–3003, may 1999.

[60] Wolfgang Fieber, Martin L. Schneider, Theresia Matt, Bernhard Kräutler, Robert Konrat, and Klaus Bister. Structure, function, and dynamics of the dimerization and DNA-binding domain of oncogenic transcription factor v-Myc11Edited by P. E. Wright. *Journal of Molecular Biology*, 307(5):1395–1410, apr 2001.

[61] W L DeLano. The PyMOL Molecular Graphics System, Version 1.8. *Schrödinger LLC*, 2002.

[62] Simon Sauvé, Luc Tremblay, and Pierre Lavigne. The NMR solution structure of a mutant of the max b/HLH/LZ free of DNA: Insights into the specific and reversible DNA binding mechanism of dimeric transcription factors. *Journal of Molecular Biology*, 2004.

[63] O Schneewind, A Fowler, and K F Faull. Structure of the cell wall anchor of surface proteins in Staphylococcus aureus. *Science (New York, N.Y.)*, 268(5207):103–6, apr 1995.

[64] W W Navarre and O Schneewind. Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in gram-positive bacteria. *Molecular microbiology*, 14(1):115–21, oct 1994.

[65] H Ton-That, G Liu, S K Mazmanian, K F Faull, and O Schneewind. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of Staphylococcus aureus at the LPXTG motif. *Proceedings of the National Academy of Sciences of the United States of America*, 96(22):12424–9, oct 1999.

[66] Kathleen W Clancy, Jeffrey A Melvin, and Dewey G McCafferty. Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers*, 94(4):385–96, 2010.

[67] H Ton-That, K F Faull, and O Schneewind. Anchor structure of staphylococcal surface proteins. A branched peptide that links the carboxyl terminus of proteins to the cell wall. *The Journal of biological chemistry*, 272(35):22285–92, aug 1997.

[68] David A Levary, Ranganath Parthasarathy, Eric T Boder, and Margaret E Ackerman. Protein-protein fusion catalyzed by sortase A. *PloS one*, 6(4):e18342, apr 2011.

[69] F Delaglio, S Grzesiek, G W Vuister, G Zhu, J Pfeifer, and A Bax. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, 6(3):277–93, nov 1995.

[70] Simon P. Skinner, Benjamin T. Goult, Rasmus H. Fogh, Wayne Boucher, Tim J. Stevens, Ernest D. Laue, and Geerten W. Vuister. Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallographica Section D Biological Crystallography*, 71(1):154–161, jan 2015.

[71] Marcin J Suskiewicz, Joel L Sussman, Israel Silman, and Yosef Shaul. Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein science : a publication of the Protein Society*, 20(8):1285–97, aug 2011.

[72] Michelle E. Kimple, Allison L. Brill, and Renee L. Pasker. Overview of affinity tags for protein purification. *Current Protocols in Protein Science*, 2013.