

IDENTIFIKATION UND CHARAKTERISIERUNG RNA GUANIN-QUADRUPLEX BINDENDER PROTEINE

-

vorgelegt von
Diplom-Biologin (t. o.)
Annekathrin Almut von Hacht
geb. in Waiblingen

von der Fakultät III – Prozesswissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktorin der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende:	Prof. Dr.-Ing. Vera Meyer
Gutachter:	Prof. Dr. Jens Kurreck
Gutachter:	Prof. Dr. Mario Mörl
Gutachter:	Prof. Dr. Peter Neubauer

Tag der wissenschaftlichen Aussprache: 9. September 2015

Berlin 2015

Inhalt

1.	Einleitung.....	1
1.1.	Guanin-Quadruplexe.....	1
1.2.	Vorkommen von G-Quadruplexen.....	4
1.2.1.	G-Quadruplexe in Telomeren	5
1.2.2.	G-Quadruplexe in Promotoren	6
1.2.3.	RNA G-Quadruplexe	6
1.3.	G-Quadruplex bindende Proteine.....	10
1.4.	In dieser Arbeit untersuchte 5'-UTR G-Quadruplexe.....	13
1.4.1.	Membrane-type matrix metalloproteinase 3.....	13
1.4.2.	Actin related Protein 2/3 complex subunit 2	13
2.	Zielsetzung.....	17
3.	Material und Methoden	18
3.1.	Materialien.....	18
3.1.1.	Geräte	18
3.1.2.	Chemikalien, Puffer und Kits	19
3.1.3.	Bakterienstämme	26
3.1.4.	Zelllinien.....	26
3.1.5.	Oligonucleotide	27
3.2.	Methoden	30
3.2.1.	Mikrobiologische Methoden	30
3.2.2.	Kultivierung eukaryotischer Zellen	30
3.2.3.	Klonierungen, PCR und Plasmidisolation.....	31
3.2.4.	Biophysikalische Charakterisierung der RNA G-Quadruplex Oligonucleotide	33
3.2.5.	Luciferase-Reporter-Assay.....	33
3.2.6.	Identifizierung RNA G-Quadruplex bindender Proteine.....	34
3.2.7.	MALDI TOF Massenspektrometrie.....	36
3.2.8.	Heterologe Proteinexpression	37

3.2.9. Electrophoretic Mobility Shift Assay	38
3.2.10. Oberflächenplasmonresonanz-Spektroskopie	38
3.2.11. Luciferase Co-Transfektionen	40
3.2.12. In vitro Transkription und In vitro Translation	40
3.2.13. SILAC Pull-down-Assay.....	41
3.2.14. Quantitative Real Time PCR.....	42
3.2.15. Statistische Auswertung	42
4. Ergebnisse.....	43
4.1. Auswahl der G-Quadruplex-Sequenzen.....	43
4.2. Charakterisierung der G-Quadruplex Oligonucleotide	45
4.2.1. Bioinformatische Analyse	46
4.2.2. CD-Spektroskopie	47
4.2.3. UV-Schmelzkurven.....	48
4.2.4. Zusätzliche Kontrolloligonucleotide	50
4.3. Einfluss der G-Quadruplexe auf die Translation	53
4.4. Identifizierung RNA G-Quadruplex bindender Proteine	56
4.4.1. Pull-down-Assay	56
4.4.2. Analyse der Pull-down-Assays	56
4.4.3. Massenspektrometrische Analyse der Proteinbanden	58
4.5. Proteinexpression	61
4.6. Interaktionsstudien.....	62
4.6.1. Nucleolin	62
4.6.2. U2AF65	68
4.6.3. EFHD2	73
4.6.4. SRSF1	75
4.7. Funktionelle Studien	76
4.7.1. Luciferase Co-Transfektionen	76
4.7.2. In vitro Translation.....	77

4.8.	Bestätigung des Pull-down Experiments mittels Shotgun Proteomics.....	78
4.9.	ARP2/3-Komplex G-Quadruplexe	83
4.9.1.	ARPC2-UTR Varianten.....	83
4.9.2.	G-Quadruplexe des ARP2/3-Komplexes	86
5.	Diskussion	88
5.1.	Charakterisierung der G-Quadruplexe.....	88
5.2.	Einfluss der G-Quadruplexe auf die Translation	90
5.3.	RNA G-Quadruplex bindende Proteine.....	91
5.3.1.	Nucleolin.....	91
5.3.2.	Spleißfaktoren	93
5.3.3.	Ribosomale Proteine	97
5.3.4.	Proteine ohne bekannte RNA-Bindeaktivität	98
5.4.	ARPC2 Spleißvarianten.....	100
5.5.	G-Quadruplexe im ARP2/3-Komplex	101
5.6.	Ausblick	102
6.	Zusammenfassung.....	104
7.	Summary.....	105
8.	Literaturverzeichnis	106
9.	Abkürzungsverzeichnis	116
10.	Danksagung	118
11.	Anhang.....	119
11.1.	Plasmidkarten	119
11.2.	SDS-PAGEs der Pull-down-Assays	121
11.2.1.	MMP16	121
11.2.2.	ARPC2	123
11.3.	Massenspektrometrie Daten	126
11.4.	Publikationen	135

1. EINLEITUNG

1.1. GUANIN-QUADRUPLExE

Guanin-Quadruplexe (G-Quadruplexe) sind ungewöhnliche Sekundärstrukturen in Nucleinsäuren. Bereits 1910 beobachtete der Chemiker Ivar Bang, dass Guanosinmonophosphat-Lösungen in millimolaren Konzentrationen ein Gel bilden (1). Die Ausbildung von helikalen Strukturen durch Guanosin-5'-Monophosphat wurde erstmals 1962 von Gellert et al. beschrieben (2). Anhand von röntgenspektroskopischen Analysen schlugen sie eine Struktur vor, in der vier Guanine durch Wasserstoffbrücken in einer Ebene gepaart sind. Diese planaren Strukturen können sich durch Van-der-Waals-Kräfte stapeln und bilden somit eine Helixstruktur. Das Vorkommen von G-Quadruplex-Strukturen in Nucleinsäuren wurde Ende der 1980er Jahre entdeckt als ein ungewöhnliches Laufverhalten G-reicher DNA Oligonucleotide mit einer Sequenz aus den Telomeren in nativen Gelen beobachtet wurde (3). In Nucleinsäuren können Sequenzen mit vier Abfolgen von mindestens zwei Guaninen Quadruplexe bilden. Dabei ordnen sich vier Basen in einer Ebene an. Diese planaren Ebenen können durch π - π Wechselwirkungen einen Stapel bilden und formen so das G-Quadruplex (4). Während die bekannteste Sekundärstruktur der Nucleinsäuren, die α -Doppelhelix, durch Watson-Crick Basenpaarungen zustande kommt, werden G-Quadruplexe durch Hoogsteen-Wechselwirkungen stabilisiert. Zwischen den vier Guaninen jeder Ebene werden insgesamt acht Wasserstoffbrückenbindungen ausgebildet (Abbildung 1). Im Gegensatz dazu wird ein Watson-Crick Basenpaar durch zwei oder drei Wasserstoffbrücken stabilisiert. Wie daher zu erwarten, zeigen G-Quadruplexe eine große thermale Stabilität und könnten in der Lage sein auch *in vivo* mit der Ausbildung einer Doppelhelix zu konkurrieren.

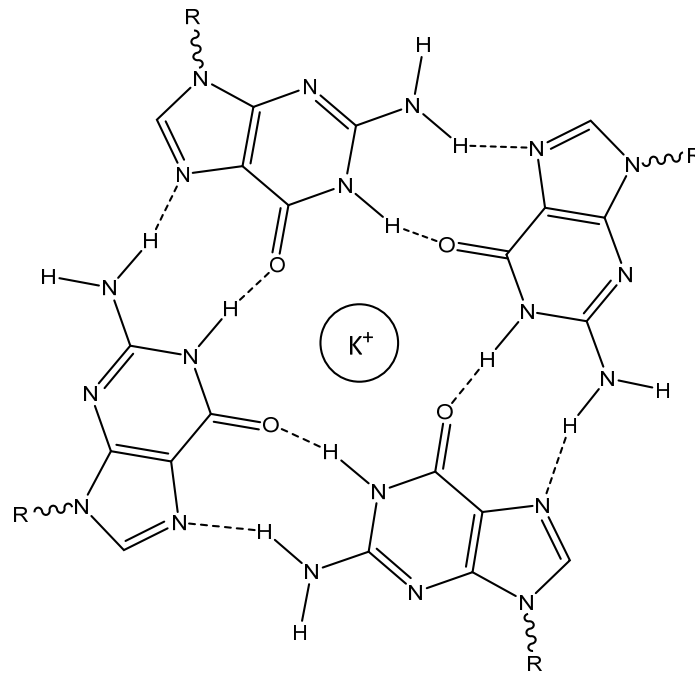


Abbildung 1: Struktur eines Guanin-Quadruplexes. Vier Guanin-Basen sind durch Hogsteen-Wechselwirkungen miteinander gepaart. Die vier Guanine bilden insgesamt acht Wasserstoffbrückenbindungen aus. Im Zentrum der Struktur neutralisiert ein Kation die negativen Ladungen der Sauerstoffatome.

Kationen spielen eine wichtige Rolle bei der Stabilisierung der G-Quadruplexe. Sie sitzen im Zentrum der Struktur und neutralisieren die negativen Ladungen der Guanin-Sauerstoffatome, die sich ansonsten abstoßen würden. Die unterschiedliche Fähigkeit der Kationen G-Quadruplexe zu stabilisieren wurde schon früh entdeckt (5). So stabilisiert das Kalium-Ion die Struktur aufgrund seines Ionenradius am besten, in einem geringeren Maß wirken auch Rb^+ , Cs^+ und Na^+ stabilisierend. Aus der Gruppe der Alkalimetalle sind Li^+ Ionen am wenigsten in der Lage G-Quadruplexe zu stabilisieren. Des Weiteren wird die Stabilisierung durch K^+ auch als physiologisch relevant angesehen. Die intrazelluläre K^+ -Konzentration ist mit etwa 140 mM recht hoch, im Gegensatz dazu beträgt die Konzentration an Na^+ -Ionen nur etwa 10 mM (6).

Je nach Direktionalität der einzelnen Nucleinsäurestränge können verschiedene G-Quadruplex Topologien entstehen. Haben alle vier Stränge die gleiche Orientierung bilden sie ein paralleles G-Quadruplex. Verlaufen zwei Stränge in die eine, zwei in die andere Richtung ist es ein antiparalleles G-Quadruplex. Hierbei sind auch alle Varianten von gemischten Strukturen möglich. Mit der Direktionalität der Stränge geht auch die Konformation der glykosidischen Bindungen zwischen Base und Zucker in *syn* oder *anti* einher. Bei parallelen G-Quadruplexen liegen die glykosidischen Bindungen in *anti*, bei antiparallelen liegen sie sowohl in *syn*- als auch in *anti*-Konformation vor. Die Konformation der glykosidischen Bindung beeinflusst die Größe und Topologie der Furchen die durch das Zucker-Phosphat-Rückgrat begrenzt werden (4, 7). Die Sequenzen zwischen den einzelnen Guanin-Abfolgen bilden die sogenannten Loops. Sie sind in der Regel 1-7 Nucleotide lang. Man unterscheidet drei

verschiedene Typen: laterale Loops, die benachbarte, antiparallele Zucker-Phosphat-Rückgrate verbinden, diagonale Loops, die gegenüberliegende, antiparallele Stränge auf derselben Oberfläche verbinden und Propeller-Loops, die benachbarte parallele Stränge auf der gegenüberliegenden Oberfläche verbinden. Die Loops haben einen Einfluss auf die Stabilität der Struktur. Je kürzer der Loop, desto stabiler sind die G-Quadruplex-Strukturen im Allgemeinen (8). Die Anzahl der Guanin-Quartette die den Stapel bilden sowie die Strangstöchiometrie, also aus wie vielen Strängen das G-Quadruplex zusammengesetzt ist, haben ebenfalls einen Einfluss auf die Stabilität.

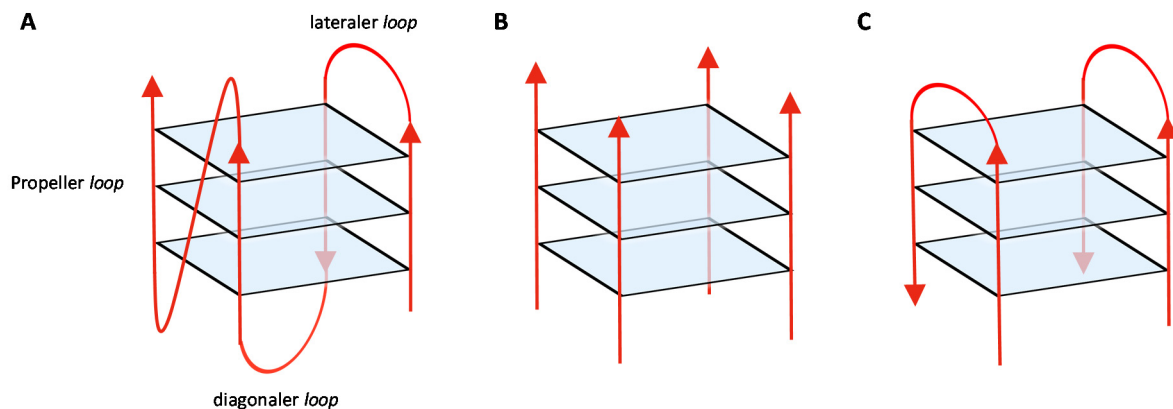


Abbildung 2: Topologie von G-Quadruplexen. (A) Mögliche Loop Typen eines G-Quadruplexes: der Propeller Loop verbindet zwei benachbarte, parallele Stränge auf den gegenüberliegenden Oberflächen, der diagonale Loop zwei gegenüberliegende antiparallele Stränge auf derselben Oberfläche und der laterale Loop zwei benachbarte, antiparallele Stränge auf derselben Oberfläche (B) tetramolekulares, paralleles G-Quadruplex (C) bimolekulares, antiparalleles G-Quadruplex

RNA G-Quadruplexe sind meist stabiler als DNA G-Quadruplexe (9). Für RNA G-Quadruplexe wurden ausschließlich parallele Strukturen beschrieben, während DNA G-Quadruplexe vielfältiger in ihren Konformationen sind. Die humane Telomersequenz gehört zu den am besten untersuchten G-Quadruplex formenden Sequenzen. Sie besteht aus tausenden von Tandem-Wiederholungen der Abfolge (GGGTTA)_n (10). Obwohl es immer dieselbe Sequenz ist, wurde eine Vielzahl an möglichen G-Quadruplex-Strukturen beschrieben, die sie einnehmen kann (11). Einige der möglichen Strukturen sind in Abbildung 3 dargestellt. Hier lässt sich gut erkennen, wie unterschiedlich die Struktur je nach Anordnung der einzelnen Stränge sein kann.

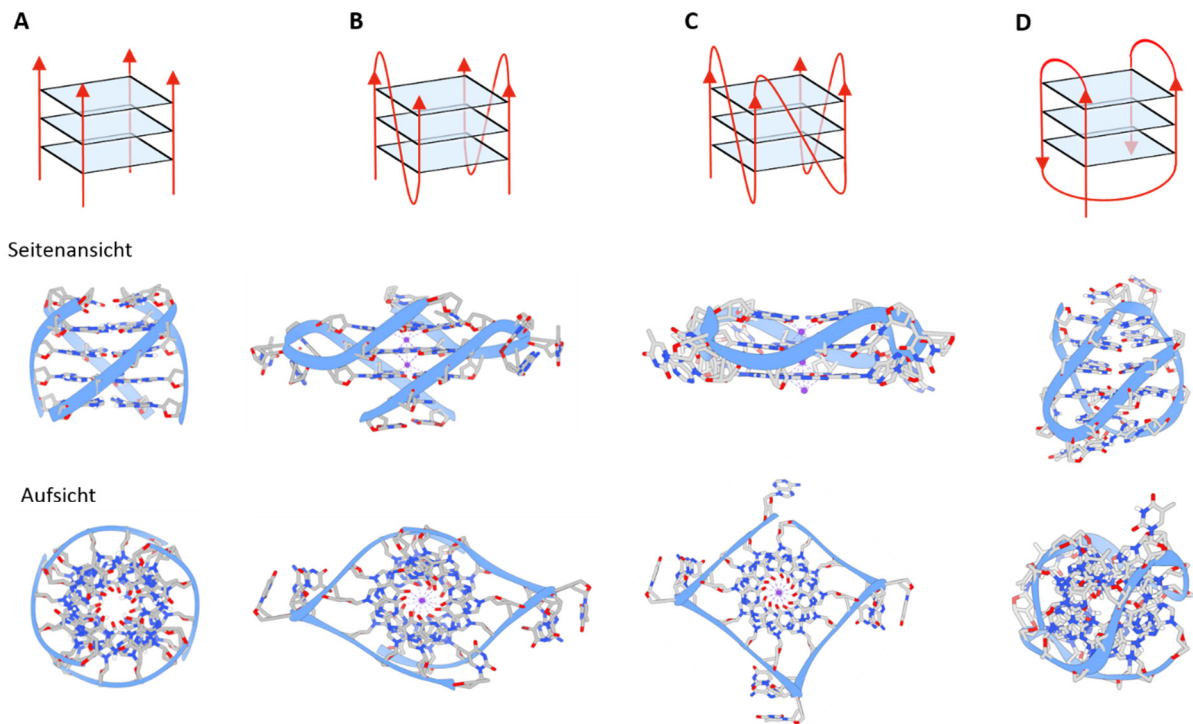


Abbildung 3: Beispiele verschiedener telomerischer G-Quadruplexe. Oben sind die Strukturen schematisch dargestellt, in der mittleren Zeile in der Seitenansicht und unten in der Aufsicht. **(A)** Tetramolekulare, parallele Struktur (PDB-ID 1NP9) **(B)** Bimolekulare, parallele Struktur mit Propeller-Loops (PDB-ID 1K8P) **(C)** Unimolekulare, parallele Struktur mit Propeller-Loops (PDB-ID 1KF1) **(D)** Unimolekulare, gemischte Struktur mit diagonalen und lateralen Loops (PDB-ID 143D).

Zum Nachweis der Ausbildung einer G-Quadruplex-Struktur stehen verschiedene Methoden zur Verfügung. Zur Charakterisierung werden häufig CD-Spektren (12) und UV-Schmelzkurven (13) aufgezeichnet, in denen die G-Quadruplexe jeweils charakteristische Profile zeigen. Ein Vorteil hierbei ist die schnelle und einfache Durchführbarkeit der Messungen. Ein Nachteil ist, dass diese Methoden nur Hinweise auf die Ausbildung einer Struktur liefern, für eine eindeutige Beschreibung der Topologie der Struktur müssen NMR (*nuclear magnetic resonance*, Kernspinresonanz) oder Röntgenkristallanalysen durchgeführt werden. Die Vielfältigkeit der Strukturen, die durch eine Sequenz ausgebildet werden können, erschweren jedoch diese Analysen. Dennoch wurden bereits einige G-Quadruplex-Strukturen über NMR-Spektroskopie aufgeklärt, z.B. die aus den Promotorregionen von KIT1 (14), Bcl-2 (15), MYC (16) und VEGF (17). Röntgenkristallstrukturen sind hauptsächlich von telomerischen DNA und RNA G-Quadruplexen bekannt (18–20).

1.2. VORKOMMEN VON G-QUADRUPLEXEN

In der Zelle findet man eine Anhäufung von potentiell G-Quadruplex bildenden Sequenzen in Telomeren, Promotorregionen von Genen und untranslatierten Regionen (UTRs) von messenger RNAs (mRNAs). Am längsten bekannt sind die telomerischen G-Quadruplexe, hier gab es auch den ersten Nachweis einer Ausbildung *in vivo* durch einen spezifischen Antikörper (21).

1.2.1. G-QUADRUPLExE IN TELOMEREN

Die Enden humaner Chromosomen, die Telomere, bestehen aus Tausenden von Wiederholungen der Sequenz GGGTTA, sowie einem G-reichen Strang, der einen 100-200 Nucleotide langen 3'-Überhang formt. Telomere schützen die Enden linearer Chromosomen vor Abbau und bewahren die genomische Integrität. Nach jeder Zellteilung verkürzen sich die Telomere und stehen daher im Verdacht, an der Alterung der Zellen beteiligt zu sein. Die Telomerase wirkt dieser Verkürzung entgegen, indem sie durch reverse Transkription ihres RNA-Anteils (*telomerase RNA component*, TERC) die 3'-Überhänge verlängert. Die Telomerase ist in sich kontinuierlich teilenden Zellen wie Keimzellen oder Stammzellen, sowie in den meisten Tumorzellen aktiv (22). In normalem, ausdifferenziertem Gewebe wird in der Regel keine Telomeraseaktivität nachgewiesen. Der G-reiche 3'-Überhang der Telomere kann einzelsträngig vorliegen, dann wird er durch das Protein hPOT1 stabilisiert, oder er kann G-Quadruplex-Strukturen ausbilden. Diese G-Quadruplexe hemmen die Telomeraseaktivität, die Enden der Chromosomen können nicht verlängert werden, weshalb die Stabilisierung dieser Strukturen durch kleine Moleküle ein interessanter neuer Ansatz zur Krebstherapie ist. G-Quadruplexe scheinen noch weitere Rollen bei der Erhaltung der Telomere zu spielen. Der RNA-Anteil der Telomerase (TERC), der als Matrize für die Verlängerung der Enden dient, kann an seinem 5'-Ende ebenfalls ein G-Quadruplex ausbilden, das Einfluss auf die Aktivität des Proteins hat (23, 24). Des Weiteren wird der C-reiche Strang der Telomere transkribiert, wobei die *telomeric repeat containing RNA* (TERRA) entsteht. Diese nicht kodierenden RNAs sind 100-9000 Nucleotide lang und an einer Reihe von zellulären Prozessen wie der Regulierung der Telomerlänge und der Inhibierung der Telomerase beteiligt und haben auch einen Einfluss auf die globale Genexpression in der Zelle (25, 26).

Die Ausbildung von G-Quadruplexen *in vivo*, in den Telomeren von Ciliaten, konnte bereits 2001 durch einen spezifischen Antikörper gezeigt werden (21). Ciliaten besitzen einen polyploiden Makronukleus, der Millionen von Chromosomen und damit auch von Telomeren enthält. Diese große Anhäufung von Telomeren ermöglichte die Visualisierung durch einen fluoreszenzmarkierten Antikörper. Die Ausbildung von G-Quadruplexen in Säugerzellen konnte lange Zeit nur durch indirekte Methoden wie ihren Einfluss auf Transkription oder Translation gezeigt werden. 2013 konnte mit einem G-Quadruplex-spezifischen Antikörper die Ausbildung von G-Quadruplexen sowohl an den Enden eukaryotischer Chromosomen als auch in diskreten Punkten außerhalb der Telomerregion gezeigt werden (27). Die Ausbildung der G-Quadruplex-Strukturen war abhängig vom Zellzyklus, während der Replikation konnten am meisten G-Quadruplex-Strukturen nachgewiesen werden. Des Weiteren ließ sich die Zahl der identifizierten Strukturen durch einen stabilisierenden Liganden erhöhen. Mit demselben Antikörper wurde auch die Ausbildung von RNA G-Quadruplexen im Cytoplasma nachgewiesen (28).

1.2.2. G-QUADRUPLExE IN PROMOTOREN

Da die DNA in der Zelle als Doppelstrang vorliegt, müsste die Watson-Crick Paarung zur Ausbildung von G-Quadruplexen aufgelöst werden. Dies geschieht während den Prozessen der Replikation, Rekombination und Transkription. Eine Anreicherung von potentiellen G-Quadruplexen wurde tatsächlich in Promotorregionen gefunden. Über 40 % der humanen Genpromotoren enthalten mindestens eine potentielle G-Quadruplex-Sequenz (29). Sie sind besonders häufig in der Region 100 Basen vor dem Transkriptionsstartpunkt. Ihr Auftreten korreliert auch mit dem Vorkommen von *nuclease hypersensitive elements* (NHEs) in der DNA. Das sind Bereiche, die besonders für die Spaltung durch Nukleasen oder chemische Modifizierungen anfällig sind. Man nimmt an, dass diese Bereiche entwundener DNA den Zugang trans-aktiver Faktoren zu Promotorbereichen erlauben und damit auf transkriptional aktive Bereiche hindeuten. Ein kritischer Punkt bei der Ausbildung von G-Quadruplexen in zellulärer DNA ist das Vorliegen der DNA als Doppelstrang. Die G-Quadruplex-Struktur würde immer mit der helikalen Struktur konkurrieren. Da Promotorregionen zur Transkription entwunden werden, ist dies ein Bereich in dem sich G-Quadruplexe ausbilden könnten. Das erste Promotor G-Quadruplex, für das eine Rolle in der Transkription beschrieben wurde, war das des Protoonkogens MYC (30). Dieses G-Quadruplex hemmt die Transkription des MYC Genes, das in bis zu 80 % der soliden Tumoren überexprimiert ist. Eine Punktmutation in der G-Quadruplex-Sequenz konnte die Transkriptionsaktivität steigern, während eine weitere Stabilisierung der Struktur durch einen kleinen Liganden die Aktivität des Promotors reduzierte. Interessanterweise zeigte eine genomweite bioinformatische Studie, dass G-Quadruplexe in Promotorregionen von Protoonkogenen angereichert sind, während sie in Promotoren von Tumorsuppressorgenen unterrepräsentiert sind (31). Weitere Beispiele für Protoonkogene mit G-Quadruplexen in der Promotorregion sind Bcl-2, NRAS, KRAS, KIT, VEGF, HIF1A und TERT. Dieses Expressionsmuster macht G-Quadruplexe zu interessanten Zielstrukturen für die Tumorthherapie. Die Strukturen von G-Quadruplexen sind einerseits sehr vielfältig, durch die verschiedenen Strangdirektionalitäten und Loops, die z.B. die Größe der Furchen beeinflussen, auf der anderen Seite sind die planar gestapelten G-Tetraden ein gemeinsames strukturelles Element. Die meisten der Liganden, die an G-Quadruplexe binden, basieren auf heteroaromatischen Ringsystemen und interagieren mit der planaren Oberfläche der G-Tetraden, wodurch sie keine hohe Selektivität zeigen.

1.2.3. RNA G-QUADRUPLExE

RNA G-Quadruplexe sind im Vergleich zu Telomer und Promotor G-Quadruplexen weniger gut untersucht, jedoch wird die Rolle, die sie in vielen Schritten des RNA-Metabolismus spielen, zunehmend deutlicher. Im Gegensatz zu den meisten DNA-Molekülen besitzen RNA-Stränge keinen

Gegenstrang, was die Ausbildung eines G-Quadruplexes erleichtern könnte. Etwa 7 % aller bekannten proteincodierenden Gene besitzen eine potentielle G-Quadruplex bildende Sequenz in ihrer 5'-UTR und etwa 10 % in der 3'-UTR (32).

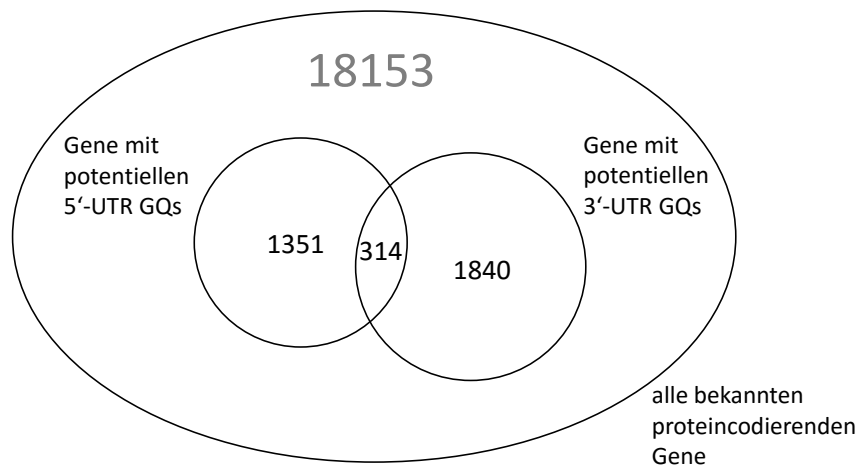


Abbildung 4: Vorkommen von potentiellen G-Quadruplex (GQ) bildenden Sequenzen in 5'- untranslatierten Regionen und in 3'-untranslatierten Regionen von Genen bezogen auf alle bekannten proteincodierenden Gene (verändert nach (32)).

Der Einfluss eines G-Quadruplexes in der 5'-UTR einer mRNA auf die Translation wurde erstmals für das Protoonkogen NRAS gezeigt. Die Struktur reduzierte die Translation eines Reportergens in einem *in vitro* Translationsassay um etwa 70 % im Vergleich zu einer mutierten Kontrolle (33). Die inhibierende Wirkung eines G-Quadruplexes auf die Translation in eukaryotischen Zellen wurde kurz darauf für die G-Quadruplex-Sequenz aus der 5'-UTR der mRNA des Zic-1 Zinkfingerproteins demonstriert (34). Die isolierte G-Quadruplex-Sequenz wurde in die 5'-UTR eines Reportergens eingefügt und bewirkte ebenfalls eine Hemmung der Translation um ca. 70 %, während der Gehalt der mRNA in der Zelle konstant blieb. Inzwischen wurden noch viele weitere Proteine beschrieben, die ein G-Quadruplex in der 5'-UTR besitzen, welches die Translation inhibiert. Einige Beispiele sind in Tabelle 1 aufgeführt.

Tabelle 1: Beispiele für 5'-UTR G-Quadruplexe, für die eine Translationsinhibierung beschrieben wurde.

Gen	Protein	Sequenz 5'-3'	Ref.
ADAM10	Disintegrin und Metalloprotease	GGGGGACGGGUAGGGGCGGGAGGUAGGGG	(35)
Bcl-2	Protoonkogen B-cell lymphoma 2	GGGGGCCGUGGGGUGGGAGCUGGGG	(36)
CCND3	<i>Cyclin D3</i>	GGGCGGCGGGCGGGCUGGGG	(37)
ERS1	Estrogen receptor 1	GGGUAGGGGCAAAGGGGCTGGGG	(38)
MMP16	Matrixmetalloproteinase	GAGGGAGGGAGGGAGAGGGA	(39)
NRAS	GTPase NRAS	GGGAGGGGCGGGUCUGGG	(33)
ZIC1	Zinc finger protein Zic1	GGGUGGGGGGGCGGGGAGGCCGGGG	(34)

Bei der Initiation der Translation bindet die kleine Untereinheit des Ribosoms (40S Untereinheit) zusammen mit verschiedenen Initiationsfaktoren an das 5'-Cap der mRNA. Dieser Komplex scannt dann in 5' → 3'-Richtung entlang der mRNA bis das Startcodon erkannt wird, hier bindet die große (60S) an die kleine Untereinheit und sie bilden das katalytisch aktive 80S Ribosom (40). Die G-Quadruplex-Struktur in der 5'-UTR könnte das Binden von Initiationsfaktoren an die RNA stören, das Scannen beeinflussen oder das Erkennen des Startcodons verhindern (41). Ein Beispiel für solch eine sterische Blockade ist das *iron responsive element*, eine Haarnadelstruktur, die sich in der 5'-UTR der mRNA des Proteins Ferritin befindet. In der Abwesenheit von Eisen binden die *iron regulatory proteins* (IRE 1 und 2) an das IRE und die Translation wird gehemmt. Es wurde gezeigt, dass diese Hemmung abhängig von der Position des Elements ist. Das IRE befindet sich 40 nt nach dem 5'-Cap, wird es weiter in Richtung Startcodon verschoben, ist die Hemmung der Translation nicht mehr effizient. Dies liegt wahrscheinlich daran, dass die Struktur die Bindung der 40S Untereinheit an das Cap blockiert. Wird das Element verschoben, ist wieder genug Platz für das Binden der kleinen Untereinheit an das 5'-Cap vorhanden (42). Befindet sich ein G-Quadruplex am 5'-Ende der 5'-UTR, könnte die Hemmung der Translation auf einem ähnlichen Mechanismus basieren. Eine Inhibition der Translation wurde aber nicht nur für G-Quadruplexe gezeigt, die sich nahe am 5'-Cap befinden, sondern auch für G-Quadruplexe an anderen Positionen in der 5'-UTR. Es scheinen also auch andere Prozesse beeinflusst zu werden. Zum Einfluss der Position des G-Quadruplexes auf die Inhibition der Translation gibt es zwei widersprüchliche Studien. Kumari et. al. zeigten, dass das NRAS G-Quadruplex die Translation nur inhibiert, wenn es sich in den ersten 50 Basen nach dem Cap befindet (43), während Halder et al. eine künstliche G-Quadruplex-Sequenz an verschiedenen Positionen einer 5'-UTR inserierten und keinen Einfluss der Position fanden (44). Sie konnten jedoch zeigen, dass das Maß der Inhibition stark von der Stabilität der G-Quadruplex-Struktur abhängt.

Im Kontrast zu den vielen 5'-UTR G-Quadruplexen, deren inhibitorische Wirkung auf die Translation beschrieben wurde, stehen die G-Quadruplexe denen eine Rolle innerhalb einer *internal ribosome entry site* (IRES) zugeschrieben wurde und die damit aktivierend auf den Translationsprozess wirken. In diesem IRES vermittelten, Cap-unabhängigen Translationsinitiationsmechanismus bindet die 40S Untereinheit des Ribosoms oder sogenannte *IRES trans acting factors* (ITAFs) direkt an die IRES Struktur in der 5'-UTR, und die Translation wird auf diese Weise initiiert. Dieser Mechanismus, der ursprünglich für die virale Genexpression beschrieben wurde, spielt eine Rolle während zellulärer Prozesse, in denen die Cap-abhängige Translation beeinträchtigt ist, wie zum Beispiel während der Mitose, Zelldifferenzierung, Apoptose, unter hypoxischen Bedingungen oder bei Nährstoffmangel (45). G-Quadruplexe sind an den IRES-Elementen des *fibroblast growth factor 2* (FGF2) und des *vascular endothelial growth factor* (VEGF) beteiligt (46, 47). Für die VEGF IRES wurde kürzlich gezeigt, dass das G-Quadruplex ein essentieller Bestandteil ist, der direkt mit der ribosomalen 40S Untereinheit

interagiert. Die Vorhersage solcher Elemente ist schwierig, da bis heute keine allgemeingültige Struktur oder Sequenz zellulärer IRES-Elemente bekannt ist.

Auch für 3'-UTR G-Quadruplexe wurde ein inhibierender Einfluss auf die Translation demonstriert. Das G-Quadruplex in der 3'-UTR des Protoonkogens Pim-1 hemmte dessen Translation um etwa 50 % (48). Für die meisten 3'-UTR G-Quadruplexe wurden jedoch andere Funktionen beschrieben. Die zwei mRNAs der Proteine PSD-95 und CaMKIIa besitzen G-Quadruplex-Sequenzen in ihrer 3'-UTR, die für die Lokalisierung dieser mRNAs in Neuronen verantwortlich sind (49). Dieses mRNA *targeting* ist dabei wichtig für die lokale Proteinsynthese. Hierbei werden mRNAs zum Beispiel in die Dendriten oder Axone von Neuronen transportiert, und die Proteine werden direkt am Wirkungsort translatiert. Dieser Mechanismus wurde für mehrere neuronale mRNAs beschrieben (50–52). Man nimmt an, dass das G-Quadruplex trans-aktive Faktoren rekrutiert, die dann eine Verbindung zu Motorproteinen, wie den Kinesinen, herstellen. Die G-Quadruplex-Sequenzen aus der 3'-UTR der neuronalen Proteine PSD-95 und CaMKIIa waren in der Lage eine Reporter-mRNA in die Dendriten zu lokalisieren. Wurden die G-Quadruplex-Sequenzen aus der 3'-UTR deletiert, reduzierte sich der Anteil der lokalisierten mRNA auf ca. 10 %. Potentielle G-Quadruplex-Sequenzen wurden in etwa 30 % der dendritischen mRNAs gefunden.

Eine weitere Rolle spielen G-Quadruplexe in der alternativen Polyadenylierung und damit einhergehend in der Verkürzung von 3'-UTRs von mRNAs (53). So können regulierende Elemente durch Verkürzung der 3'-UTR entfernt werden, z.B. wird durch alternative Polyadenylierung die *FXR1 (fragile X-related mental retardation autosomal homolog 1)* mRNA verkürzt und damit eine miRNA Bindestelle entfernt, was letztendlich zu einer erhöhten Genexpression führt. Die Interaktion des *heterogeneous nuclear ribonucleoprotein H/F (hnRNP H/F)* mit potentiell G-Quadruplex bildenden Sequenzen fördert die alternative Polyadenylierung.

Nicht nur auf die eukaryotische Genexpression haben G-Quadruplexe einen Einfluss, immer häufiger werden sie auch als regulatorische Elemente in viralen Nucleinsäuren identifiziert. So wurde eine G-reiche Sequenz in der codierenden Region der *Epstein-Barr virus-encoded nuclear antigen 1 (EBNA1)* mRNA beschrieben, die G-Quadruplexe ausbilden kann. Auf diese Weise wird die Translation des Proteins EBNA1 herunterreguliert, was für die latente Infektion mit dem Virus wichtig sein könnte. Ein hoher Gehalt an Virusprotein führt zu einer schnellen Erkennung des Virus durch das Immunsystem. Wurde die G-Quadruplex-Struktur mit einem Antisense Oligonucleotid destabilisiert, erhöhte dies die Proteinexpression und damit auch die frühzeitige Erkennung durch Immunzellen (54). Des Weiteren wurden regulatorische G-Quadruplexe in den Nucleinsäuren von HIV, Simian Virus 40 und dem humanen Papillomavirus beschrieben (55).

Ein weiteres Gebiet, in dem G-Quadruplexen zunehmend eine Rolle zugeschrieben wird, ist das alternative Spleißen. Alternatives Spleißen trägt erheblich zur Diversität unseres Genoms bei. Das gesamte Ausmaß wird erst nach und nach bekannt. Man nimmt an, dass beinahe jede mRNA, die aus mehreren Exonen besteht, alternativ gespleißt wird (56). Über die Regulation der alternativen Spleißprozesse ist noch wenig bekannt. Es ist anzunehmen, dass nicht jeder der über 100.000 alternativen Spleißprozesse von spezifischen Faktoren reguliert wird, und es daher allgemeine regulatorische Proteine geben muss. So werden immer wieder SR-(*serine-arginine-rich*) Proteine und hnRNPs (*heterogeneous nuclear ribonucleoproteins*) als Regulatoren identifiziert (57). Es wird angenommen, dass stabile Sekundärstrukturen in der prä-mRNA Spleißstellen verdecken können und so das Spleißen beeinflussen (58). Auf diese Weise könnten G-Quadruplexe auf das Spleißen einwirken. Auf der anderen Seite könnten sie als Erkennungsstrukturen für Proteine dienen, die dann entweder aktivierend oder inhibierend auf die Nutzung einer Spleißstelle wirken könnten. Das *fragile X mental retardation protein* (FMRP) bindet an eine G-Quadruplex-Struktur in seiner eigenen mRNA und wirkt so als Exon-Spleiß-Enhancer, führt also zur Inklusion des entsprechenden Exons (59). Auch für das humane Telomerase (hTERT) Gen wurde gezeigt, dass die Stabilisierung einer G-Quadruplex-Struktur Einfluss auf das Spleißen hat. Die Stabilisierung des G-Quadruplexes durch einen Liganden in einer Lungenkarzinomzelllinie führte zur Herunterregulierung der Telomeraseaktivität durch eine Veränderung des Spleißens hin zur katalytisch inaktiven Form von hTERT (60). Ein weiteres Beispiel ist das *β -Site amyloid precursor protein cleaving enzyme 1* (BACE1), das die Spaltung des *β -Site amyloid precursor proteins* katalysiert, welches an der Alzheimer Krankheit beteiligt ist. Hier ist eine G-Quadruplex-Struktur in Exon 3 an der Auswahl der Spleiß-Stelle beteiligt. Durch Mutation dieser G-Quadruplex-Sequenz wurde die Nutzung der konstitutiven Spleißstelle reduziert und eine katalytisch inaktive Variante des Proteins gebildet (61).

1.3. G-QUADRUPLIX BINDENDE PROTEINE

Da G-Quadruplexe an so vielen wichtigen Schritten der Genexpression beteiligt sind, liegt es nahe anzunehmen, dass es *trans*-aktive Faktoren gibt, die auf diese Regulation Einfluss nehmen. Während RNA G-Quadruplex bindende Proteine noch weniger untersucht sind, sind bereits eine Reihe an DNA G-Quadruplex bindenden Proteinen bekannt. Dies könnte damit zusammenhängen, dass die DNA Struktur schon länger bekannt und daher besser untersucht ist. Diese G-Quadruplex Binder zeigen oft hohe Affinitäten und Spezifitäten zu ihren Zielstrukturen. Die Existenz dieser spezifisch bindenden Proteine kann auch Hinweise auf die Relevanz bzw. auf das Vorkommen von G-Quadruplexen *in vivo* liefern. Tabelle 2 zeigt eine Auswahl an bereits identifizierten DNA oder RNA G-Quadruplex bindenden Proteinen.

Tabelle 2: Eine Auswahl an bereits beschriebenen DNA und RNA G-Quadruplex bindenden Proteinen.

Protein	Erkennungssequenz	Funktion	Referenz
<i>Nucleolin</i>	MYC Promotor, DNA, parallel, unimolekular	stabilisierend	(62)
<i>Lipopolysaccharid responsive factor LR1</i>	Immunglobulin <i>Switch</i> -Region, DNA, parallel	stabilisierend	(63)
<i>Myoblast determination protein (MyoD)</i>	Regulatorische Sequenzen in muskelspezifischen Genen, DNA, parallel		(64, 65)
<i>Fragile X mental retardation Protein (FMRP)</i>	<i>fragile X repeat sequence</i> in der eigenen mRNA, antiparallel; mehrere andere RNA G-Quadruplex-Sequenzen	stabilisierend	(66) (67)
<i>Heterogenous nuclear ribonucleoprotein A1 (hnRNP A1)</i>	KRAS Promotor G-Quadruplex, DNA, parallel	destabilisierend	(68)
hnRNP A1, hnRNP A2, CBF-A	<i>fragile X repeat sequence</i> , RNA, antiparallel		(69)
<i>Protection of Telomere 1 (Pot1)</i>	Telomerische G-Quadruplexe, DNA	destabilisierend	(70)
<i>RNA Helicase associated with AU-rich element (RHAU)</i>	RNA-Komponente der Telomerase (TERC), verschiedene weitere G-Quadruplex-Strukturen, uni- und tetramolekular	Helikase	(24, 71)
<i>Werner syndrome ATP-dependent helicase (WRN-Helikase)</i>	antiparallel, DNA, bimolekular	Helikase	(72)
<i>Bloom syndrome protein, (BLM-Helikase)</i>	parallel, DNA, unimolekular	Helikase	(73)

Einen Hinweis auf die Bedeutung von G-Quadruplexen *in vivo* liefern zwei Krankheiten, deren Ursache in der fehlenden Expression der BLM-Helikase (Bloom-Syndrom) und der WRN-Helikase (Werner-Syndrom) liegen (74). Sie werden mit deren fehlender DNA G-Quadruplex Helikase-Aktivität in Zusammenhang gebracht. Das Bloom-Syndrom geht einher mit einem erhöhten Auftreten von Leukämien, soliden Tumoren und Sarkomen, während das Werner-Syndrom durch frühzeitige Alterung und dem Auftreten damit einhergehender Erkrankungen wie Diabetes, Arteriosklerose und Osteoporose gekennzeichnet ist. Beide Erkrankungen gehören zu den Chromosomenbruchsyndromen. Die Helikasen könnten für die Auflösung von DNA G-Quadruplex-Strukturen verantwortlich sein, die während genomischer Prozesse wie Replikation oder Rekombination entstehen (72, 73).

Nucleolin wurde erstmals als Teil des heterodimeren Proteins LR1 als G-Quadruplex bindendes Protein beschrieben (63). Dieses B-Zell-spezifische Protein, bestehend aus Nucleolin und hnRNP D, konnte mit hoher Affinität an die G-Quadruplex DNA aus dem nichtkodierenden Strang der Immunglobulin *Switch*-Region binden. In dieser Region findet durch Rekombination der Klassenwechsel von Immunglobulinen statt, dabei zeigen diese Regionen eine Neigung zur G-Quadruplex Bildung. Des Weiteren wurde Nucleolin über eine Affinitätschromatographie mit HeLa Zelllysats *in vitro* als Binder des MYC Promotor G-Quadruplexes identifiziert. Nucleolin konnte das MYC G-Quadruplex in spektrometrischen Messungen stabilisieren, und die Überexpression von Nucleolin reduzierte die MYC Promotoraktivität in einem Luciferase-Reporter-Assay (62).

Das wahrscheinlich am besten untersuchte RNA G-Quadruplex bindende Protein ist das *fragile X mental retardation protein* (FMRP). Die gestörte Expression dieses Proteins steht im Zusammenhang mit dem Fragilen-X-Syndrom, einer der häufigsten Ursachen erblicher kognitiver Behinderung. FMRP bindet an ein G-Quadruplex in der codierenden Region seiner eigenen mRNA. Eine *in vitro* Selektion FMRP bindender RNA-Sequenzen zeigte, dass es bevorzugt Sequenzen mit G-Quadruplex Motiven bindet (67). Durch Immunpräzipitation von FMRP enthaltenden Ribonucleoproteinkomplexen aus Mäusegehirn konnte gezeigt werden, dass etwa 70 % der assoziierten mRNAs G-Quadruplex-Sequenzen enthielten (75). Es sind zwei mRNAs bekannt, an die FMRP bindet, die ein G-Quadruplex in der 5'-UTR besitzen, *MAP1B* und *PP2A*. Beide Proteine werden in Abwesenheit von FMRP stärker exprimiert, was darauf hindeutet, dass FMRP an die 5'-UTR G-Quadruplexe bindet und durch deren Stabilisierung die Translation hemmt (76, 77).

Die hohe thermodynamische Stabilität von G-Quadruplexen *in vitro* deutet darauf hin, dass auch *in vivo* spezielle Proteine benötigt werden, um diese Strukturen aufzulösen. Die humane DEAD-Box-RNA-Helikase RHAU (*RNA Helicase associated with AU-rich element*) ist in der Lage G-Quadruplex-Strukturen *in vitro* aufzulösen (71). Sie kann sowohl tetramolekulare als auch unimolekulare DNA oder RNA G-Quadruplexe auflösen, zeigt jedoch eine höhere Affinität für die RNA-Struktur. Eine *in vitro* Analyse der Ziel-RNAs von RHAU zeigte, dass die Mehrheit G-Quadruplex-Sequenzen enthält. Es fand sich eine Anreicherung der RNA-Komponente der Telomerase, TERC. RHAU war in der Lage über TERC mit dem Telomerase Holoenzym zu interagieren (24). Ein *in vivo* Beweis für die G-Quadruplex Helikase Aktivität von RHAU fehlt allerdings noch.

1.4. IN DIESER ARBEIT UNTERSUCHTE 5'-UTR G-QUADRUPELXE

1.4.1. MEMBRANE-TYPE MATRIX METALLOPROTEINASE 3

Matrix Metalloproteinasen (MMPs) sind Zink-abhängige Endopeptidasen, die extrazelluläre Matrixproteine abbauen. Die Gruppe der MMPs besteht aus 16 sekretierten und 7 membrangebundenen Proteasen. Die Funktionen der MMPs sind vor allem assoziiert mit dem Umbau der extrazellulären Matrix. Sie reichen von der Morphogenese während der embryonalen Entwicklung, der Ossifikation der Knochen und der Angiogenese über Prozesse zur normalen Erhaltung von Geweben wie der Wundheilung und -Reparatur oder der angeborenen Immunantwort (78). Die hier untersuchte *matrix metalloproteinase 16* (*Membrane-type matrix metalloproteinase 3*, MMP16) wird in vielen Tumorgeweben hoch exprimiert. Dies wurde z.B. für Nierenkarzinome (79), Prostatakrebs (80) und Melanome (81) gezeigt. Die hohe Expression korreliert auch mit der Invasivität der Tumore, wo die Fähigkeit zum Umbau der extrazellulären Matrix wichtig ist (82).

In der MMP16 5'-UTR wurde ein G-Quadruplex beschrieben welches die Translation inhibiert und so als cis-regulatorisches Element für die Expression des Proteins wirkt. Die Ausbildung einer parallelen G-Quadruplex-Struktur *in vitro* wurde durch CD-Spektroskopie, in der G-Quadruplexe ein charakteristisches Spektrum zeigen, sowie durch RNase T1 *footprinting* Experimente nachgewiesen. In einem *in-cellulo*-Reporter-Assay inhibierte das G-Quadruplex in der Vollängen-UTR die Translation um etwa 60 % im Vergleich zu einer Kontrolle, in der nur die G-Quadruplex-Sequenz mutiert war (39). Des Weiteren konnte das Porphyrin TmPyP4 die Struktur destabilisieren. Dies konnte sowohl anhand des CD-Spektrums als auch durch erhöhte Translation einer Luciferase im Reporter-Assay gezeigt werden (83).

1.4.2. ACTIN RELATED PROTEIN 2/3 COMPLEX SUBUNIT 2

Actin related protein 2/3 complex subunit 2 (ARPC2) ist eine der 7 Untereinheiten des ARP2/3-Komplexes (*actin related protein 2/3* Komplex). Er ist beteiligt an der Verzweigung und an der Nukleation von Aktinfilamenten. Der Komplex besteht aus den Untereinheiten ARP2 und ARP3, die strukturell verwandt mit monomerem Aktin sind, sowie den Proteinen ARPC1-5.

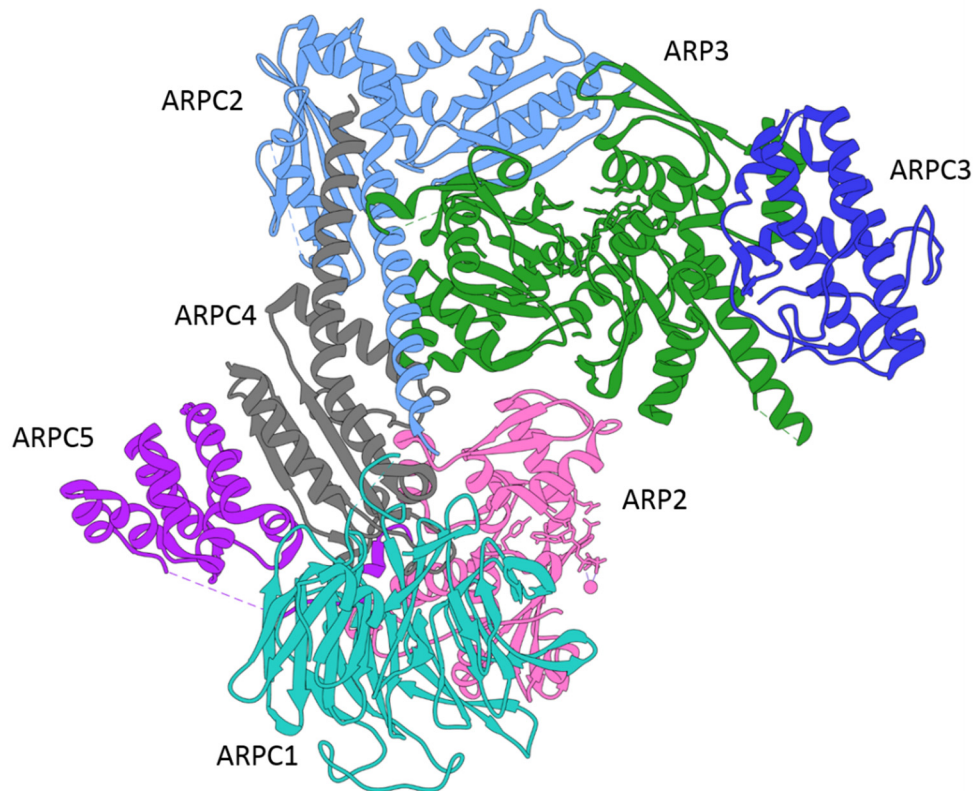


Abbildung 5: Röntgenkristallstruktur des ARP2/3 Komplexes aus *Bos taurus* (PDB-ID 1TYQ). Der Komplex besteht aus den sieben Untereinheiten ARP2, ARP3 und aus den ARP-Komplex Untereinheiten ARPC1-5.

Der ARP2/3-Komplex ist beteiligt an der Verzweigung des Aktins, das als F-Aktin (filamentöses Aktin) ein lineares Polymer bildet. Die zwei Untereinheiten ARP2 und ARP3 weisen eine hohe Homologie zu Aktin auf. Sie können zusammen mit Aktinmonomeren einen Nukleationskeim für die Bildung eines neuen Aktinpolymers bilden. Zur Verzweigung des Aktinnetzwerks bindet der ARP2/3-Komplex an das Aktin-Mutterfilament und initiiert die Bildung eines Tochterfilaments, welches sich dann in einem charakteristischen 70° Winkel abzweigt. Die Proteine ARPC2 und ARPC4 bilden den Kontakt zum Mutterfilament, die genauen Funktionen der Proteine ARPC1, 3 und 5 sind unbekannt. Lokalisierung und Funktionen des ARP2/3-Komplexes in der Zelle sind in Abbildung 6 gezeigt.

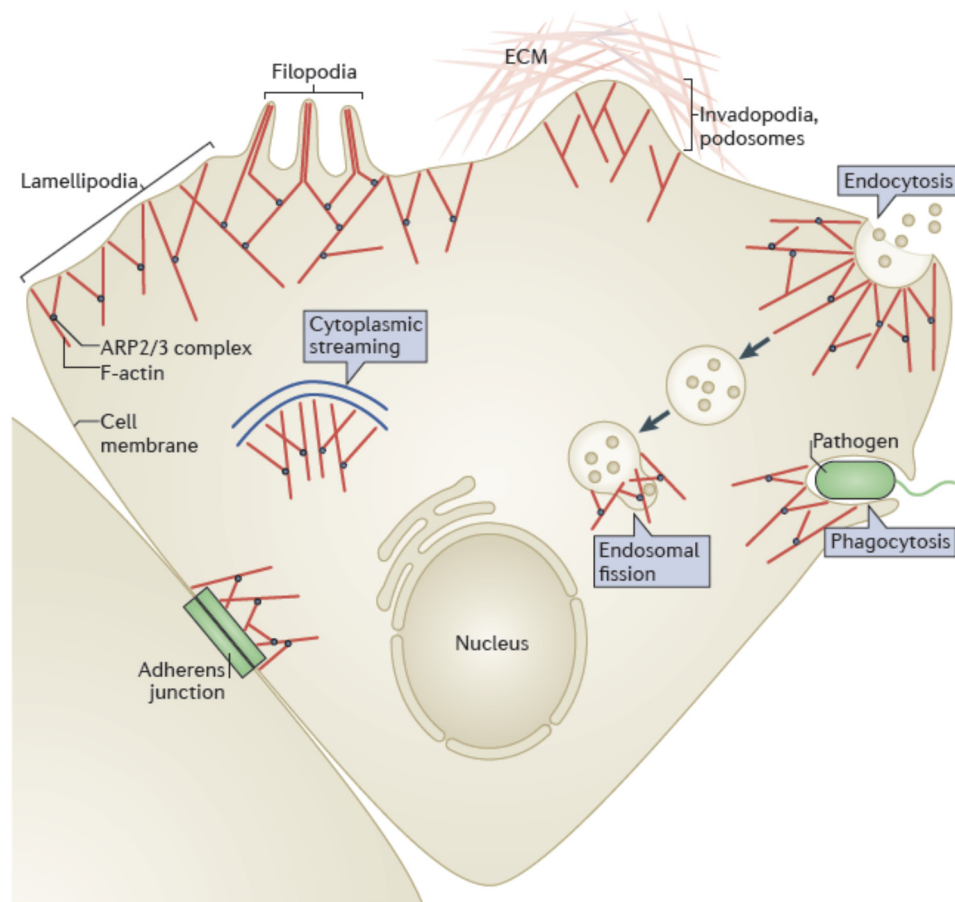


Abbildung 6: Zelluläre Funktionen des ARP2/3-Komplexes (aus 79). Der Komplex ist beteiligt an der Migration von Zellen durch die Ausbildung von Lamellipodien und Filopodien, an der Aufnahme von Partikeln oder Pathogenen durch Endocytose und Phagozytose, an der Cytoplasmaströmung und am Aufbau von Invadopodien und Podosomen, die der Zellinvasion und dem Abbau extrazellulärer Matrix dienen. ARP2/3 ist auch lokalisiert an *adherens junctions* durch die Zell-Zell-Kontakte ausgebildet werden.

Der Komplex ist an Prozessen beteiligt, die mit der Umstrukturierung des Cytoskeletts und daraus resultierenden Bewegungsprozessen assoziiert sind. Ein Fehlen des ARP2/3-Komplexes hat daher Einfluss auf die Migration von Zellen, wobei hier verschiedene Effekte beschrieben wurden. Es gibt Studien, die zeigen, dass ARP2/3 defiziente Mausfibroblasten keine Lamellipodien ausbilden und daher nicht migrieren (85). Andere beschreiben, dass nur die gerichtete Bewegung beeinträchtigt ist (86, 87) oder dass die Abwesenheit von ARP2/3 keinen Einfluss auf die Migration hat (88). Alle sieben mRNAs des Komplexes werden an die Lamellipodien oder Filopodien transportiert und dort translatiert, was die lokale Expression des ARP2/3-Komplexes zu regulieren scheint (89). Wird die Lokalisation der Untereinheit ARP2 gehemmt, stört dies die gerichtete Migration (90).

Es wurde gezeigt, dass die Aktivität des ARP2/3-Komplexes durch den epidermalen Wachstumsfaktor (*epidermal growth factor*, EGF) reguliert wird (91). Die Stimulation des EGF-Rezeptors führt zur Weiterleitung des Wachstumssignals über die Phosphorylierung von Tyrosinresten in Zielmolekülen. Ein Haupteffekt der Stimulierung durch EGF ist der Umbau des Cytoskeletts. In der Studie, die die

Zielproteine der EGF-abhängigen Phosphorylierung in einem globalen Ansatz identifizierte, wurde die Phosphorylierung aller Untereinheiten des ARP2/3-Komplexes nach Stimulierung durch EGF nachgewiesen. Die Aktivierung zeigte ein Maximum 5 Minuten nach der EGF-Stimulierung, sowie einen weiteren Anstieg nach 20 min.

Für sich alleine zeigt der ARP2/3-Komplex eine geringe Aktivität, für seine Aktivierung sind sogenannte *nucleation promoting factors* (NPFs) notwendig. Es ist eine Reihe dieser aktivierenden NPFs bekannt von denen die meisten zu den Wiskott-Aldrich Syndrom Proteinen (WASP) gehören. Neben den aktivierenden Faktoren sind jedoch auch Proteine bekannt die den ARP2/3-Komplex inhibieren. Des Weiteren könnte die ARP2/3 Aktivität durch funktionale Interaktionen mit anderen Aktin-bindenden Proteinen reguliert werden.

ARPC2 besitzt in seiner 5'-UTR eine G-reiche Sequenz, die das Potential hat ein G-Quadruplex zu bilden. Interessanterweise werden von ARPC2 in der NCBI *Nucleotide* Datenbank (<http://www.ncbi.nlm.nih.gov/nucore>) zwei Spleißvarianten vorhergesagt, die sich lediglich in ihren 5'-UTR Sequenzen unterscheiden (Abbildung 7). Die eine Variante enthält das potentielle G-Quadruplex und ist etwas länger, die zweite Variante ist kürzer und das G-Quadruplex wurde herausgespleißt.

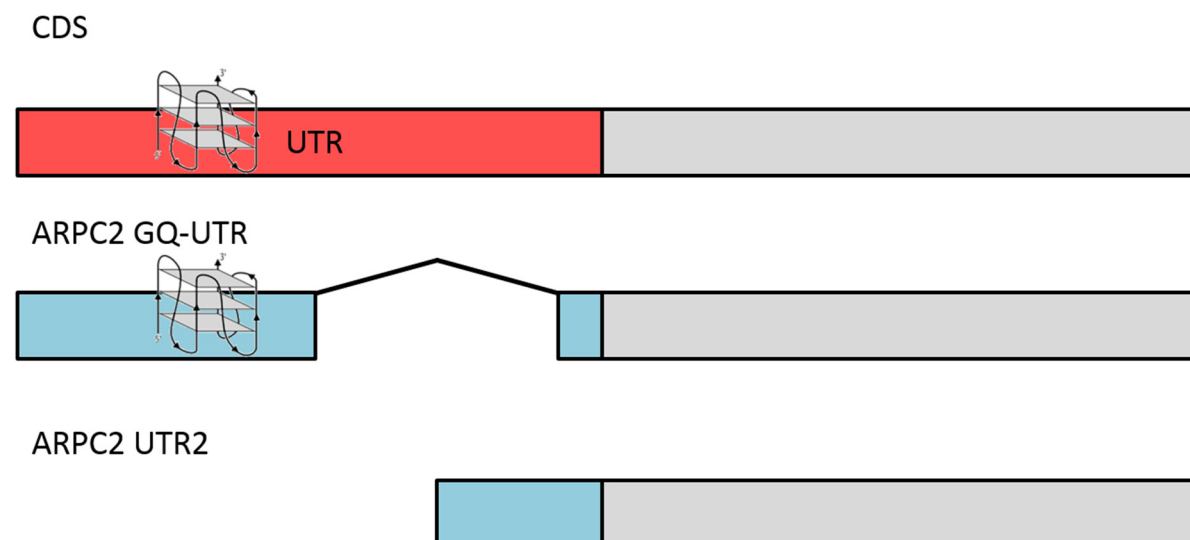


Abbildung 7: 5'-UTR Spleißvarianten des ARPC2 Gens. Die obere Darstellung zeigt schematisch die codierende Sequenz (CDS), die beiden Spleißvarianten sind darunter dargestellt. Die ARPC2 GQ-UTR Variante enthält die G-Quadruplex-Sequenz, dahinter ist ein Stück der UTR Sequenz bis kurz vor dem Startcodon herausgespleißt. Die ARPC2 UTR2 Variante enthält nur den hinteren Teil der UTR Sequenz. Die beiden Sequenzen sind nur in den 8 Basen vor dem Startcodon identisch.

Eine weitere Auffälligkeit des ARP2/3-Komplexes ist, dass sechs der sieben Untereinheiten eine G-Quadruplex-Sequenz in ihrer 5'-UTR oder in ihrer 3'-UTR besitzen. Die Untereinheiten ARP3, ARPC2 und ARPC5 besitzen ein potentielles G-Quadruplex in ihrer 5'-UTR, ARP2, ARPC1B und ARPC4 in der 3'-UTR. ARP2 besitzt sogar zwei G-reiche Sequenzen in der 3'-UTR. Keine der Untereinheiten hat eine G-Quadruplex-Sequenz in der 3'- und der 5'-UTR.

2. ZIELSETZUNG

G-Quadruplexe können als regulatorische Elemente in Nucleinsäuren dienen. Während ihre Funktionen in Telomeren und auch in DNA Promotoren schon länger erforscht werden, wird ihre Bedeutung in mRNAs erst nach und nach erkannt. G-Quadruplexe in der 5'-UTR einer mRNA können unter anderem die Translation inhibieren und bieten damit eine zusätzliche Möglichkeit zur Regulation der Genexpression. Es gibt Beispiele, wie das des G-Quadruplex aus der 5'-UTR der Zic-1 mRNA, in denen die Translation um bis zu 80 % gehemmt wird. Hier liegt es nahe anzunehmen, dass eine solche starke Inhibition auch regulatorischen Faktoren unterliegt. Ein Ziel dieser Arbeit war es daher, Proteine zu identifizieren, die spezifisch an RNA G-Quadruplexe binden können. Dies sollte am Beispiel zweier G-Quadruplexe aus der MMP16 (*matrix metalloproteinase 16*) und ARPC2 (*actin related protein 2/3 complex subunit 2*) 5'-UTR erfolgen. So sollten spezifische, aber auch generelle G-Quadruplex bindende Proteine gefunden werden.

Die beiden ausgewählten G-Quadruplexe sollten in der verwendeten Zelllinie exprimiert werden, um sicherzugehen, dass auch die entsprechenden regulatorischen Proteine gebildet werden. Deshalb sollte die Expression der entsprechenden mRNAs zunächst durch RT-PCR nachgewiesen werden. Während das MMP16 G-Quadruplex schon in der Literatur beschrieben wurde, war über die ARPC2 G-Quadruplex-Sequenz noch nichts bekannt. Daher sollte die Ausbildung eines G-Quadruplexes durch diese Sequenz mittels CD-Spektroskopie sowie durch die Aufnahme von UV-Schmelzkurven nachgewiesen werden. Des Weiteren sollte der Einfluss, sowohl der isolierten Struktur, als auch der Gesamt-UTR Sequenz, auf die Translation in einem Luciferase-Reporter-Assay bestimmt werden.

G-Quadruplex bindende Proteine sollten in Pull-down-Assays isoliert und anschließend massenspektrometrisch identifiziert werden. Um diese Bindungen näher zu charakterisieren, sollten ausgewählte Proteine heterolog in *Escherichia coli* produziert und über einen Polyhistidin-Tag aufgereinigt werden. Die Interaktionen zwischen Proteinen und G-Quadruplexen sollten durch Electrophoretic Mobility Shift Assays sowie Oberflächenplasmonresonanzmessungen näher charakterisiert werden.

Da die G-Quadruplexe nach ihrem Einfluss auf die Translation ausgewählt wurden, sollte überprüft werden, ob die gefundenen Proteine die Inhibition beeinflussen können. Dafür sollten *in vitro* Translationsassays sowie *in cellulo* Co-Transfektionen mit Reporter-Plasmiden durchgeführt werden. Von ARPC2 werden in der NCBI *Nucleotide* Datenbank zwei Spleißvarianten vorhergesagt, die sich nur in ihrer 5'-UTR unterscheiden und von denen nur eine das G-Quadruplex enthält. Die Expression dieser beiden Varianten sollte nachgewiesen und deren Regulation untersucht werden.

Der ARP2/3 Komplex besteht aus 7 Untereinheiten, von denen eine ARPC2 ist. Sechs dieser Untereinheiten besitzen eine G-Quadruplex-Sequenz in der 5'- oder 3'-UTR. Der Einfluss dieser G-Quadruplexe auf die Translation sollte bestimmt werden.

3. MATERIAL UND METHODEN

3.1. MATERIALIEN

3.1.1. GERÄTE

Brutschränke	C 150 (Binder GmbH, Tuttlingen) HERAcell 240 (Thermo Scientific Biosciences, St. Leon-Roth)
CD-Spektrometer	J-810 Spektropolarimeter (Jasco, Groß-Umstadt)
Drehrad	Stuart Rotator SB3 (Bibby Scientific, Staffordshire, UK)
Elektrophoresekammern (Polyacrylamidgele)	Mini-Protean Tetra System (Bio-Rad Laboratories, München)
Elektrophoresekammern (Agarosegele)	HE33 Mini Subunit (Hoefer, Holliston, USA) Wide Mini Sub Cell (Bio-Rad Laboratories, München)
Feinwaage	Satorius Universal (Satorius, Göttingen)
Fluoreszenzmikroskop	Axiovert 200 M (Carl Zeiss, Jena)
Geldokumentation	UV Solo TS Imaging System (Biometra, Göttingen) ChemiDoc XR+Imaging System (Bio-Rad Laboratories, München)
Luminometer	Lumat LB 9507 (Berthold, Bad Wildbad)
Massenspektrometer	Ultraflex-II TOF/TOF (Bruker Daltonics, Bremen, Germany)
Mikroskop	Zeiss AxioVision (Carl Zeiss, Jena)
Oberflächenplasmonresonanz-Spektrometer	Biacore-X, Biacore-2000 (GE Healthcare, Freiburg)
qPCR Gerät	CFX96 Real Time System (Bio-Rad Laboratories, München)
Schüttelinkubator	Innova 42 (New Brunswick Scientific, Eppendorf, Hamburg)

Spektrophotometer	Nano Drop 2000 (Thermo Scientific Biosciences, St. Leon-Roth)
Thermocycler	Personal Thermocycler (Biometra, Göttingen)
Thermomixer	Eppendorf comfort (Eppendorf, Hamburg)
Ultraschallsonde	Sonifier 250 (Branson, Danbury, USA)
UV Photometer	V-650 mit Peltier Element (Jasco, Groß-Umstadt)
Vakuumzentrifuge	Concentrator Plus (Eppendorf, Hamburg)
Zellkulturbank	MSC Advantage 1.8 (Thermo Scientific Biosciences, St. Leon-Roth)
Zentrifugen	Heraeus Biofuge Fresco (Rotor: 3325B) (Thermo Scientific Biosciences, St. Leon-Roth) Sorvall RC-5B PLUS (Rotor: GSA) (Thermo Scientific Biosciences, St. Leon-Roth) Tischzentrifuge, Centrifuge 5415 C (Eppendorf, Hamburg) Sigma 3-16PK (Rotor: 11133) (Sigma, Osterode am Harz)

3.1.2. CHEMIKALIEN, PUFFER UND KITS

Alle Standardchemikalien wurden, soweit nicht anders angegeben, von den Firmen Carl Roth (Karlsruhe), Merck (Darmstadt), Roche Diagnostics (Risch, CH) und Sigma-Aldrich (St. Louis, USA) bezogen.

BAKTERIENKULTIVIERUNG

Ampicillin	Stammlösung: 100 mg/mL in H ₂ O Endkonzentration: 100 µg/mL
Kanamycin	Stammlösung: 50 mg/mL in H ₂ O Endkonzentration: 50 µg/µL
LB-Agarplatten	1,5 % Euro-Agar, 1 % Bacto-Trypton, 0,5 % Hefeextrakt, 0,5 % NaCl
LB-Medium	1 % Bacto-Trypton, 0,5 % Hefeextrakt, 0,5 % NaCl

HERSTELLUNG CHEMISCH KOMPETENTER E. COLI

Lösung 1	100 mM Rubidiumchlorid, 35 mM Kaliumacetat, 10 mM Calciumchlorid, 0,5 mM Lithiumchlorid, 15 % (w/v) Glycerin, 45 mM Manganchlorid, pH 5,8
Lösung 2	10 mM MOPS, 75 mM Calciumchlorid, 10 mM Rubidiumchlorid, 15 % (w/v) Glycerin, pH 7

KULTIVIERUNG EUKARYOTISCHER ZELLEN

Alle Medien und Medienzusätze wurden bezogen von BioWest SAS (Nuaille, Frankreich).

Ablösen adhärenter Zellen	Trypsin-EDTA Lösung
Medium für HEK293 Zellen	500 mL DMEM (Dulbecco's Modified Eagle Medium, Low Glucose) 10 % FKS (Fötales Kälber Serum) 5 mL 100x Penicillin/Streptomycin 5 mL L-Glutamin (Endkonzentration 2 mM) 5 mL Non Essential Amino Acids (NEAA) 3,5 mL Glucose (Endkonzentration 4,5 mg/mL)
Medium für HeLa Zellen	500 mL DMEM (Dulbecco's Modified Eagle Medium, Low Glucose) 10 % FKS (Fötales Kälber Serum) 5 mL 100x Penicillin/Streptomycin 5 mL L-Glutamin (Endkonzentration 2 mM) 5 mL Non Essential Amino Acids (NEAA)
Medium für MCF7 Zellen	500 mL RPMI 10 % FKS (Fötales Kälber Serum) 5 mL 100x Penicillin/Streptomycin 5 mL L-Glutamin (Endkonzentration 2 mM) 5 mL Non Essential Amino Acids (NEAA) 5 mL Pyruvat
Puffer für die Zellkultur	DPBS (Dulbecco's Phosphate Buffered Saline)

Transfektionsreagenzien	<i>Lipofectamine</i> 2000 (Life Technologies, Darmstadt) Polyethylenimin (PEI), (Polysciences, Eppelheim)
-------------------------	--

KLONIERUNGEN, PCR UND PLASMIDISOLATIONEN

Dephosphorylierung	Fast Alkaline Phosphatase, (Thermo Scientific Biosciences, St. Leon-Roth) Antarctic Phosphatase (New England Biolabs, Frankfurt am Main)
DNA Ligation	T4 DNA Ligase (Thermo Scientific Biosciences, St. Leon-Roth)
DNA-Größenstandard	1 KB DNA Ladder, 100 bp DNA Ladder (New England Biolabs, Frankfurt am Main)
DNA-Laufpuffer	
10 x TBE	1 M TRIS-Base, 890 mM Borsäure, 20 mM EDTA, pH 8 mit NaOH
50 x TAE	2 M TRIS-Base, 100 mM EDTA, pH 8 mit Essigsäure
DNA-Polymerasen	Phusion DNA Polymerase (Thermo Scientific Biosciences, St. Leon-Roth) Pfu DNA Polymerase (Promega, Mannheim) Taq (Rapidozym, Berlin)
DNase	RQ1 DNase (Promega, Mannheim)
Gel Extraktion	Invisorb Fragment CleanUp (Stratec, Berlin)
Oligonucleotid Annealing-Puffer	100 mM Kaliumacetat, 30 mM HEPES-KOH pH 7,4, 2 mM Magnesiumacetat
PCR Aufreinigung	QIAquick PCR Purification Kit (Qiagen, Hilden)
Phosphorylierung	Polynucleotidkinase (Thermo Scientific Biosciences, St. Leon-Roth)
Plasmid Midipräparation	Nucleobond Xtra Midi Plus Kit (Macherey-Nagel, Düren)
Plasmid Minipräparation	Invisorb Spin Plasmid Mini Two Kit (Stratec, Berlin)
Restriktionsenzyme	New England Biolabs (Frankfurt am Main)
Reverse Transkriptase	RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific Biosciences, St. Leon-Roth)
RNA Extraktion	RNeasy Mini Kit (Qiagen, Hilden)

LUCIFERASE-REPORTER-ASSAY

Dual-Luciferase-Assay	Dual-Luciferase-Reporter-Assay System (Promega, Mannheim)
Renilla-Luciferase-Assay	Renilla-Luciferase-Assay Kit (Promega, Mannheim)

PULL-DOWN-ASSAY

10 x SDS Laufpuffer	0,25 M TRIS, 1,92 M Glycine, 1 % SDS, pH ~ 8,4
4 x TCA Lösung	40% (w/v) Trichloressigsäure in H ₂ O
5 x Laemmli Probenpuffer	250 mM TRIS-HCl pH 6,8
	500 mM DTT
	10 % (w/v) SDS
	10 % (v/v) Glycerin
	0,1 % (w/v) Bromphenolblau
Blockingpuffer	10 mM TRIS-HCl pH 7,5
	100 mM KCl
	0,1 mM Na-EDTA
	1 mM DTT
	0,01 % (v/v) Triton X-100
	0,1 % (w/v) nukleasefreies BSA
	0,02 % (w/v) tRNA aus <i>Sacharomyces cerevisiae</i>
Coomassie Entfärbelösung	10 % Ethanol, 2 % ortho-Phosphorsäure
Dialyseschlauch	ZelluTrans (Carl Roth, Karlsruhe)
Faltungspuffer	10 mM TRIS-HCl pH 7,5
	100 mM KCl
Kolloidale Coomassie Färbelösung	0,02% Coomassie (CBB-G250), 5% Aluminiumsulfat-(14-18)-Hydrat, 10% Ethanol und 2% ortho-Phosphorsäure
Proteaseinhibitor	Complete (Roche, Mannheim)
Protein-Marker	SeeBlue Plus2 Pre-Stained Standard (Life Technologies, Darmstadt)
	Unstained Protein Molecular Weight Marker (Thermo Scientific Biosciences, St. Leon-Roth)
Proteinreinigungspuffer A	10 mM TRIS-HCl pH 7,5
	10 mM KCl
	1,5 mM MgCl ₂
	0,5 mM DTT

Proteinreinigungspuffer B	300 mM Tris-HCl pH 7,5 1,4 M KCl 3 mM MgCl ₂
Proteinreinigungspuffer C	20 mM TRIS-HCl pH 7,5 100 mM KCl 20 % (v/v) Glycerin 0,5 mM DTT 0,2 mM Na-EDTA
Pull-down-Waschpuffer (verschiedene KCl- Konzentrationen)	10 mM TRIS-HCl pH 7,5 1 mM DTT 0,1 mM Na-EDTA 0,01 % Triton X-100 400, 600, 800, 1200, 2000 und 2600 mM KCl
RNase Inhibitor	Ribolock (Thermo Scientific Biosciences, St. Leon-Roth)
SDS Polyacrylamid-Gele	Sammelgel: 4 % Acrylamid/Bisacrylamid, 0,125 M TRIS-HCl pH 6,8, 0,1 % SDS, 0,1 % APS (Ammoniumperoxodisulfat), 0,1 % TEMED (Tetramethylethyldiamin), Trenngel: 7,5-15 % Acrylamid/Biscrylamid, 0,1% SDS, 0,03 % APS, 0,1 % TEMED, 0,375 M TRIS-HCl pH 8,8
Streptavidin-Agarose-Beads	Streptavidin-Agarose aus <i>Streptomyces avidinii</i> (Sigma- Aldrich, München)

TRYPSIN IN-GEL-VERDAU

Carbamidomethylierung	55 mM Iodacetamid in 100 mM NH ₄ HCO ₃
Reduktionslösung	100 mM DTT in 100 mM NH ₄ HCO ₃
Trypsinlösung	12,5 µg/mL Trypsin (bovin, Sequenzierungs-Reinheitsgrad) in 25 mM NH ₄ HCO ₃ (Trypsin-Stammlösung 1:80 mit Verdaupuffer verdünnt)
Trypsin-Stammlösung	1mg/mL in 1mM HCl (Trypsin sequencing grade, Sigma-Aldrich, München)
Verdaupuffer (Spaltung)	25 mM NH ₄ HCO ₃ pH 7,9 – 8,5

Alle Lösungen wurden in H₂O (HPLC-Reinheitsgrad, Sigma-Aldrich) angesetzt.

HETEROLOGE PROTEINEXPRESSION

Dialysepuffer I	20 mM TRIS pH 7,5 150 mM NaCl 10 % Glycerin 0,1% Triton X-100 1 mM DTT
Dialysepuffer II	10 mM Natriumphosphatpuffer pH 7,5 150 mM NaCl 10 % Glycerin 0,1 % Triton X-100 1 mM DTT
Dialyseschlauch	ZelluTrans (Carl Roth, Karlsruhe)
Expressionsmedium	EnPresso™ Tablet Cultivation Set (BioSilta, Cambridgeshire, Großbritannien)
Leersäulen	Polypropylene Columns, 1 mL (Qiagen, Hilden)
Nickel-NTA Agarose	Protino Ni-NTA Agarose (Macherey-Nagel)
NPI-10	50 mM NaH ₂ PO ₄ 300 mM NaCl 10 mM Imidazol pH 8 einstellen mit NaOH
NPI-20	50 mM NaH ₂ PO ₄ 300 mM NaCl 20 mM Imidazol pH 8 einstellen mit NaOH
NPI-250	50 mM NaH ₂ PO ₄ 300 mM NaCl 250mM Imidazol pH 8 einstellen mit NaOH
Proteinkonzentrationsbestimmung	BCA Protein Assay Kit (Thermo Scientific Biosciences, St. Leon-Roth)

ELECTROPHORETIC MOBILITY SHIFT ASSAY

4 x EMSA Puffer	40 mM TRIS pH 7,5, 4 mM EDTA, 4 mM DTT, 16 µg/mL BSA, 100 mM KCl
Laufpuffer für native PAGE	1 x TRIS-Borat-EDTA Puffer (TBE)
Native Polyacrylamidgele	10 % Acrylamid/Bisacrylamid, 1 x TBE, 0,1 % TEMED, 0,1 % APS
Nucleinsäuredetektion	GelRed (Biotium, Hayward, USA)

OBERFLÄCHENPLASMONRESONANZ-SPEKTROSKOPIE

Immobilisierung	Amine Coupling Kit (GE Healthcare, Freiburg)
Immobilisierungspuffer	HBS-EP (GE Healthcare, Freiburg) 0,01 M HEPES pH 7,4 3 mM EDTA 0,005 % v/v Surfactant P20
Laufpuffer	HBS-EP: 10 mM TRIS-HCl pH 7,5 im Verhältnis 9:1
Sensorchips	CMD-500 L (Xantec, Düsseldorf)

IN VITRO TRANSKRIPTION UND IN VITRO TRANSLATION

<i>In vitro</i> Transkriptionskit	Ampliscibe™ T7-Flash™ Transcription Kit (epicentre, Madison, USA)
<i>In vitro</i> Translationskit	Flexi® Rabbit Reticulocyte Lysate System (Promega, Mannheim)
Reinigung der RNA Produkte aus der Transkription	RNeasy Mini Kit (Qiagen, Hilden)
DNase Verdau	RQ1 DNase (Promega, Mannheim)

RNA Agarosegel

10 x Formaldehyd-Puffer	200 mM 3-[N-morpholino]propanesulfonic acid (MOPS) 50 mM Natriumacetat 10 mM EDTA pH to 7.0 mit NaOH
-------------------------	---

Formaldehyd Agarosegel	1 x Formaldehyd-Puffer
	1 % Agarose
	Nukleasefreies Wasser

SILAC PULL-DOWN-ASSAY

Komponenten für SILAC-Medium	SILAC-DMEM (DMEM ohne Lysin und Arginin)
	Dialysiertes FCS (PAA Laboratories, Coelbe)
	L-Lysin:2HCl, 13C6, 15N2 (EURISO-TOP, Saarbrücken)
	L-Arginin:HCl, 13C6 (EURISO-TOP, Saarbrücken)
Ultrafiltration	Amicon Ultra 0,5 mL <i>Filters</i> 3K (Merck Millipore, Darmstadt)

QUANTITATIVE REAL TIME PCR

Reverse Transkription	Maxima H Minus First Strand cDNA Synthesis Kit with dsDNase (Thermo Scientific Biosciences, St. Leon-Roth)
Polymerase	Maxima Sybr Green qPCR Master Mix (Thermo Scientific Biosciences, St. Leon-Roth)

3.1.3. BAKTERIENSTÄMME

- *Escherichia coli* Top10 (Life Technologies, Darmstadt)
- *Escherichia coli* BL21 (DE3), (Life Technologies, Darmstadt)

3.1.4. ZELLINIEN

Hek293	Human Embrionic Kidney 293 Zellen, ACC 305 (Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Braunschweig)
HeLa	Human cervix carcinoma, ACC57 (DSMZ, Braunschweig)
MCF7	Brust-Adenokarzinom (DSMZ, Braunschweig)

3.1.5. OLIGONUCLEOTIDE

Die verwendeten Oligonucleotide wurden von TIB Molbiol (Berlin) oder Life Technologies (Darmstadt) synthetisiert. Biotinylierte Oligonucleotide wurden von Purimex (Giebenstein) synthetisiert.

Tabelle 3: Oligonucleotide zur Klonierung in den psiCHECK-2 Vektor. Die Überhänge zur Klonierung der Oligonucleotide sind unterstrichen.

Oligonucleotid Name	Sequenz 5' → 3'
ARPC2 mut sense	<u>CTAGAGCCG</u> TAGACTGAGCGAAGACCGAGCTTGT
ARPC2 mut antisense	<u>CTAGACAAGCTCGGTCTTCG</u> CTCAGTCTACGGCT
ARP2a sense	<u>TCGAGGGGATGGGGGGTGGTT</u> CGGGATGGGTGGGC
ARP2a antisense	<u>GGCCGCC</u> CAACCATCCCGAACCACCCCATCCC
ARP2b sense	<u>TCGAGGGGGGGAGGGTAACAATGGGC</u>
ARP2b antisense	<u>GGCCGCC</u> ATTGTTACCCTCCCCC
ARP3 sense	<u>CTAGCGGGAGGAGGCCGCGCGGGCAGGGCGGGG</u>
ARP3 antisense	<u>CTAGCCCCGCCCTGCCCCGCGCGGCCTCCTCCCG</u>
ARPC1B sense	<u>TCGAGGGAAGCGGGGAGAGGGGTCAGGGC</u>
ARPC1B antisense	<u>GGCCGCC</u> CTGACCCCTCTCCCGCTTCCC
ARPC4 sense	<u>TCGAGGGAGTTGGGTTGGGGTGGGC</u>
ARPC4 antisense	<u>GGCCGCC</u> CAACCACTCCC
ARPC5 sense	<u>CTAGCGGGTAGTGGGTTGCTGGGCTGGG</u>
ARPC5 antisense	<u>CTAGCCCAGCCCAGCAACCACTACCCG</u>
3xGGG mutante 3UTR sense	<u>TCGAGGGAAGCAAAGAGAGGGGTCAGGGC</u>
3xGGG mutante 3UTR antisense	<u>GGCCGCC</u> CTGACCCCTCTTTGCTTCCC
3xGGG mutante 5UTR sense	<u>CTAGCGGGAAGCAAAGAGAGGGGTCAGGG</u>
3xGGG mutante 5UTR antisense	<u>CTAGCCCTGACCCCTCTTTGCTTCCCG</u>

Tabelle 4: Primer zur Klonierung in pET28a. Die BamHI und NdeI Restriktionsschnittstellen sind unterstrichen.

Oligonucleotid Name	Sequenz 5' → 3'
Nucleolin fw	GGAATTCC <u>ATATGG</u> TGAAGCTCGCGAAGG
Nucleolin Δ1-284 fw pET	GGAATTCC <u>ATATGG</u> CCAAACAGAAAGCAGC
Nucleolin rev	GCGGGATCCCTATTCAAACCTTCGTCTTC
EFHD2 fw	GGAATTCC <u>ATATGG</u> CACGGACGAGC
EFHD2 rev	GTA <u>GGATCC</u> TCAGGCTCGGCCTCC
SRSF1 fw	GGAATTCC <u>ATATGT</u> CGGGAGGTGGTG
SRSF1 rev	GCGGGATCCCTTATGTACGAGAGCGA
SRSF9 fw	GGAATTCC <u>ATATGT</u> CGGGCTGGGCG
SRSF9 rev	GTA <u>GGATCC</u> TCAGTAGGGCCTGAAAG
U2AF65 fw	GGAATTCC <u>ATATGT</u> CGGACTTCGACGAG
U2AF65 rev	GTA <u>GGATCC</u> CTACCAGAAAGTCCCGG
MAOM fw	GGAATTCC <u>ATATGTT</u> GTCCTCGGTTAAG
MAOM rev	GTA <u>GGATCC</u> CTATTCTGTTATCACAGG
Ybx fw	GGAATTCC <u>ATATGAG</u> CAGCGAGGCCGAG
Ybx rev	GTA <u>GGATCC</u> TTACTCAGCCCCGCCCTG

Tabelle 5: Oligonucleotide zur Klonierung der identifizierten G-Quadruplex bindenden Proteine in pcDNA3.1. Die NheI-Schnittstellen sind unterstrichen.

Oligonucleotid Name	Sequenz 5' → 3'
Nucleolin Δ1-284 fw pc	ATTGCTAGC <u>ATGG</u> CAAACAGAAAGCAGC
Ybx fw pcDNA3.1	ATTGCTAGC <u>ATGAG</u> CAGCGAGGCCGAGAC
EFHD2 fw pcDNA	ATTGCTAGC <u>ATGG</u> CACGGACGAGC
U2AF65 fw pcDNA	ATTGCTAGC <u>ATGT</u> CGGACTTCGACGAG
SRSF1 fw pcDNA	ATTGCTAGC <u>ATGT</u> CGGGAGGTGGTG
SRSF9 fw pcDNA	ATTGCTAGC <u>ATGT</u> CGGGCTGGGCG
Nucleolin fw pcDNA	ATTGCTAGC <u>ATGG</u> TGAAGCTCGCGAAGG
MAOM pcDNA oS	ATTGCTAGC <u>ATGTT</u> GCACATAAAAGAAAAAG
MAOM pcDNA mS	ATTGCTAGC <u>ATGTT</u> GTCCTCGGTTAAG

Tabelle 6: Primer zum Nachweis der Expression von ARPC2 und MMP16 mittels RT-PCR.

Oligonucleotid Name	Sequenz 5'→3'
ARPC2 G-Quadruplex fw	CTAGAGCCGGGGGCTGG
ARPC2 G-Quadruplex rev	CACCTCCAGCAGGATCATGG
ARPC2 P3 fw	ATGATCCTGCTGGAGGTGAA
ARPC2 P3 rev	CCATCACTTTTGTCTCTCC
MMP16_cdr_for	AAAATGGCAAACGTGATGTGG
MMP16_cdr_rev	GGCTCATCTGAGTCAAATGG
MMP16_for	TACTGCTGACCCTGTCTTCG
MMP16_rev	TTCTCCCTCTCTCCCTCTCC

Tabelle 7: Biotinylierte RNA Oligonucleotide. Diese am 3'-Ende biotinylierten Oligonucleotide wurden zur Charakterisierung der G-Quadruplex-Strukturen, zur Identifizierung von Bindungspartnern in Pull-down-Assays und für Interaktionsstudien genutzt.

Oligonucleotid Name	Sequenz 5'→3'
MMP16 GQ	AACGAGGGAGGGAGGGAGAGGGAGAGA-Biotin
MMP16 mut	AACGAGAGAGAGAGAGAGAGAGAGAGA-Biotin
MMP16 GQ kurz	GGGAGGGAGGGAGAGGG-Biotin
MMP16 mut kurz	GAGAGAGAGAGAGAGAG-Biotin
MMP16 2xG	AACGAGAGAGGGAGAGAGAGGGAGAGA-Biotin
ARPC2 GQ	AGCCGGGGGCGUGGGCGGGACCGGGCUUGU-Biotin
ARPC2 mut2	AGCCGUAGACUGAGCGAAGACCGAGCUUGU-Biotin
ARPC2 GQ kurz	GGGGGCGUGGGCGGGACCGGG-Biotin
ARPC2 mut2 kurz	GUAGACUGAGCGAAGACCGAG-Biotin
ARPC2 2xG	AGCCGUAGACUGGGCGAAGACCGGGCUUGU-Biotin
NRAS GQ	GGGAGGGGCGGGUCUGGGUG-Biotin
NRAS mut	GAGAGGAGCGAGUCUGAGUG-Biotin
Bcl-2 GQ	GGGGGCCGUGGGGUGGGAGCUGGGG-Biotin
Bcl-2 mut	GGAGGCCGUGAGAUGAGAGCUGAGG-Biotin

Tabelle 8: Primer für die qPCR für den Nachweis und die Quantifizierung der ARPC2 Spleißvarianten.

Oligonucleotid Name	Sequenz 5'→3'
ARPC2 GQ fw	GTGGGTGTGAGAGCGGAAGTG
ARPC2 GQ rev	GCCTGAACCCGCCTGC
ARPC2 UTR2 fw	CTCCCTTACCCACCTCACC
ARPC2 UTR2 rev	ACTGCTCCGGTTTGTCTCC
ARPC2 CDR	AGAAGAGGGCAAGGAAGGAG
ARPC2 CDR	TGGCTAAAGAGGACCTGTGG

3.2. METHODEN

3.2.1. MIKROBIOLOGISCHE METHODEN

HERSTELLUNG CHEMISCH KOMPETENTER E. COLI

Zur Herstellung kompetenter *E. coli* wurde die Rubidiumchloridmethode verwendet. Hierzu wurden 150 mL einer *E. coli* Kultur in der logarithmischen Wachstumsphase pelletiert (10 min, 4°C, 1500 g), 10 min auf Eis abgekühlt und in 15 mL eiskalter Lösung 1 resuspendiert. Die Suspension wurde 10 min auf Eis inkubiert und wieder abzentrifugiert. Anschließend wurde das Pellet in 6 mL Lösung 2 resuspendiert, zu je 100 µL aliquotiert und sofort in flüssigem Stickstoff eingefroren. Die kompetenten *E. coli* wurden bei -80°C gelagert.

TRANSFORMATION VON PLASMIDEN

Zur Transformation wurden die kompetenten Zellen langsam auf Eis aufgetaut. Die zu transformierende DNA wurde zugegeben und der Ansatz für 20 min auf Eis inkubiert. Es folgte ein Hitzeschock für 1 min bei 42°C. Nach 3 min Inkubation auf Eis wurden 700 µL LB Medium zugegeben. Um die Antibiotikaresistenz auszubilden, wurden die Bakterien mindestens 45 min bei 37°C im Schüttelinkubator inkubiert. Zur Selektion positiver Transformanten wurden die Bakterien auf LB-Agar Platten mit dem entsprechenden Antibiotikum ausplattiert und über Nacht bei 37°C inkubiert.

3.2.2. KULTIVIERUNG EUKARYOTISCHER ZELLEN

KULTIVIERUNG

Die eukaryotischen Zelllinien wurden in der Regel in 75 cm² Kulturflaschen mit dem entsprechenden Medium kultiviert. Sie wurden je nach Konfluenz alle 2 bis 3 Tage passagiert. Die Zellen wurden im Brutschrank bei 37°C und 5 % CO₂ inkubiert.

TRANSFEKTION VON PLASMIDEN IN EUKARYOTISCHE ZELLINIEN

24 h vor der Transfektion wurden die eukaryotischen Zellen in 24-Well-Platten so ausgesät, dass sie nach 24 h 60-70 % konfluent waren. Für die Aussaat wurde antibiotikafreies Medium benutzt. Zur Transfektion mit *Lipofectamine* 2000 wurden 2 µL *Lipofectamine* pro Ansatz sowie die Plasmid-DNA in je 50 µL serumfreien RPMI verdünnt. Die beiden Lösungen wurden gemischt und 20 min bei Raumtemperatur inkubiert. Anschließend wurde der Ansatz langsam auf die Zellen gegeben.

Zur Transfektion mit PEI wurde PEI-Lösung (2,5 mg PEI/mL) und DNA in jeweils 50 µL 150 mM NaCl-Lösung verdünnt. Pro µg zu transfizierender DNA wurde dabei ein µL der PEI-Lösung eingesetzt. Die beiden Ansätze wurden vorsichtig gemischt, 20 min bei Raumtemperatur inkubiert und tropfenweise auf die Zellen gegeben.

3.2.3. KLONIERUNGEN, PCR UND PLASMIDISOLATION

KLONIERUNGEN IN psiCHECK-2

Zur Untersuchung des Einflusses verschiedener Sequenzen auf die Translation wurde der Dual-Luciferase-Reporter-Vektor psiCHECK-2 verwendet (Anhang, Abbildung 41). Die isolierten G-Quadruplex-Sequenzen, die entsprechenden mutierten Sequenzen und die Vollängen-UTRs wurden in die 5'-UTR der Renilla-Luciferase kloniert. Dazu wurde die NheI Schnittstelle benutzt. Für Insertionen der ARP2/3 Komplex Sequenzen in die 3'-UTR wurden die XhoI und NotI Schnittstellen benutzt. Zur Klonierung von kurzen Sequenzen wurden sense und antisense Oligonucleotid zunächst in Annealingpuffer (100 mM KAc; 30 mM HEPES-KOH pH 7,4; 2 mM MgAcetat) 4 min auf 95°C erhitzt und langsam auf Raumtemperatur abgekühlt, um sie zusammenzulagern. Anschließend wurden die Oligonucleotide unter Verwendung einer Polynucleotidkinase phosphoryliert. Der Vektor psiCHECK-2 wurde mit NheI verdaut und anschließend dephosphoryliert (Fast AP). Ca. 50 ng Vektor wurden mit 10 pmol des phosphorylierten Oligonucleotids über Nacht ligiert. Am nächsten Tag wurden die Ansätze in *E. coli* Top10 transformiert. Plasmid DNA aus positiven Transformanten wurde mittels Plasmid Mini Präparation isoliert und per Restriktionsverdau analysiert. Der korrekte Einbau der Oligonucleotide wurde durch Sequenzierung überprüft.

KLONIERUNGEN IN pET28A

Die in den Pull-down-Experimenten identifizierten Proteine wurden für die Produktion in *E. coli* in den Expressionsvektor pET28a kloniert (Anhang, Abbildung 42). Die Sequenzen wurden so kloniert, dass sie einen N-terminalen His-tag erhielten, der der späteren Reinigung des Proteins dienen sollte. Dazu wurde mit dem RNeasy Mini Kit RNA aus Hek293 Zellen isoliert. Nach einem DNase Verdau wurde die mRNA mit dem RevertAid H Minus First Strand cDNA Synthesis Kit mit Oligo dT Primern nach Herstellerangaben in cDNA umgeschrieben. Zur Klonierung wurden die cDNA Sequenzen mit genspezifischen Primern, die die Restriktionsschnittstellen BamHI und NdeI (siehe Tabelle 4) enthielten, amplifiziert. Für die Amplifikation der EFHD2 und Ybx Sequenzen wurde der PCR 1 % DMSO zugesetzt, da die Sequenzen GC reich waren. Der Vektor und die PCR Produkte wurden mit den entsprechenden Restriktionsenzymen verdaut und der Vektor wurde anschließend dephosphoryliert.

Beide DNA Fragmente wurden über ein Agarosegel gereinigt, über Nacht ligiert und in kompetente *E. coli* TopTen transformiert. Plasmid DNA aus positiven Transformanten wurde mittels Plasmid Mini Präparation isoliert und per Restriktionsverdau analysiert. Die erhaltenen Klone wurden durch Sequenzierung überprüft.

KLONIERUNGEN IN pcDNA3.1

Zur Transfektion in eukaryotische Zellen wurden die identifizierten G-Quadruplex bindenden Proteine in den eukaryotischen Expressionsvektor pcDNA3.1 kloniert (Anhang, Abbildung 43). Dazu wurden die für die Proteine codierenden Sequenzen mit Primern, die die NheI und BamHI Schnittstellen enthielten, aus den entsprechenden pET28a Expressionsvektoren amplifiziert. Die Sequenzen der *forward* Primer mit den NheI Schnittstellen sind in

Tabelle 5 aufgelistet, die verwendeten *reverse* Primer waren dieselben wie für die Klonierung in pET28a (Tabelle 4). PCR Produkt und das Plasmid pcDNA3.1 wurden mit den Restriktionsenzymen NheI und BamHI verdaut. Das Plasmid wurde dephosphoryliert und beide DNA Fragmente wurden über ein 1%iges Agarosegel gereinigt. Nach der Gelextraktion mit dem Invisorb Fragment CleanUp Kit wurden die Konzentrationen der Fragmente auf einem weiteren Agarosegel abgeschätzt. Etwa 50 ng Vektor und die dreifache molare Menge an Insert wurden über Nacht bei 16°C ligiert. Am nächsten Tag wurden die Ansätze in *E. coli* Top10 transformiert. Plasmid DNA aus positiven Transformanten wurde mittels Plasmid Mini Präparation isoliert und per Restriktionsverdau analysiert. Die erhaltenen Klone wurden durch Sequenzierung überprüft.

NACHWEIS DER GENEXPRESSION MITTELS RT PCR

Die gesamt RNA aus den verwendeten eukaryotischen Zelllinien wurde mit Hilfe des RNeasy Mini Kits isoliert. Nach einem DNase Verdau wurde die RNA unter Verwendung von random Hexamer Primern revers transkribiert. Als Kontrolle diente ein Ansatz ohne die Zugabe von Reverser Transkriptase. Mit der so erhaltenen cDNA wurden die PCRs zum Nachweis der Genexpression durchgeführt. Die verwendeten Primer sind in Tabelle 6 gelistet. Zur Amplifikation der Sequenzen wurde die Taq Polymerase verwendet. Die erhaltenen PCR Produkte wurden anschließend auf einem Agarosegel aufgetrennt und mit Ethidiumbromid sichtbar gemacht.

3.2.4. BIOPHYSIKALISCHE CHARAKTERISIERUNG DER RNA G-QUADRUPLPLEX OLIGONUCLEOTIDE

AUSBILDUNG DER SEKUNDÄRSTRUKTUR

Zur Ausbildung der G-Quadruplex-Struktur wurde das jeweilige Oligonucleotid in Faltungspuffer 5 min auf 95°C erhitzt. Anschließend wurde es im Thermocycler mit einer Rate von 2°C/min auf Raumtemperatur abgekühlt.

CD-SPEKTROSKOPIE

Die CD-Spektren wurden auf einem Jasco J-810 Spektropolarimeter aufgenommen. Die Oligonucleotide wurden auf eine Konzentration von 5 µM in Faltungspuffer (100 mM KCl) verdünnt. Vor der Messung wurden die Oligonucleotide wie zuvor beschrieben zur Ausbildung des G-Quadruplex erhitzt und langsam abgekühlt. Die Schichtdicke der Küvette betrug 1 mm, die Aufnahmegeschwindigkeit 50 nm/min. Die Spektren wurden bei 20°C aufgenommen.

UV-SCHMELZKURVEN

Die Oligonucleotide wurden in Faltungspuffer (1-100 mM KCl) auf eine Konzentration von 1 µM verdünnt. Die UV-Schmelzkurven wurden auf einem UV-Photometer mit Peltier-Element zur Temperaturkontrolle in Quarzküvetten aufgenommen. Die Proben wurden mit Mineralöl überschichtet. Die Heiz- bzw. Kühlrate betrug 1° C min⁻¹. Jede Messung bestand aus fünf Zyklen aus Heizen und Kühlen, also aus Schmelzen und Annealen der Oligonucleotide.

3.2.5. LUCIFERASE-REPORTER-ASSAY

TRANSFEKTION VON HEK293 ZELLEN UND ZELLYSE

Die Hek293 Zellen für den Luciferase-Assay wurden 24 h vor der Transfektion der psiCHECK-2 Plasmide in 24-Well Platten ausgesät. Pro Well wurden 2x10⁵ Zellen in 500 µL M5 Medium ohne Antibiotikum ausgesät. Die Zellen sollten bei der Transfektion etwa 60-70 % konfluent sein. Zur Transfektion von Plasmiden in eukaryotische Zellen wurde das Transfektionsreagenz *Lipofectamine* 2000 nach Herstellerangaben verwendet. Es wurden in der Regel 0,8 µg des Reporterplasmids transfiziert. Nach 24 h wurde das Medium abgesaugt und die Zellen mit 100 µL *Passive Lysis Buffer* (Promega) bedeckt, zur Lyse wurden die Zellen mindestens 30 min bei Raumtemperatur geschüttelt. Die Luciferase-Aktivität wurde entweder sofort bestimmt oder die Proben wurden bei -80°C gelagert.

BESTIMMUNG DER LUCIFERASE-AKTIVITÄT

Die Bestimmung der Luciferase-Aktivität erfolgte auf dem Lumat LB 9507 (Berthold) mit dem Dual-Luciferase-Assay Kit (Promega). Für jede Messung wurden 20 µL des Zelllysats und 50 µL je Substrat verwendet. Die Aktivität jeder Luciferase wurde über 10 s bestimmt. Jede Probe wurde mindestens dreimal gemessen.

3.2.6. IDENTIFIZIERUNG RNA G-QUADRUPLIX BINDENDER PROTEINE

PULL-DOWN-ASSAY

HEK293 oder HeLa Zellen wurden durch Trypsinierung geerntet, das Pellet wurde einmal mit PBS gewaschen und 1:1 in Proteinreinigungspuffer A aufgenommen. Nach 15 min auf Eis wurden die Zellen abzentrifugiert, in zwei Pelletvolumen Proteinreinigungspuffer A und einem Volumen Proteinreinigungspuffer B aufgenommen, und es wurde 7 x Proteaseinhibitorlösung zugegeben. Die Zellen wurden mit einer Ultraschallsonde auf Eis aufgeschlossen (15 s Ultraschall, 15 s Pause, 6 Wiederholungen, Output 4). Unlösliche Zellbestandteile wurden durch Zentrifugation (20 min, 4 °C, 16000 x g) abgetrennt. Der Überstand wurde über Nacht gegen 2 L Proteinreinigungspuffer C bei 4 °C dialysiert (*Molecular weight cut off*: 12 kDa). Der Proteinextrakt wurde bei – 80°C gelagert.

Die Streptavidin-Agarose-Beads wurden mit BSA und tRNA aus *Saccharomyces cerevisiae* blockiert, um unspezifische Bindungen zu verhindern. Für jeden Ansatz wurden 40 µL (einer 50%igen Suspension) *Beads* eingesetzt. Die *Beads* wurden abzentrifugiert (200 x g, 1 min, 4 °C), der Überstand wurde mit einer 1 mL Spritze mit einer Kanüle mit einem Durchmesser von 0,4 mm abgenommen und die *Beads* wurden dreimal mit 400 µL Waschpuffer [100 mM] KCl gewaschen. Anschließend wurden sie in 400 µL Blockingpuffer aufgenommen und für eine Stunde bei 4°C im Drehrad (10 RpM) inkubiert. Nach dem Blockieren wurden die *Beads* dreimal mit Waschpuffer [300 mM KCl] gewaschen, dazu wurden je 400 µL Puffer zugegeben, 5 min im Drehrad (4 °C, 10 RpM) inkubiert und 1 min abzentrifugiert (200 x g, 1 min, 4 °C). Die Streptavidin-Agarose-Beads wurden in 300 µL Waschpuffer [300 mM KCl] aufgenommen. Je Ansatz wurden 1,4 nmol der biotinylierten RNA Oligonucleotide in 10 mM TRIS-HCl pH 7,5, 100 mM KCl und 0,1 mM Na-EDTA in einem Volumen von 100 µL zugegeben. Die Oligonucleotide wurden vorher im Thermocycler der Quadruplex-Formation unterzogen. Zur Immobilisierung der biotinylierten Oligonucleotide an die Streptavidin-Agarose wurden die Ansätze über Nacht im Drehrad inkubiert (10 rpm, 4 °C). Am nächsten Tag wurden die *Beads* dreimal mit 500 µL Waschpuffer [400 mM KCl] gewaschen. Im Anschluss wurden die *Beads* in 200 µL des Waschpuffers aufgenommen und 300 µL des Proteinextrakts wurden zugegeben. Dem Proteinextrakt wurden vor der Zugabe 90 U RNase Inhibitor beigemischt, um zelluläre RNasen zu inhibieren. Die Bindung der

Proteine an die immobilisierte RNA erfolgte für 1 h bei 30°C im Schüttelinkubator. Der Überstand wurde abgenommen (*Flow through*) und die *Beads* wurden nun mit Waschpuffern mit steigenden KCl-Konzentrationen (400, 600, 800, 1200, 2000 und 2600 mM) gewaschen. Für jeden Waschschrift wurden 500 µL des entsprechenden eiskalten Puffers zugegeben, 5 min im Drehrad bei 4°C inkubiert, die Streptavidin-Agarose-*Beads* wurden herunterzentrifugiert und der Überstand wurde mit einer Spritze abgenommen. Mit jedem Puffer wurde dieser Schritt dreimal wiederholt und die so erhaltenen Fraktionen vereint.

Die eluierten Proteinfractionen wurden im Anschluss auf SDS-PAGEs analysiert. Hierfür wurden die Proteine aus dem Volumen von insgesamt 1,5 mL je Waschschrift durch Trichloressigsäure ausgefällt. Zu jeder Probe wurden 500 µL der 4x TCA Lösung gegeben und die Ansätze wurden eine Minute auf dem Vortexmischer gemischt. Nach 15 min Inkubation auf Eis wurden die gefällten Proteine durch Zentrifugation pelletiert (15000 g, 4°C, 15 min), und das Pellet wurde mit 500 µL eiskaltem Aceton gewaschen. Dieser Waschschrift wurde noch einmal wiederholt. Nach erneutem Zentrifugieren wurde das Aceton abgenommen und die Proteinpellets ca. 30 min bei Raumtemperatur getrocknet. Die Pellets wurden dann in einem möglichst kleinen Volumen (etwa 20 µL) 1 M TRIS-Base pH~10,5 aufgenommen. Es wurde eine entsprechende Menge an 5 x Laemmli Probenpuffer zugegeben. Falls der Puffer sich gelb färbte (saurer pH) wurde noch TRIS-Base zugegeben bis der Puffer wieder blau wurde. Die so konzentrierten Proben wurden 5 min auf 95°C erhitzt, auf 10%igen SDS-PAGEs aufgetrennt und die Proteine wurden mit kolloidaler Coomassie Färbung sichtbar gemacht.

KOLLOIDALE COOMASSIEFÄRBUNG

Die SDS-PAGEs wurden mit der kolloidalen Coomassiefärbung gefärbt. Diese Färbemethode ist wesentlich sensitiver als die normale Coomassiefärbung. Die Sensitivität ist mit der der Silberfärbung vergleichbar. Im Gegensatz zur Silberfärbung, bei der die Proteine im Gel fixiert werden, lassen sich Proteinbanden, die mit kolloidaler Coomassiefärbung sichtbar gemacht wurden, massenspektrometrisch untersuchen. Für die Färbung wurden die Gele nach dem Lauf erst mindestens zweimal für 10 min in deionisiertem Wasser gewaschen, um das SDS zu entfernen. Anschließend wurde die Färbelösung auf das Gel gegeben und über Nacht gefärbt. Je nach Intensität der Färbung und des Hintergrundes wurde mit Entfärber entfärbt.

3.2.7. MALDI TOF MASSENSPEKTROMETRIE

TRYPsin IN-GEL-VERDAU

Zur Analyse der Proteinbanden mittels Massenspektrometrie wurde ein In-Gel Trypsinverdau durchgeführt. Coomassie gefärbte Proteinbanden wurden aus der SDS-PAGE auf einer sauberen Unterlage mit einem frischen Skalpell ausgeschnitten. Die Banden wurden in etwa 1x1 mm große Stücke zerteilt. In einem 500 µL Eppendorfgefäß wurden 20 µL einer Lösung aus Acetonitril und NH_4HCO_3 (1:1) zu den Gelstücken gegeben und die Ansätze wurden für 15 min bei Raumtemperatur und unter Schütteln inkubiert. Anschließend wurde der Überstand abgenommen und die Gelstücke wurden in 20 µL Acetonitril inkubiert um ihnen die Flüssigkeit zu entziehen. Nach Entfernen des Acetonitrils wurden die Gelstücke in der Vakuumzentrifuge für ca. 10 min getrocknet. Anschließend wurden 20 µL der Reduktionslösung zugegeben. Die Ansätze wurden für 30 min bei 56°C inkubiert. Beim Abnehmen des Überstandes wurde möglichst genau das Volumen bestimmt, das die Gelstücke aufnehmen können, um später einen abnehmbaren Überstand für die Messung zu erhalten. Um das Reduktionsmittel zu entfernen, wurden die Gelstücke wieder mit Acetonitril bedeckt. Zur Carbamidomethylierung wurden 20 µL einer 55 mM Lösung von Iodacetamid in 100 mM NH_4HCO_3 zugegeben und die Gelstücke wurden für 20 min bei Raumtemperatur und im Dunkeln inkubiert. Der Überstand wurde abgenommen und die Gelstücke wurden 15 min mit 100 mM NH_4HCO_3 unter Schütteln gespült. Die Flüssigkeit wurde durch zweimalige Behandlung mit Acetonitril vollständig entfernt. Die Gelstücke wurden in der Vakuumzentrifuge wieder für ca. 10 min getrocknet. Anschließend wurde so viel Trypsinlösung (Trypsinstammlösung 1:80 mit Verdau-puffer verdünnt) zugegeben, dass ein abnehmbarer Überstand entstand, also das oben bestimmte Volumen zuzüglich 3 µL. Zuerst wurden die Ansätze 30 min auf Eis inkubiert, dann bei 37°C über Nacht. Am nächsten Tag wurden die Gefäße kurz zentrifugiert, um die Lösung am Boden zu sammeln und nochmal für 30 min bei 37°C geschüttelt. Der Verdauüberstand wurde abgenommen, wobei mindestens 1 µL für die Messung benötigt wurde. War nicht genug Überstand vorhanden, wurde die nötige Menge Wasser zugegeben und die Ansätze wurden wieder für 30 min inkubiert. Zur Nachextraktion der im Gel verbliebenen Peptide, die oft besser kristallisieren aufgrund des geringeren Salzgehaltes, wurden 10 µL 40 % Acetonitril/0.1 % TFA zugegeben und nochmals mindestens über Nacht bei Raumtemperatur geschüttelt. Die Überstände wurden abgenommen. Bis zur Messung wurden die Proben bei -20°C aufbewahrt.

MASSENSPEKTROMETRIE¹

Die Massen der durch Trypsin in-Gel-Verdau erhaltenen Peptide wurden durch Matrix unterstützte Laser Desorption/Ionisation *time of flight* Massenspektrometrie (MALDI-TOF-MS) auf einem Ultraflex-II TOF/TOF Instrument (Bruker Daltonics, Bremen, Germany) mit einem 200 Hz solid-state Smart beam™ Laser analysiert. Das Massenspektrometer wurde im positiven Ionenmodus betrieben. Die Massenspektren wurden über einen Messbereich von einem Masse zu Ladungsverhältnis (m/z) von 600-4000 aufgenommen. Als Matrix wurde α -Cyano-4-hydroxy-zimtsäure (CHCA) verwendet. Die Proteinproben wurden auf einen metallischen Träger gespottet und mit der Matrix kokristallisiert. MS/MS Spektren ausgewählter Peptide wurden im LIFT Modus aufgenommen. Die Datenbanksuche wurde mit Mascot (Matrix Science Ltd., <http://www.matrixscience.com>) durchgeführt. Die Massentoleranz wurde normalerweise auf ± 50 ppm gesetzt und eine fehlende Spaltung wurde zugelassen.

3.2.8. HETEROLOGE PROTEINEXPRESSION

EXPRESSION IN E. COLI

Die pET28a Expressionsplasmide wurden in den *E. coli* Stamm BL21 (DE3) transformiert. Zur Expression wurde in der Regel das EnPresso *Tablet Cultivation Set* nach Herstellerangaben verwendet. Dazu wurde eine 1 mL LB Vorkultur angeimpft. Nach 6-8 h Inkubation bei 37°C wurden die EnPresso Medium-Tabletten in 50 mL sterilem Wasser gelöst, 25 μ L *EnZ I'm*, 100 μ g/mL Kanamycin und 500 μ L der Vorkultur wurden zugegeben. Diese Kultur wurde über Nacht bei 30°C im Schüttelinkubator inkubiert. Am nächsten Tag wurde die *Booster* Tablette, weitere 50 μ L *EnZ I'm* und 200 μ M IPTG zur Induktion zugegeben. Nach 24 h Inkubation bei 30°C wurde die Kultur durch Zentrifugation geerntet. Das Pellet wurde entweder sofort zur Proteinreinigung weiterverwendet oder bei -20°C gelagert.

REINIGUNG DER PROTEINE ÜBER DEN POLYHISTIDIN-TAG

Zum Zellaufschluss wurde das Pellet in 3 mL NPI-10 resuspendiert. Es wurde eine Spatelspitze Lysozm zugegeben und die Suspension wurde 30 min auf Eis inkubiert. Anschließend folgte der Aufschluss mit der Ultraschallsonde: 15 sec Ultraschall, 15 sec Pause 10 Wiederholungen (*Output*: 4). Zelltrümmer wurden 30 min bei 10000 g und 4°C abzentrifugiert, der Überstand wurde abgenommen und das Pellet wurde verworfen. Eine leere Säule wurde mit 1 mL Säulenmaterial (Ni-NTA Agarose) befüllt und das Säulenmaterial wurde mit 5 mL NPI-10 äquilibriert. Der Überstand aus der Zentrifugation wurde auf

¹ Die massenspektrometrische Analyse wurde von Dr. Christoph Weise an der FU Berlin durchgeführt.

die Säule gegeben. Die Säule wurde viermal mit 4 mL NPI-20 gewaschen. Das Protein wurde in vier 0,5 mL Fraktionen mit NPI-250 Puffer eluiert. Um das Imidazol aus den Fraktionen zu entfernen, wurde über Nacht gegen 2 L Dialysepuffer dialysiert. Die Reinheiten wurden auf SDS-PAGEs überprüft. Proteinkonzentrationen wurden mit Hilfe des BCA Kits bestimmt. Die Proteine wurden entweder sofort weiterverwendet oder sie wurden bei -20°C gelagert.

3.2.9. *ELECTROPHORETIC MOBILITY SHIFT ASSAY*

Für jeden Ansatz wurden 20 pmol Oligonucleotid in Faltungspuffer gefaltet. Anschließend wurden Protein und 4xEMSA Puffer zugegeben und die Ansätze wurden mit H₂O auf ein Volumen von 20 µL aufgefüllt. Die Ansätze wurden für 30 min bei 30°C inkubiert. Um sie auf das Gel aufzutragen, wurden 3 µL Glycerin zugegeben. Die Proben wurden dann 2 h bei 10 mA pro Gel auf einer nativen PAGE aufgetrennt. Die Nucleinsäuren wurden durch 30 min Färbung mit GelRed Lösung sichtbar gemacht.

3.2.10. *OBERFLÄCHENPLASMONRESONANZ-SPEKTROSKOPIE*

Für die Ermittlung von Bindungsparametern mittels Oberflächenplasmonresonanz-Spektroskopie wurde ein Bindungspartner, der Ligand (hier das Protein), auf einem Sensorchip immobilisiert. Der Sensorchip bestand aus einer Glasoberfläche mit einer dünnen Goldschicht, auf der eine Dextranmatrix aufgebracht war, über die der Ligand kovalent gebunden werden konnte. Die Detektionseinheit auf der Rückseite des Chips bestand aus einer Lichtquelle, einem Prisma und einem Detektor. In der Flusszelle wurde der Analyt, das RNA Oligonucleotid, am immobilisierten Liganden vorbeigeleitet (Abbildung 8). Hierbei wird die Masseveränderung gemessen, die während der Interaktion des Analyten mit dem Liganden auf der Chipoberfläche auftritt. Die Signalstärke wird in Resonanzeinheiten (*Response Units*, RU) ausgedrückt und ist proportional zur gebundenen Masse. Aus dem Messsignal lassen sich die Bindungsparameter bestimmen.

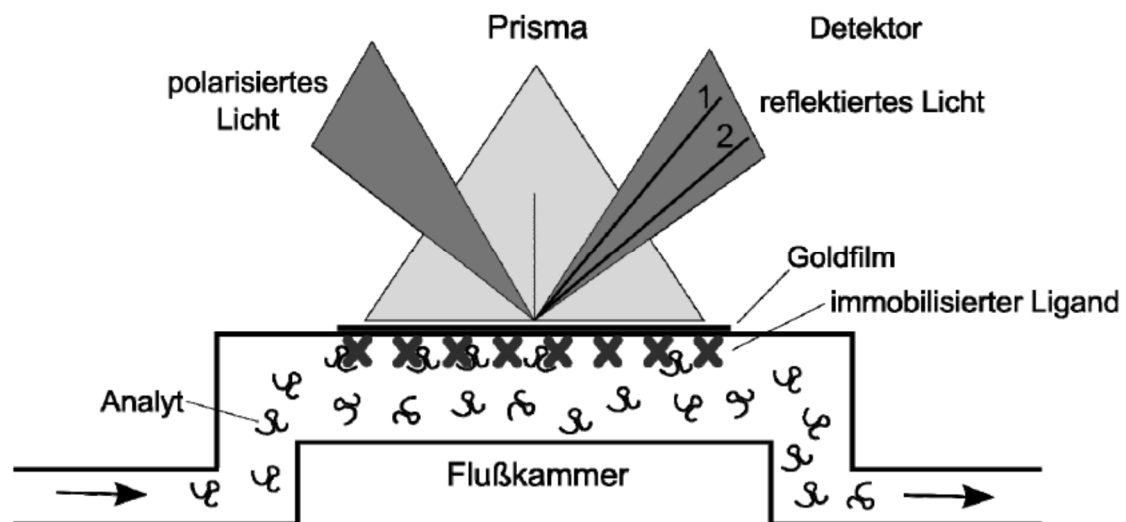


Abbildung 8: Prinzip der Oberflächenplasmonresonanz. Ein Bindungspartner wird als Ligand auf der Sensorchipoberfläche immobilisiert, der Analyt wird in der Flusszelle am Liganden vorbeigeleitet. Interaktionen zwischen den beiden Molekülen werden als Massenveränderungen auf der Oberfläche durch die Detektionseinheit gemessen.

Oberflächenplasmonresonanz-Experimente wurden auf den Biacore-X oder Biacore-2000 SPR-Instrumenten ausgeführt. Alle Messungen wurden bei 25°C durchgeführt. Die Proteine wurden auf Carboxymethyldextran (CMD 500 L) Sensorchips über Amin-Kopplung immobilisiert. Die Immobilisierung wurde in HBS-EP Laufpuffer bei einer Flussrate von 5 µL/min durchgeführt. Die Oberfläche des CMD 500 Chips wurde mit 35 µL einer Mischung aus 0,02 M L-Ethyl-3-(3-dimethylaminopropyl)carbodiimid Hydrochlorid (EDC) und N-Hydroxysuccinimid (NHS) aktiviert. Anschließend wurde Protein, verdünnt in HBS-EP Laufpuffer, injiziert, bis etwa 1000 Response Units erreicht waren. Die aktivierte Oberfläche wurde dann durch die Injektion von 35 µL Ethanolamin (pH 8,5) blockiert. Ein aktivierter und blockierter Kanal, in dem kein Protein gebunden war, diente als Referenzkanal. Die Bindungsexperimente wurden in einer 1:10 Mischung aus Faltungspuffer, in dem die Oligonucleotide gelöst waren und HBS-EP Puffer durchgeführt. Dies diente dazu "bulk-shift" Effekte zu minimieren, die durch die Änderung der optischen Dichte bei der Injektion eines anderen Puffers als dem Laufpuffer zustande kommen. 30 µL jedes Oligonucleotids wurden mit einer Flussrate von 20 µL/min, also über 90 s, über beide Kanäle injiziert. Aus dem Verlauf dieser Bindung ließ sich die Assoziationsgeschwindigkeitskonstante (k_a) ermitteln. Der Injektion des Analyten folgt die Dissoziationsphase, aus der die Dissoziationsgeschwindigkeitskonstante (k_d) ermittelt werden konnte. Zur Korrektur der Messungen auf Hintergrundsignale wurden die entsprechenden Signale der Referenzzellen für jede Injektion, sowie eine Pufferinjektion abgezogen. Das gemessene Signal für die Assoziation und die Dissoziation bei verschiedenen Analytkonzentrationen wurde mit Hilfe der Software BIAevaluation (Version 4.1) an das Langmuir Bindungsmodell angepasst ("Fit"). Dieses Modell beruht auf einer 1:1 Bindung der Interaktionspartner, die beide nur eine Bindungsstelle besitzen. Die

Dissoziationskonstante kann dann aus den Geschwindigkeitskonstanten nach der Gleichung $K_D = k_d/k_a$ berechnet werden. Die Oligonucleotide wurden in Konzentrationen von 6-1000 nM eingesetzt. Zur Regeneration der Oberfläche wurden nach jedem Oligonucleotid 10 µL einer 1 M KCl-Lösung injiziert.

3.2.11. *LUCIFERASE CO-TRANSFEKTIONEN*

Um den Einfluss der identifizierten Proteine auf die Translation zu untersuchen wurden Co-Transfektionen durchgeführt. Dazu wurden je 0,4 µg der Luciferase-Reporterplasmide mit 0,8 µg der proteincodierenden Plasmide in Hek293 Zellen mit Hilfe des Transfektionsreagenz *Lipofectamine 2000* transfiziert. Die Zellen wurden 48 h nach der Transfektion geerntet und die Luciferase-Aktivitäten wurden wie in Abschnitt 3.2.5 beschrieben bestimmt. Zur Kontrolle wurde ein GFP codierendes Plasmid transfiziert. Die gemessenen Renilla-Luciferase-Aktivitäten wurden auf die Firefly-Aktivitäten normiert. Die so erhaltenen Werte wurden auf die Aktivitäten der mutierten Kontrolle bezogen.

3.2.12. *IN VITRO TRANSKRIPTION UND IN VITRO TRANSLATION*

IN VITRO TRANSKRIPTION

Die Renilla-Luciferase-Reporter Konstrukte wurden mit dem psiCHECK-2 Vektor als Matrize und mit Hilfe des Ampliscribe T7 Transkriptionskits transkribiert. Die psiCHECK-2 Konstrukte, die als Template dienen sollten, wurden zunächst mit dem Restriktionsenzym XhoI linearisiert. Die Ansätze setzten sich wie folgt zusammen:

1 µg	Template DNA
4 µL	5xTranscription Buffer
1 µL	10 mM ATP
1 µL	10 mM CTP
1 µL	10 mM GTP
1 µL	10 mM UTP
2 µL	100 mM DTT
10 U	T7 RNA Polymerase

ad 20 µL

Die Transkription erfolgte für 2 h bei 37°C. Die Template DNA wurde durch einen DNase Verdau abgebaut (2 h, 37°C) und die Produkte wurden mit dem RNeasy Mini Kit gereinigt. Nach der Transkription wurden die Produkte auf einem RNA Agarosegel mit Formaldehyd überprüft. Die Konzentrationen wurden mit dem Nanodrop bestimmt.

IN VITRO TRANSLATION

Zur *in vitro* Translation wurde das Flexi® Rabbit Reticulocyte Lysate System verwendet. Als Template wurden 0,6 pmol der RNA aus der *in vitro* Transkription eingesetzt. Zuerst wurden RNA und Protein gemischt und für 10 min bei Raumtemperatur inkubiert, damit sich die Bindung zwischen RNA und Protein schon ausbilden konnte, bevor die restlichen Komponenten zugegeben wurden. Vom Protein wurde die doppelte molare Menge (1,2 pmol) zugegeben. Zur Normierung wurde jede RNA vor dem Pipettieren der Ansätze mit der gleichen Menge Firefly RNA aus dem Kit versetzt. Ein Ansatz war wie folgt zusammengesetzt:

7 µL	Retikulozytenlysat
0,1 µL	<i>Amino Acid Mixture, minus Leucine</i>
0,1 µL	<i>Amino Acid Mixture, minus Methionine</i>
0,28 µL	2,5 M KCl
0,2 µL	Ribolock RNase Inhibitor (40 U/mL)
<u>200 ng Template RNA mit Protein</u>	
ad 10 µL	

Die Ansätze wurden für 90 min bei 30°C inkubiert. Von jedem Ansatz wurden Triplikate angefertigt. Anschließend wurden die Aktivitäten der beiden Luciferasen mindestens in Duplikaten gemessen.

3.2.13. SILAC PULL-DOWN-ASSAY

Das *stable isotope labeling with amino acids in cell culture* (SILAC) ist eine Methode, bei der Proteine durch den Einbau von Aminosäuren mit schweren Isotopen markiert werden. Dies geschieht durch die Zugabe der entsprechenden Aminosäuren in das Zellkulturmedium. In diesem Experiment wurden die Aminosäuren L-Lysin:2HCl, 13C6, 15N2 und L-Arginin:HCl, 13C6 verwendet. SILAC-DMEM, das kein Lysin und kein Arginin enthielt, wurde mit den markierten Aminosäuren versetzt. Dem Medium wurde FCS zugesetzt, aus dem die Aminosäuren mittels Dialyse entfernt wurden. Hek293 Zellen wurden für 7 Zellteilungen in diesem Medium kultiviert. Eine andere Population Hek293 Zellen wurde in Medium mit unmarkierten Aminosäuren kultiviert.

Mit dem Proteinextrakt dieser beiden Populationen wurden Pull-down-Assays durchgeführt. Der Extrakt mit den schweren Aminosäuren wurde mit den G-Quadruplex-Sequenzen inkubiert und der mit den unmarkierten Proteinen mit den mutierten Sequenzen. Wie im vorherigen Pull-down-Assay wurden die Streptavidin-Agarose-Beads mit Pull-down-Waschpuffer mit steigenden KCl-Konzentrationen gewaschen. Jedoch wurde nicht in mehreren Schritten eluiert, sondern die *Beads* wurden mit Puffer bis 600 mM KCl gewaschen und anschließend wurde in einem Schritt mit Puffer mit 2600 mM KCl eluiert. Die Elutionsfraktionen von der G-Quadruplex und der entsprechenden mutierten

Sequenz wurden 1:1 gemischt. Um das KCL zu entfernen und die Proben einzuengen, wurde eine Ultrafiltration mit einer Amicon Säule durchgeführt. Anschließend wurden die Proben auf eine SDS-PAGE aufgetragen und etwa 0,5 cm in das Gel laufen gelassen. Das Gel wurde dann mit der kolloidalen Coomassielösung gefärbt und das Gelstück mit der Probe ausgeschnitten. Die Analyse der Proben fand dann im Labor von Prof. Juri Rappsilber an der University of Edinburgh, Schottland statt.

3.2.14. *QUANTITATIVE REAL TIME PCR*

Für die quantitative Real Time PCR (qPCR) wurde zunächst RNA aus den entsprechenden eukaryotischen Zellen mit Hilfe des RNeasy Mini Kits (Qiagen) isoliert. Nach einem DNase Verdau wurde die mRNA mit Hilfe des Maxima H Minus First Strand cDNA Synthesis Kits in cDNA umgeschrieben. Die Reverse Transkription wurde mit Oligo dT Primern durchgeführt. Für jede Probe wurde eine Kontrolle ohne Reverse Transkriptase angesetzt. Die cDNA wurde in der Regel 1:40 verdünnt und direkt in die qPCR eingesetzt. Für die qPCR wurde der Maxima Sybr Green qPCR Master Mix verwendet. Die Primer sind in

Tabelle 8 aufgeführt. Für jede Probe wurde die qPCR in Triplikaten durchgeführt. Als Kontrollen dienten die Ansätze ohne Reverse Transkription sowie ohne Template. Die qPCR Daten wurden mit der Software Bio-Rad CFX Manager 3.1 ausgewertet.

3.2.15. *STATISTISCHE AUSWERTUNG*

Die Signifikanzanalyse von Messergebnissen erfolgte mittels einfaktorieller Varianzanalyse (ANOVA, *analysis of variance*), zum Vergleich der Varianzen zwischen mehr als zwei Gruppen. Das Signifikanzniveau wurde mit dem Sidak-Test bestimmt (92). Die folgenden Signifikanzstufen wurden dabei vergeben:

- * $P \leq 0.05$
- ** $P \leq 0.01$
- *** $P \leq 0.001$
- **** $P \leq 0.0001$

Die statistische Auswertung wurde mit dem Programm GraphPad Prism (Version 6.01) durchgeführt.

4. ERGEBNISSE

4.1. AUSWAHL DER G-QUADRUPLPLEX-SEQUENZEN

In dieser Arbeit sollten Proteine identifiziert werden, die G-Quadruplexe erkennen können. Um sowohl spezifische, als auch generelle Binder zu finden, wurden zwei G-Quadruplex-Sequenzen für die anschließenden Pull-down-Experimente ausgewählt. Die G-Quadruplex-Sequenz aus der 5'-UTR der *matrix metalloproteinase 16* (MMP16) war bereits publiziert und wurde aufgrund ihrer Fähigkeit, ein stabiles G-Quadruplex zu formen, ausgewählt (39). Über die G-reiche Sequenz aus der 5'-UTR des *actin related protein 2/3 complex subunit 2* (ARPC2) war noch nichts bekannt. Abbildung 9 zeigt die Sequenzen der beiden 5'-UTRs der mRNAs mit den Positionen der potentiell G-Quadruplex bildenden Regionen.

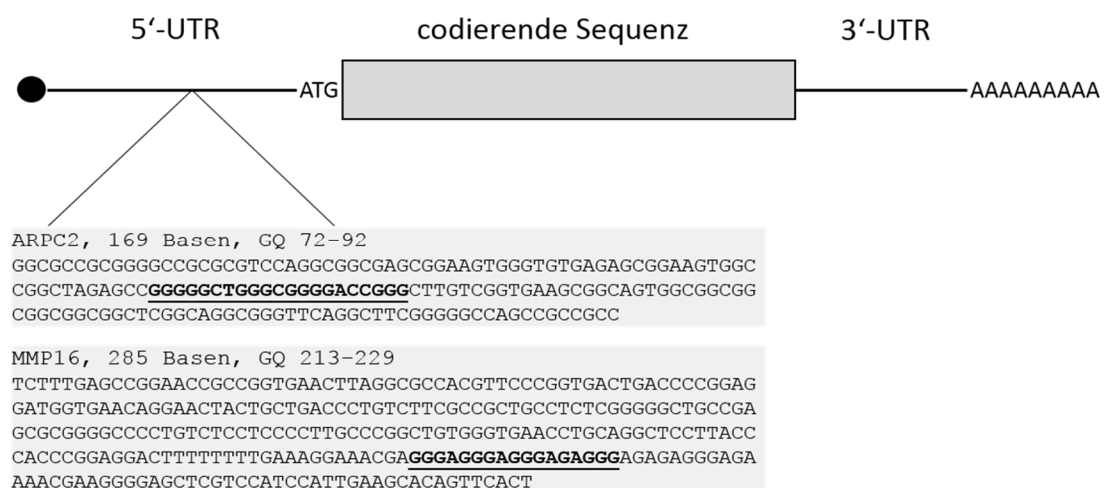


Abbildung 9: Schematische Darstellung der mRNAs von ARPC2 und MMP16. In den grauen Kästen sind die Sequenzen der 5'-UTRs und die Positionen der potentiell G-Quadruplex bildenden Regionen gezeigt. Die G-Quadruplex-Sequenzen sind fett gedruckt und unterstrichen. NCBI Referenz Sequenzen: ARPC2 NM_152862.2; MMP16 NM_005941.4

Der Vergleich der beiden 5'-UTR Sequenzen zeigt, dass die G-Quadruplex-Sequenz in der Mitte der 169 Basen langen ARPC2-UTR liegt, 77 Basen vor dem Startcodon, während die MMP16 G-Quadruplex-Sequenz sich etwas näher am Translationsstartpunkt befindet. Die MMP16 5'-UTR ist 285 Basen lang und die G-Quadruplex-Sequenz liegt etwa 50 Basen vor dem Startcodon.

Die Konservierung der MMP16 und ARPC2 G-Quadruplex-Sequenzen zwischen verschiedenen Spezies ist in Tabelle 9 und Tabelle 10 dargestellt. Die Konservierung der Sequenz innerhalb verschiedener Spezies deutet darauf hin, dass sie eine wichtige Rolle in biologischen Prozessen spielt. Auffällig ist, vor allem für MMP16, dass die Position der Sequenz sehr stark konserviert ist.

Tabelle 9: Konservierung der MMP16 G-Quadruplex-Sequenz in verschiedenen Spezies. Guanin-Abfolgen, die mutmaßlich an der Ausbildung des G-Quadruplexes beteiligt sind, sind unterstrichen.

Spezies	Sequenz	Position
<i>Homo sapiens</i> (Menschen)	<u>GGGAGGGAGGGAGAGGG</u>	-70
<i>Gorilla gorilla</i> (Gorillas)	<u>GGGAGGGAGGGAGAGGG</u>	-70
<i>Pan troglodytes</i> (Schimpansen)	<u>GGGAGGGAGGGAGAGGG</u>	-70
<i>Elephantulus edwardii</i> (Rüsselspringer)	<u>GGGAGGGAGGGTGAGGG</u>	-70
<i>Felis catus</i> (Katzen)	<u>GGGAGGGAGGGAGAGGG</u>	-70
<i>Canis lupus</i> (Hunde)	<u>GGGAGGGAGGGAGAGGG</u>	-70
<i>Erinaceus europaeus</i> (Igel)	<u>GGGAGGGAGGGTGAGGG</u>	-70

Tabelle 10: Konservierung der ARPC2 G-Quadruplex-Sequenz. Guanin-Abfolgen, die mutmaßlich an der Ausbildung des G-Quadruplexes beteiligt sind, sind unterstrichen.

Spezies	Sequenz	Position
<i>Homo sapiens</i> (Menschen)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-98
<i>Papio anubis</i> (Paviane)	<u>GGGGGCTGGGCGGG-GACTGGG</u>	-98
<i>Gorilla gorilla</i> (Gorillas)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-195
<i>Pan troglodytes</i> (Schimpansen)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-98
<i>Rattus norvegicus</i> (Ratten)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-95
<i>Felis catus</i> (Katzen)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-95
<i>Monodelphis domestica</i> (Beutelmurmel)	<u>GGGGGCTGGGCGGGAC-CGGG</u>	-306

Um ein realistisches Modellsystem für die Identifizierung G-Quadruplex bindender Proteine zu nutzen, wurden Gene mit 5'-UTR G-Quadruplexen ausgewählt, die in der untersuchten Zelllinie, im vorliegenden Fall in Hek293 Zellen, exprimiert werden. So sollte sichergestellt werden, dass auch die entsprechenden regulatorischen Proteine in den Zellen gebildet werden. Zum Nachweis der Expression der Gene wurde Gesamt-RNA aus kultivierten Zellen isoliert und mittels reverser Transkription in cDNA umgeschrieben. Anschließend wurden die entsprechenden Sequenzen mit Hilfe spezifischer Primer amplifiziert (Abbildung 10).

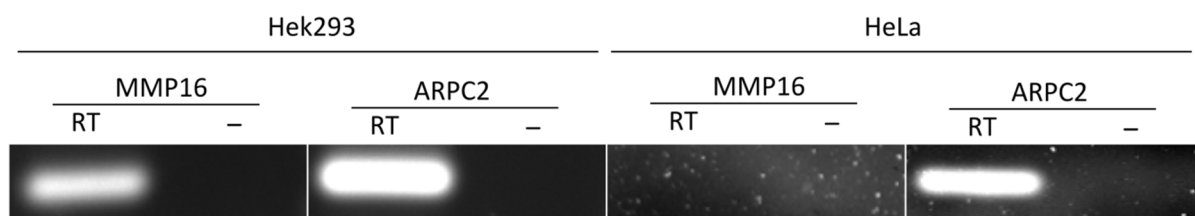


Abbildung 10: Agarosegel einer RT-PCR zum Nachweis der Expression der G-Quadruplex-Sequenzen in Hek293 und in HeLa Zellen. Die Kontrollen für die Reverse Transkription, denen keine Reverse Transkriptase zugesetzt wurde, sind mit einem Minus-Zeichen markiert. Die Amplifikation wurde in 30 Zyklen durchgeführt.

Die Expression der Gene MMP16 und ARPC2 in Hek293 Zellen konnte mittels RT-PCR nachgewiesen werden. Durch die Negativkontrolle ohne Reverse Transkriptase konnten Verunreinigungen durch genomische DNA ausgeschlossen werden. In HeLa Zellen, die für Verifikationsexperimente eingesetzt wurden, wurde nur ARPC2 exprimiert, eine Expression von MMP16 konnte nicht nachgewiesen werden.

Für die folgenden Experimente zur Charakterisierung der G-Quadruplex-Struktur und zur Identifizierung von G-Quadruplex bindenden Proteinen wurden Oligonucleotide mit den entsprechenden Sequenzen designt (Tabelle 11). Diese Oligonucleotide enthielten auch einen Teil der G-Quadruplex flankierenden Sequenz, um eine möglichst natürliche Faltung zu erreichen. Für die Kontrollsequenzen (MMP16 mut, ARPC2 mut und ARPC2 2xG) wurden die Guanin-Basen ausgetauscht, die wahrscheinlich an der Bildung einer G-Quadruplex-Struktur beteiligt sind, so dass die Ausbildung eines G-Quadruplexes theoretisch nicht mehr möglich ist. In der Kontrolle ARPC2 2xG wurden zwei Guanin-Triplets unverändert gelassen, um die Sequenz und den Guanin Gehalt nicht zu stark zu verändern und so eine Kontrolle zu erhalten, deren Sequenz möglichst ähnlich zur natürlichen ist, die aber kein G-Quadruplex ausbilden kann. Auf diese Weise sollten Proteine, die eine Affinität für G-reiche Sequenzen, aber nicht speziell zur G-Quadruplex-Struktur haben, ausgeschlossen werden.

Tabelle 11: Biotinylierte G-Quadruplex RNA Oligonucleotide und die entsprechenden mutierten Sequenzen. Diese Oligonucleotide wurden zur Charakterisierung der G-Quadruplex-Strukturen sowie für die Pull-down-Assays designt. Nucleotide, die vermutlich an der Bildung des G-Quadruplex beteiligt sind, sind unterstrichen. Mutierte Nucleotide in den Kontrollsequenzen sind fett gedruckt.

Oligonucleotid Name	Sequenz 5' → 3'
MMP16 GQ	AACGAGGGAGGGAGGGAGAGGGAGAGA-Biotin
MMP16 mut	AACGAGAGAGAGAGAGAGAGAGAGAGA-Biotin
ARPC2 GQ	AGCCGGGGGCUGGGCGGGACCGGGCUUGU-Biotin
ARPC2 mut	AGCC <u>GUAG</u> ACUGAGCGAAGACCGAGCUUGU-Biotin
ARPC2 2xG	AGCC <u>GUAG</u> ACUGGGCGAAGACCGGGCUUGU-Biotin

4.2. CHARAKTERISIERUNG DER G-QUADRUPLEX OLIGONUCLEOTIDE

Zunächst sollte die Faltung der ausgewählten guaninreichen Oligonucleotide charakterisiert werden. Während für die MMP16 Sequenz die Ausbildung einer G-Quadruplex-Struktur bereits in der Literatur beschrieben wurde (39), gab es für die ARPC2-Struktur bislang keine entsprechenden Analysen. Daher wurden die guaninreichen Sequenzen mittels bioinformatischer und spektroskopischer Methoden untersucht.

4.2.1. BIOINFORMATISCHE ANALYSE

Die ausgewählten Sequenzen wurden mit Hilfe bioinformatischer Methoden überprüft. Die Ausbildung potentieller G-Quadruplex-Strukturen einer Sequenz wurde durch den QGRS (*Quadruplex forming G-rich sequences*) mapper (<http://bioinformatics.ramapo.edu/QGRS/index.php>) analysiert (93). Für jede Sequenz wird ein G-Score berechnet. Der G-score wird durch die Loops, die für die Ausbildung eines stabilen G-Quadruplex Motifs möglichst kurz und gleich lang sein sollten, sowie die Anzahl der Guanine beeinflusst. Je höher der G-Score, desto wahrscheinlicher ist die Ausbildung des G-Quadruplexes. Die Ergebnisse der Analyse des ARPC2 G-Quadruplex Oligonucleotids sind in Tabelle 12 dargestellt.

Tabelle 12: QGRS Mapper-Analyse (<http://bioinformatics.ramapo.edu/QGRS/index.php>) der ARPC2 G-Quadruplex-Sequenz. Es wurden sechs mögliche G-Quadruplex-Strukturen gefunden. Die jeweils beteiligten Nucleotide sind unterstrichen und fett gedruckt. Die Länge der entsprechenden Sequenz, sowie der G-Score, der die Wahrscheinlichkeit der Ausbildung der Struktur angibt, sind aufgeführt.

Länge	Sequenz	G-Score
21	<u>GGGG</u> GCUG <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	39
21	<u>GGGG</u> GCUG <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	40
20	<u>GGGG</u> CU <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	39
20	<u>GGGG</u> CU <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	41
19	<u>GGG</u> CU <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	39
19	<u>GGG</u> CU <u>GGG</u> C <u>GGG</u> GACCG <u>GG</u>	41

Die Analyse mit dem QGRS Mapper ergab sechs mögliche G-Quadruplex-Strukturen die das ARPC2 Oligonucleotid einnehmen könnte. Für das MMP16 Oligonucleotid wurde nur eine Struktur gefunden, die einen G-Score von 41 hat. Die ARPC2 Sequenz besaß längere Abfolgen von Guaninen, wodurch sich mehr mögliche Kombinationen ergaben.

Des Weiteren wurde mit Hilfe des *mfold* Web Servers (<http://mfold.rna.albany.edu/> (94), Version *mfold_util* 4.6, Datum der Analyse: Oktober 2014) überprüft, ob die Oligonucleotide andere Sekundärstrukturen einnehmen können. G-Quadruplex-Strukturen konnten durch den *mfold* Algorithmus zum Zeitpunkt der Analyse nicht vorhergesagt werden. Hier ergab sich eine mögliche, weitere Sekundärstruktur für das ARPC2 Oligonucleotid mit einer Änderung der freien Enthalpie (ΔG) von -10,8 kcal/mol (Abbildung 11).

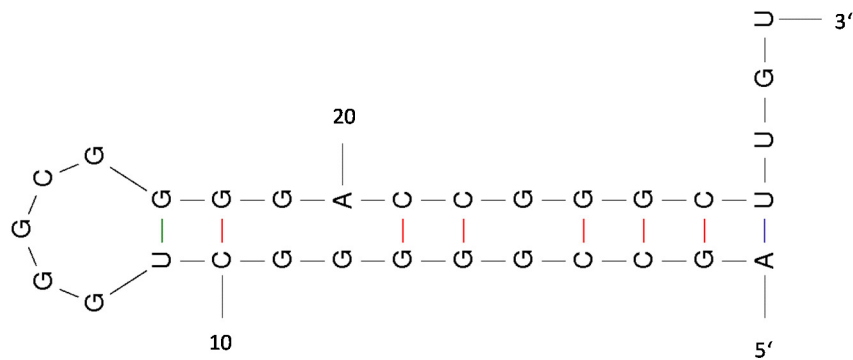


Abbildung 11: Mögliche Sekundärstruktur des ARPC2 GQ Oligonucleotids, generiert mit dem mfold Web Server (<http://mfold.rna.albany.edu/>). Diese Haarnadelstruktur besitzt ein ΔG von -10,8 kcal/mol. 5'- und 3'-Ende der Sequenz sind markiert.

Für die anderen in dieser Arbeit verwendeten Sequenzen konnten keine Sekundärstrukturen gefunden werden, die aufgrund ihres ΔG -Wertes relevant wären.

Das ARPC2 Oligonucleotid könnte damit sieben mögliche Strukturen einnehmen, während sich für die MMP16 Sequenz nur eine G-Quadruplex-Struktur ergab.

4.2.2. CD-SPEKTROSKOPIE

Die Ausbildung von G-Quadruplex-Strukturen lässt sich mittels spektroskopischer Methoden überprüfen. Dabei ist eine Standardmethode zum Nachweis von Sekundärstrukturen die Circular dichroismus (CD)-Spektroskopie. Hierbei wird die Absorption von zirkulär polarisiertem Licht durch das zu untersuchende Molekül gemessen, woraus man Rückschlüsse auf die Sekundärstruktur schließen kann.

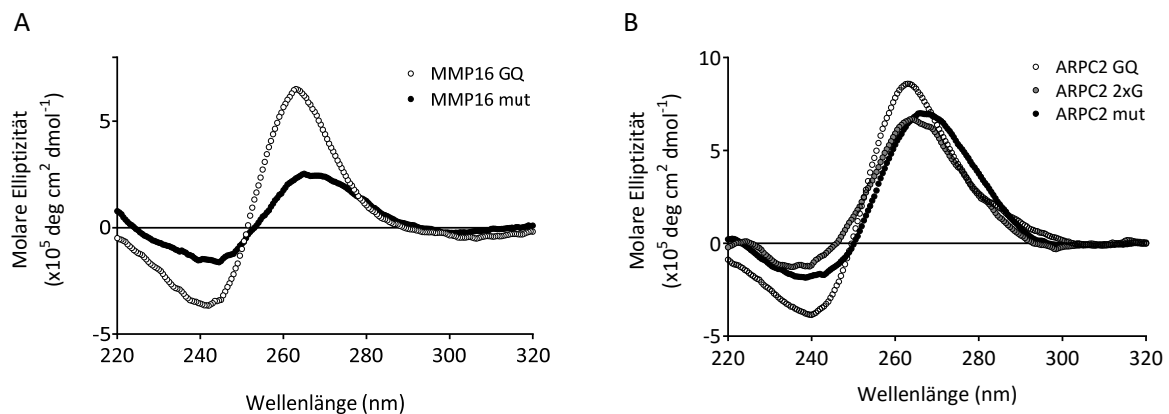


Abbildung 12: CD-Spektren der Oligonucleotide MMP16 GQ und mut (A) sowie ARPC2 GQ, 2xG und mut (B). Die Oligonucleotide wurden vor der Messung in Gegenwart von 100 mM KCl auf 95°C erhitzt und langsam abgekühlt um die Ausbildung der G-Quadruplex-Struktur zu begünstigen. Die Messungen wurden bei 20°C und mit einer Scan-Geschwindigkeit von 50 nm/min durchgeführt. Die CD-Spektren wurden mindestens fünfmal aufgenommen, gezeigt ist hier jeweils ein repräsentatives Spektrum.

Abbildung 12 zeigt die CD-Spektren der in dieser Arbeit verwendeten Oligonucleotide MMP16 GQ und mut (A) sowie ARPC2 GQ und den Kontrollen ARPC2 2xG und mut (B).

Das Spektrum des MMP16 GQ Oligonucleotids zeigte einen positiven Peak bei 260 nm und einen negativen Peak bei 240 nm, was charakteristisch für ein paralleles G-Quadruplex ist (12). Das entsprechende mutierte Oligonucleotid MMP16 mut zeigte diese Merkmale nicht so ausgeprägt wie das G-Quadruplex Oligonucleotid, was darauf hindeutet, dass es keine G-Quadruplex-Struktur ausbilden kann.

Das CD-Spektrum des ARPC2 G-Quadruplex Oligonucleotids zeigt dieselben Charakteristika wie das des MMP16 G-Quadruplex Oligonucleotids (positiver Peak bei 260 nm, negativer bei 240 nm). Die Spektren der beiden ARPC2 Kontrollen sehen jedoch recht ähnlich aus. Sie zeigen auch Peaks bei ca. 260 und 240 nm, diese sind nur etwas weniger ausgeprägt und leicht versetzt. Da die Spektren der Kontrollsequenzen nicht signifikant von dem Spektrum des ARPC2 G-Quadruplex Oligonucleotids abweichen, ist es schwierig, hieraus Schlüsse auf die Sekundärstruktur zu ziehen. CD-Spektren lassen nicht mit Sicherheit auf die Ausbildung einer G-Quadruplex-Struktur schließen, deshalb wurden zur weiteren Charakterisierung UV-Schmelzkurven angefertigt.

4.2.3. UV-SCHMELZKURVEN

Eine weitere Methode zur Analyse der Ausbildung von G-Quadruplexen ist die Aufnahme von UV-Schmelzkurven, die ein charakteristisches sigmoidales Schmelzprofil aufweisen. Im Gegensatz zu klassischen Nucleinsäureschmelzkurven, die bei 260 nm aufgenommen werden, wird das Schmelzverhalten von G-Quadruplexen bei 295 nm untersucht. Bei dieser Wellenlänge zeigen sie den

größten Unterschied in der Absorption zwischen dem gefalteten und dem ungefalteten Zustand (13). Aus diesen Spektren lassen sich auch die Schmelztemperaturen der G-Quadruplexe bestimmen.

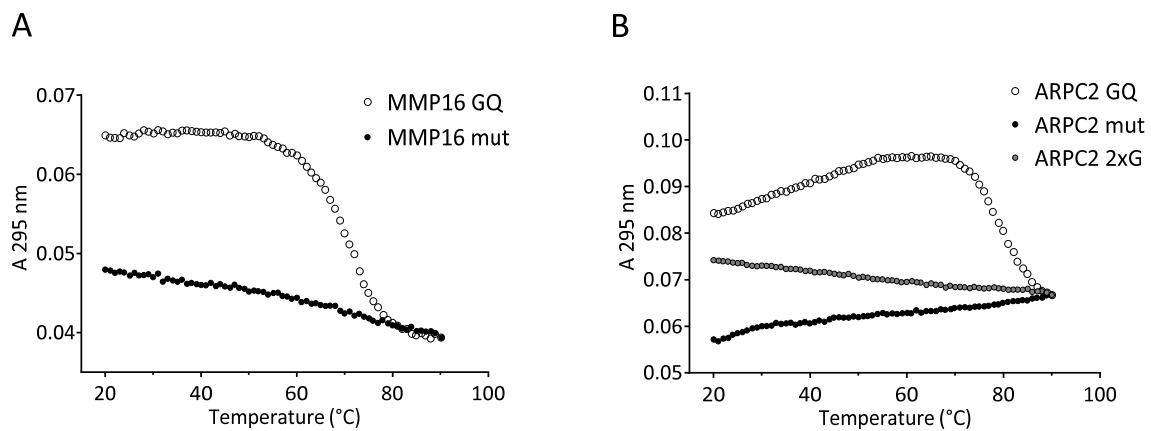


Abbildung 13: UV-Schmelzkurven der Oligonucleotide. Die MMP16 Oligonucleotide wurden bei einer KCl-Konzentration von 10 mM gemessen, da die Struktur bei 100 mM KCl nicht entfaltet werden konnte. Die Schmelzkurven der ARPC2 Oligonucleotide wurden in TRIS-Puffer mit 100 mM KCl aufgezeichnet.

Die Schmelzkurve des guaninreichen MMP16 Oligonucleotids zeigt das charakteristische sigmoide Schmelzprofil einer G-Quadruplex-Struktur (Abbildung 13A). Die Struktur konnte bei einer KCl-Konzentration von 100 mM nicht aufgeschmolzen werden. Bei einer KCl-Konzentration von 10 mM wurde eine Schmelztemperatur von $69,2 \pm 0,2^\circ\text{C}$ bestimmt. Das mutierte Oligonucleotid zeigt kein solches Schmelzprofil, kann also keine entsprechende Sekundärstruktur ausbilden.

Die ARPC2 G-Quadruplex-Sequenz zeigte ebenfalls eine sigmoide Schmelzkurve, während die beiden Kontrollsequenzen kein solches Verhalten zeigten (Abbildung 13B). Die Schmelztemperatur des ARPC2 Oligonucleotids betrug $76,5 \pm 0,2^\circ\text{C}$ bei einer KCl-Konzentration von 100 mM.

G-Quadruplexe können nicht nur innerhalb eines Nucleinsäurestrangs, also intramolekular, sondern auch zwischen mehreren Strängen ausgebildet werden. Die Abhängigkeit der Schmelztemperatur von der Konzentration des Oligonucleotids gibt Aufschluss über die Ausbildung eines solchen intermolekularen G-Quadruplexes.

Für MMP16 wurde die Unabhängigkeit der Schmelztemperatur von der Oligonucleotidkonzentration bereits beschrieben (39). Um zu untersuchen, ob ARPC2 ein intra- oder intermolekulares G-Quadruplex bildet, wurde die Schmelztemperatur bei verschiedenen Konzentrationen an Oligonucleotid bestimmt (1, 5, 10 μM).

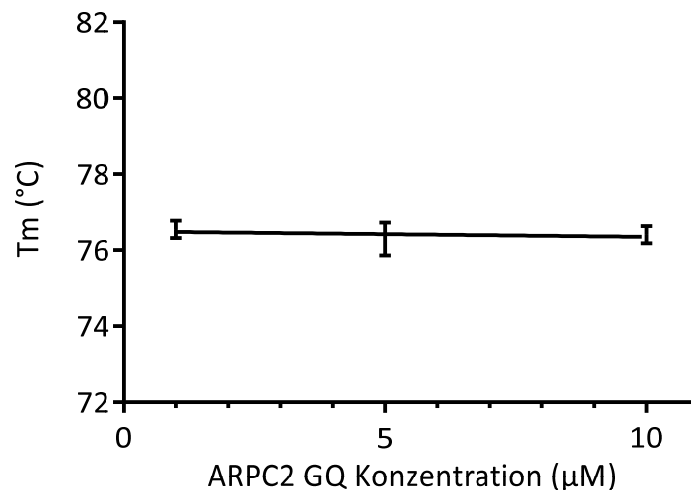


Abbildung 14: Schmelztemperatur des ARPC2 G-Quadruplex Oligonucleotids (ARPC2 GQ) in Abhängigkeit der Konzentration.. Die Schmelztemperatur von 1, 5 und 10 µM ARPC2 Oligonucleotid wurde bei 295 nm und einer KCl-Konzentration von 100 mM bestimmt. Die Temperaturen wurden jeweils aus acht Messungen gemittelt.

Abbildung 14 zeigt die Schmelztemperatur des ARPC2 G-Quadruplex Oligonucleotids in Abhängigkeit der Konzentration. Die Schmelztemperatur ändert sich nicht mit Zunahme der Konzentration an Oligonucleotid. Bei einem intermolekularen G-Quadruplex würde die Schmelztemperatur mit Zunahme der Konzentration steigen. Dieses Ergebnis zeigt, dass ARPC2 kein intermolekulares G-Quadruplex bildet und dass das gebildete G-Quadruplex somit intramolekular sein muss.

4.2.4. ZUSÄTZLICHE KONTROLLOLIGONUCLEOTIDE

Die Nucleotide, die das G-Quadruplex flankieren und zur Simulation der natürlichen Umgebung in das Oligonucleotiddesign miteinbezogen wurden, schienen zumindest bei ARPC2 G-Quadruplex einen möglichen Einfluss auf die Ausbildung des G-Quadruplex zu haben. Daher wurden die entsprechenden RNA Oligonucleotide auch ohne die flankierenden Sequenzen spektroskopisch untersucht.

Des Weiteren wurden die G-Quadruplex-Sequenzen statt als RNA auch als entsprechende DNA Oligonucleotide untersucht. Hier wurden die Sequenzen mit den flankierenden Nucleotiden gewählt. Diese Oligonucleotide sollten in späteren Bindungsexperimenten als Kontrollsequenzen verwendet werden.

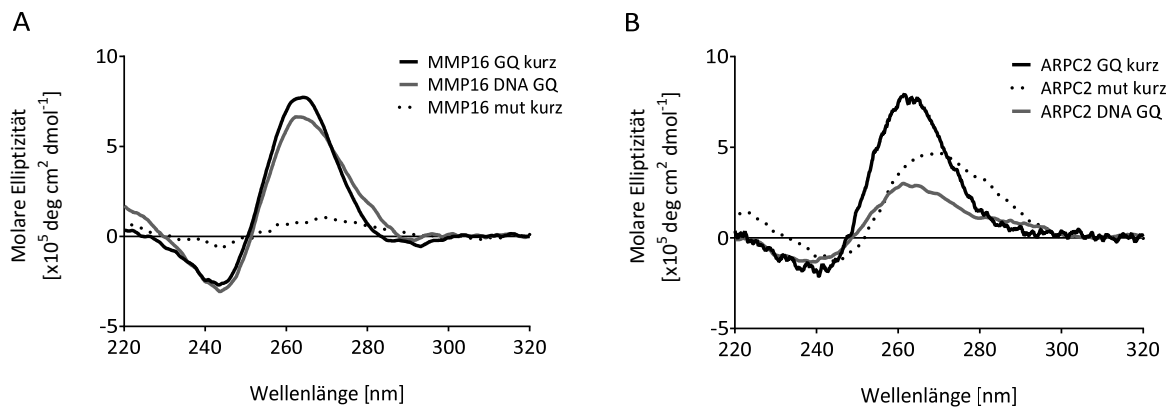


Abbildung 15: CD Spektren der kurzen G-Quadruplex RNA Oligonucleotide und mutierten Kontrollen, sowie der entsprechenden DNA Oligonucleotide. Zur Ausbildung der Sekundärstruktur wurden die Oligonucleotide vor den Messungen in TRIS-HCl-Puffer pH 7,5 mit 100 mM KCl auf 95°C erhitzt und langsam auf Raumtemperatur abgekühlt. Die Messungen wurden bei 20°C und mit einer Scan-Geschwindigkeit von 50 nm/min durchgeführt. Die CD-Spektren wurden mindestens fünfmal aufgenommen, gezeigt ist hier jeweils ein repräsentatives Spektrum.

Die CD Spektren der Kontroll-Oligonucleotide sind in Abbildung 15 dargestellt. Das kurze MMP16 Oligonucleotid (Abbildung 15 A) zeigt das typische CD Profil der parallelen G-Quadruplexe, mit einem positiven Peak bei 260 nm und einem negativen Peak bei 240 nm. Das entsprechende DNA Oligonucleotid weist ein fast deckungsgleiches Spektrum auf, was ebenfalls auf die Ausbildung eines G-Quadruplex hindeutet. Das mutierte MMP16 Oligonucleotid zeigt kein solches Spektrum und scheint somit auch keine entsprechende Sekundärstruktur auszubilden.

Das ARPC2 G-Quadruplex Oligonucleotid ohne die flankierenden Sequenzen zeigt auch das typische CD Spektrum paralleler G-Quadruplexe (Abbildung 15 B). Im Gegensatz zu den längeren ARPC2 Sequenzen, kann man hier einen deutlichen Unterschied zwischen dem G-Quadruplex und dem mutierten Oligonucleotid erkennen. Die Peaks sind viel weniger stark ausgeprägt und leicht verschoben. Das CD-Spektrum des DNA Oligonucleotids zeigt auch nur schwach ausgeprägte Peaks bei 260 und 240 nm, es scheint also auch kein G-Quadruplex ausbilden zu können.

Auch diese Oligonucleotide wurden des Weiteren durch die Aufnahme von UV-Schmelzkurven charakterisiert.

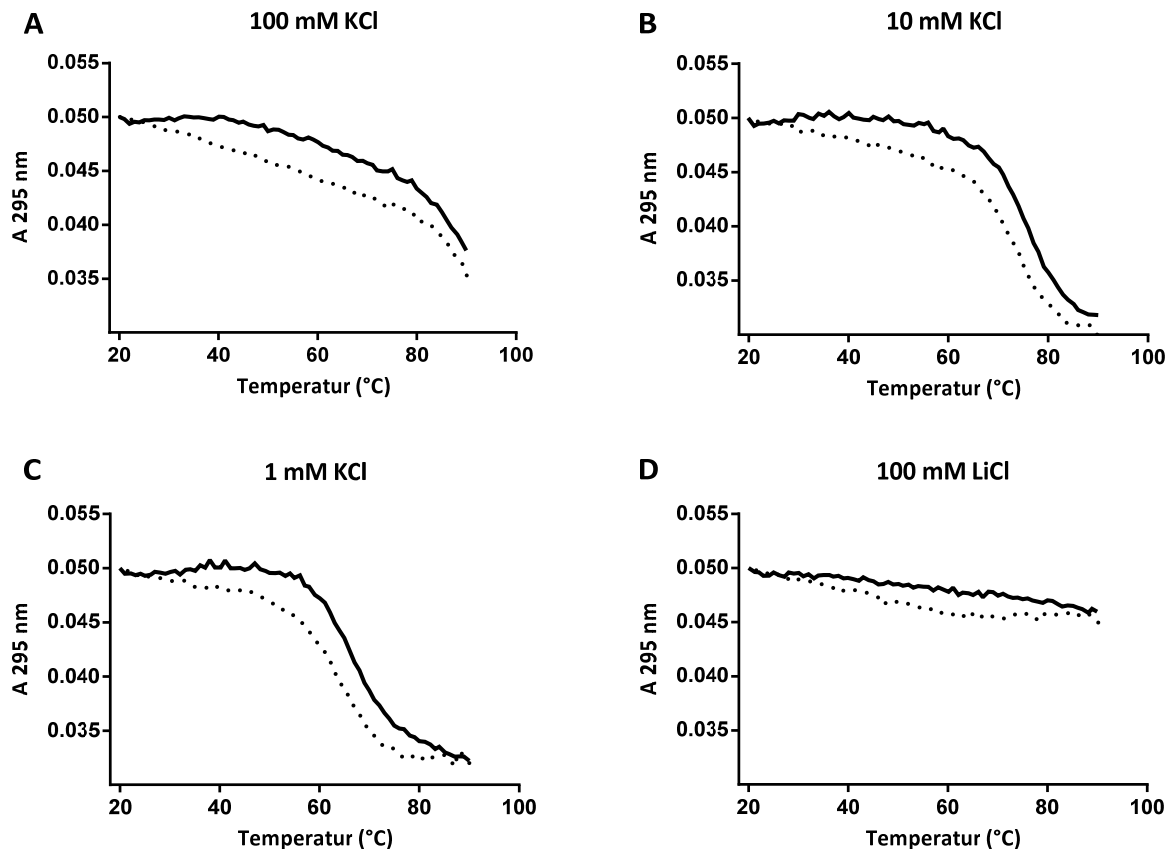


Abbildung 16: UV-Schmelzkurven des kurzen ARPC2 G-Quadruplex Oligonucleotids bei verschiedenen KCl-Konzentrationen und mit 100 mM LiCl als Kontrolle. Gezeigt sind repräsentative Kurven aus jeweils fünf Schmelzyklen. Die Messungen wurden bei 295 nm durchgeführt. Die durchgehenden Linien zeigen die Schmelzkurve, die gestrichelten Linien die Annealing-Kurve. Die Schmelztemperatur bei 100 mM KCl konnte nicht bestimmt werden, bei 10 mM KCl betrug sie 78°C, bei 1 mM KCl 63°C.

Abbildung 16 zeigt Schmelzkurven des kurzen ARPC2 G-Quadruplex Oligonucleotids bei verschiedenen KCl-Konzentrationen und mit 100 mM LiCl statt KCl im Puffer. Ohne die flankierenden Sequenzen hatte das ARPC2 Oligonucleotid eine höhere Schmelztemperatur als mit den flankierenden Sequenzen. Die Schmelztemperatur bei 100 mM KCl konnte nicht bestimmt werden, da sich die Struktur im gemessenen Temperaturbereich nicht vollständig auflöste. Bei einer Konzentration von 10 mM KCl betrug sie 78°C und bei 1 mM KCl 63°C.

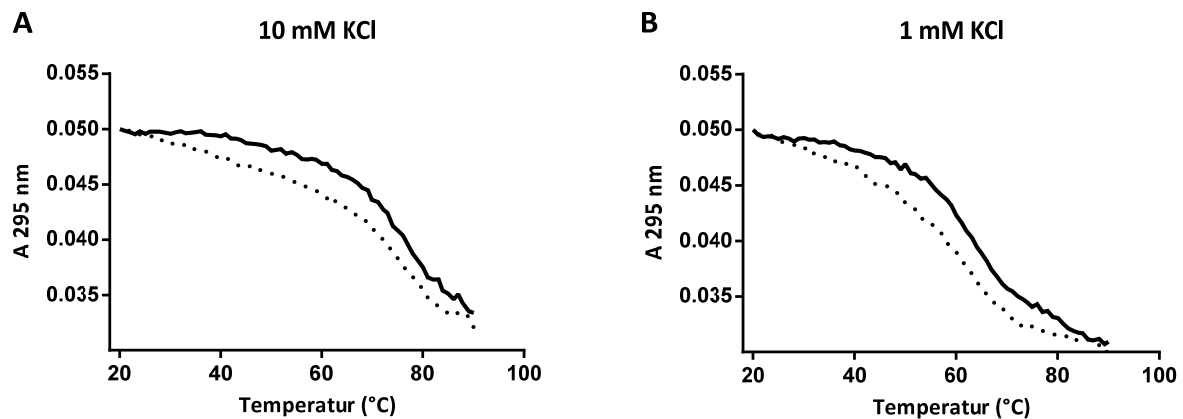


Abbildung 17: UV-Schmelzkurven des kurzen MMP16 G-Quadruplex Oligonucleotids bei 10 und 1 mM KCl. Gezeigt sind repräsentative Kurven aus jeweils fünf Schmelzyklen. Die Messungen wurden bei 295 nm durchgeführt. Die durchgehenden Linien zeigen die Schmelzkurve, die gestrichelten Linien die Annealing-Kurve.

Die Schmelzkurven des MMP16 Oligonucleotids ohne die flankierenden Sequenzen sind in Abbildung 17 gezeigt. Auch hier ist die Schmelztemperatur höher als bei dem Oligonucleotid mit den flankierenden Basen. Bei einer Konzentration von 1 mM KCl betrug sie 63°C.

4.3. EINFLUSS DER G-QUADRUPELX AUF DIE TRANSLATION

Der Einfluss bestimmter Nucleinsäuresequenzen auf die Translation kann mit Hilfe des Dual-Luciferase-Reporter-Assays überprüft werden. Hierfür wurde das Plasmid psiCHECK-2 benutzt, das sowohl die Renilla- als auch die Firefly-Luciferase exprimiert. Die Expressionskassette ist in Abbildung 18 dargestellt. Die zu untersuchenden G-Quadruplex oder Kontroll-Sequenzen wurden in die 5'-UTR der Renilla-Luciferase kloniert. Es wurden sowohl Plasmide mit den isolierten ARPC2 und MMP16 G-Quadruplex-Sequenzen konstruiert, als auch mit der 169 bp langen ARPC2 Vollängen-UTR und der entsprechenden mutierten Variante. Der Einfluss der MMP16 UTR auf die Translation wurde bereits beschrieben (39).

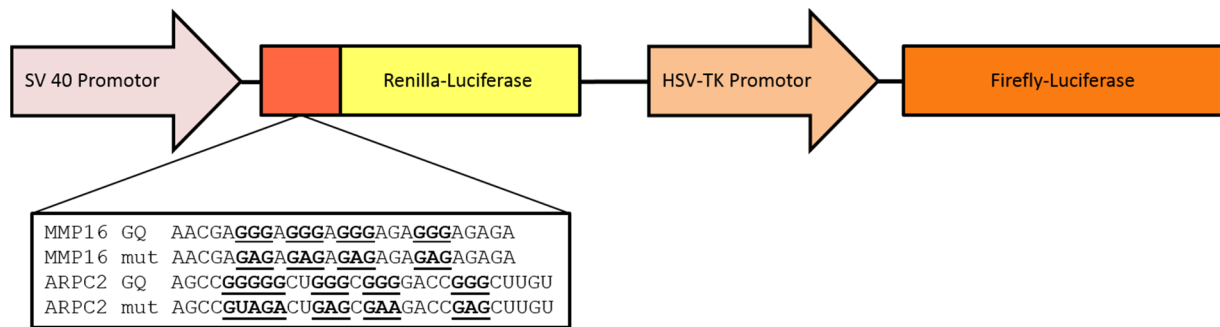


Abbildung 18: Luciferase-Expressionskassette des psiCHECK-2 Vektors. Die G-Quadruplex und Kontrollsequenzen wurden vor die Renilla-Luciferase kloniert (rot markiert). Im Kasten sind die Sequenzen gezeigt, die inseriert wurden. Die Nucleotide, die mutmaßlich an der Ausbildung des G-Quadruplex beteiligt sind sowie die entsprechenden mutierten Sequenzen, sind fett gedruckt und unterstrichen. Die Renilla-Luciferase wird von einem SV 40 (*simian virus 40*) Promotor exprimiert, die Firefly-Luciferase von einem HSV-TK (herpes simplex virus thymidine kinase) Promotor.

Der Einfluss der inserierten Sequenz auf die Translation konnte so direkt durch die Zunahme oder Abnahme der Renilla-Luciferase-Aktivität bestimmt werden. Unterschiede in den Transfektionseffizienzen oder Zellzahlen konnten durch Normierung der Renilla-Luciferase-Aktivität auf die Firefly-Luciferase-Aktivität ausgeglichen werden.

Der in Abbildung 18 beschriebene psiCHECK-2 Reportervektor mit den entsprechenden inserierten Sequenzen in der 5'-UTR der Renilla-Luciferase wurde in Hek293 Zellen transfiziert. Die Aktivitäten der beiden Luciferasen wurden 24 Stunden nach der Transfektion mit Hilfe eines Luminometers bestimmt.

Abbildung 19 zeigt die Ergebnisse der Luciferase-Assays für die isolierten G-Quadruplex-Sequenzen (A) und für die ARPC2 Vollängen-UTR (B). Die gemessenen Renilla-Luciferase-Aktivitäten wurden in Verhältnis zu den Firefly-Aktivitäten gesetzt. Alle Werte wurden auf die Aktivität des psiCHECK-2 Vektors normiert.

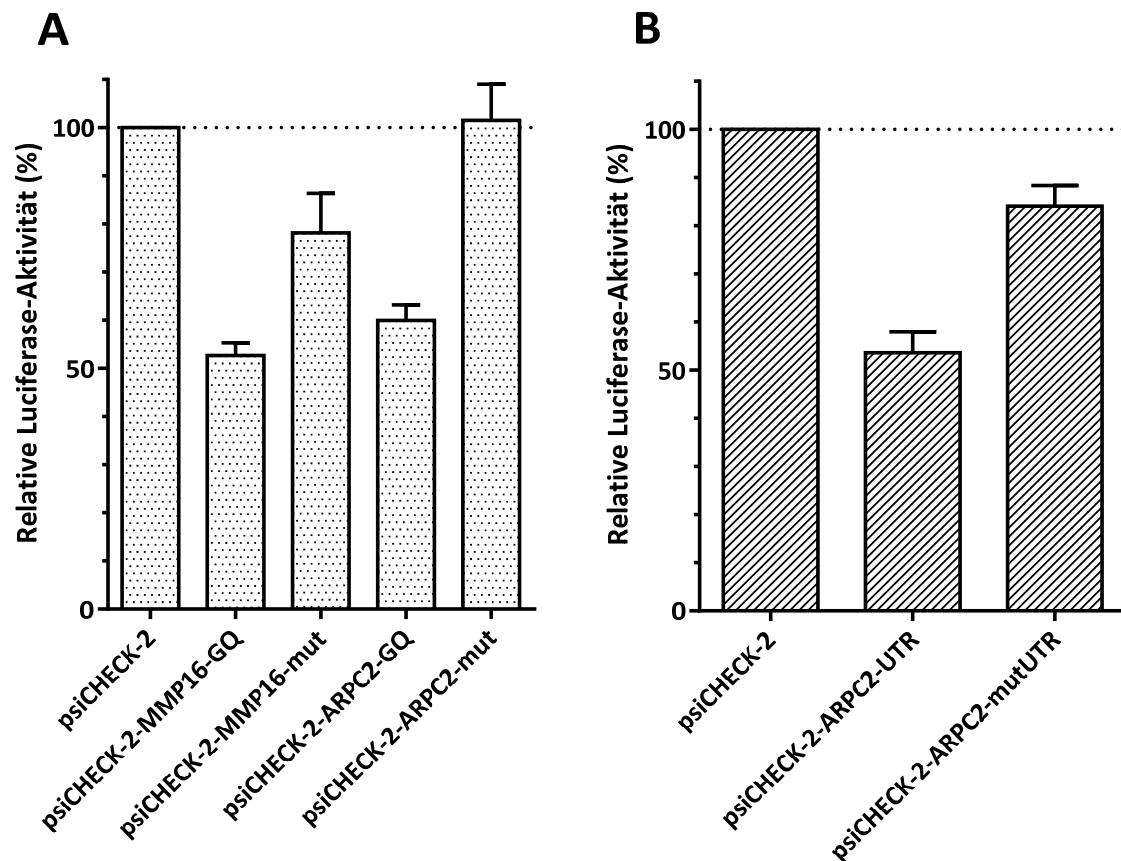


Abbildung 19: Dual-Luciferase Assay zur Untersuchung des Einflusses der G-Quadruplex-Sequenzen auf die Translation. Dargestellt sind Mittelwerte aus drei unabhängigen Experimenten in Duplikaten mit den jeweiligen Standardabweichungen. Die Luciferase-Aktivitäten wurden in Triplikaten gemessen. **(A)** zeigt die Wirkung der isolierten G-Quadruplex-Sequenzen und **(B)** die der RPC2 Vollängen-5'-UTR auf die Translation der Renilla-Luciferase. Alle Werte wurden auf die Firefly Luciferase-Aktivität normiert. Die Aktivität des psiCHECK-2 Vektors wurde auf 100 % gesetzt. Die Experimente wurden dreimal in Duplikaten durchgeführt, die Luciferase-Aktivitäten wurden mindestens zweimal bestimmt.

Die Insertion der MMP16 G-Quadruplex-Sequenz senkte die Renilla-Luciferase-Aktivität auf 53 %. Die mutierte MMP16 Sequenz bewirkte im Vergleich zum psiCHECK-2 Vektor ohne Insert eine Reduktion der Luciferase-Aktivität auf etwa 78 %.

Die RPC2 G-Quadruplex-Sequenz hemmte die Expression der Renilla-Luciferase um ca. 40 % und die entsprechende mutierte Sequenz RPC2 mut hatte im Vergleich zum psiCHECK-2 Vektor keinen Einfluss auf die Expression.

Die Insertion der RPC2 Vollängen-UTR führte zu einer Reduktion der Luciferase-Aktivität auf 54 %, während die Variante der UTR, in der nur die G-Quadruplex-Sequenz mutiert wurde, die Expression auf 84 % reduziert hat.

Alle G-Quadruplex-Sequenzen haben also zu einer Inhibition der Luciferase-Expression geführt. Die mutierten Varianten hatten keinen, oder einen geringeren Einfluss auf die Expression.

4.4. IDENTIFIZIERUNG RNA G-QUADRUPLIX BINDENDER PROTEINE

4.4.1. *PULL-DOWN-ASSAY*

Zur Identifizierung RNA G-Quadruplex bindender Proteine wurden Pull-down-Assays durchgeführt. Hierbei wurden Proteine aus zellulärem Proteinlysats selektiert, die eine hohe Affinität zu den G-Quadruplex-Sequenzen zeigten. Unspezifisch bindende Proteine sollten durch Vergleiche mit den Kontrollsequenzen ausgeschlossen werden. In jeweils getrennten Ansätzen wurden die biotinylierten MMP16 und ARPC2 G-Quadruplex Oligonucleotide (MMP16 GQ, ARPC2 GQ), sowie die entsprechenden mutierten Varianten (MMP16 mut, ARPC2 mut) an Streptavidin-Agarose-Beads immobilisiert. Diese Beads wurden anschließend mit Gesamtproteinlysats aus Hek293 oder HeLa Zellen inkubiert. In einem zusätzlichen Ansatz wurde das Zelllysats zuerst mit den mutierten Oligonucleotiden inkubiert. Dieses Lysats, aus dem unspezifisch an RNA bindende Proteine bereits eliminiert sein sollten, wurde dann auf die Beads mit den G-Quadruplex Oligonucleotiden gegeben. In allen Ansätzen wurden nicht-bindende Proteine in den Waschschritten entfernt. Proteine, die an die Oligonucleotide binden konnten, wurden mit steigenden KCl-Konzentrationen eluiert. Die Eluate wurden durch TCA-Fällung konzentriert. Anschließend wurden die gesamten Eluate in einer SDS-PAGE aufgetrennt und mittels kolloidaler Coomassiefärbung sichtbar gemacht.

4.4.2. *ANALYSE DER PULL-DOWN-ASSAYS*

Die Auswertung der Pull-down-Assays erfolgte nach der Auftrennung der selektierten Proteine auf SDS-PAGEs. Durch optischen Vergleich wurden für jede Elutionsfraktion Proteine ausgewählt, die nur für die G-Quadruplex Sequenzen angereichert waren, und die bei den entsprechenden Kontrollsequenzen auf dem Gel nicht erkennbar waren. Manche Proteinbanden, die wiederholt spezifisch für die mutierten Oligonucleotide auftraten, wurden ebenfalls analysiert. Abbildung 20 zeigt Beispiele von Pull-down-Assays mit den MMP16 (A) und ARPC2 (B, C) Oligonucleotiden.

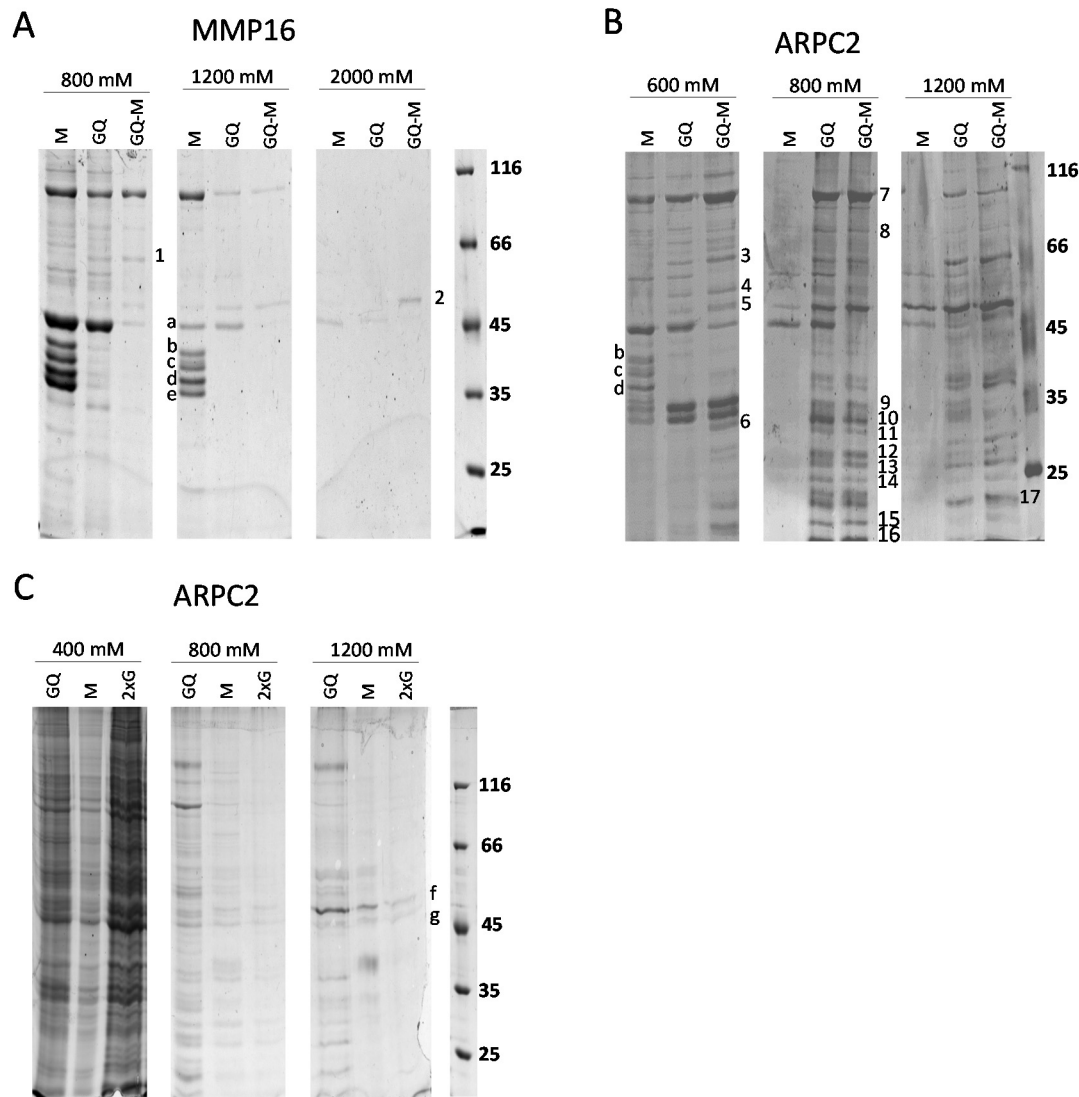


Abbildung 20: SDS-PAGEs der Pull-down-Assays (A) mit den MMP16 Oligonucleotiden, (B) mit den ARPC2 Oligonucleotiden und (C) ARPC2 mit der 2xG Kontrolle. Markiert sind die Banden, die ausgeschnitten und massenspektrometrisch identifiziert wurden. Über den einzelnen Fraktionen sind die KCl-Konzentrationen des Elutionsschritts angegeben. 1. ME2 (66 kDa); 2. YB-1 (36 kDa); 3. U2AF65 (54 kDa); 4. hnRNPH (50 kDa); 5. YB-1 (36 kDa) hnRNPF (46 kDa); 6. RPS2 (32 kDa); 7. Nucl (76 kDa); 8. RBM14 (70 kDa); 9. SRSF1 (28 kDa); 10. SRSF1 (28 kDa) RPS6 (29 kDa); 11. RPL7 (29 kDa); 12. SRSF9 (26 kDa), RPS6 (29 kDa); 13. SRSF9 (26 kDa), RPL14 (23,5 kDa); 14. RPL10 (25 kDa) RPS9 (23 kDa); 15. RPL26 (17 kDa); 16. RPL27a (17 kDa); 17. RPS9 (23 kDa); a. Actin (42 kDa); b. hnRNPA2B1 (37 kDa); c. hnRNPA2B1 (37 kDa) and hnRNPA3 (39 kDa); d. hnRNP A2/B1 (37 kDa); e. hnRNPA1 (39 kDa); f. YB-1 (36 kDa); g. YB-1/Actin (36 kDa, 42 kDa).

Für das ARPC2 G-Quadruplex-Oligonucleotid waren in allen Versuchen die meisten Proteinbanden erkennbar. Dies galt sowohl für den Vergleich mit den ARPC2 Kontrolloligonucleotiden (ARPC2 mut und ARPC2 2xG) als auch für den Vergleich mit den MMP16 Sequenzen. An das ARPC2 2xG Oligonucleotid, welches als Kontrolle für Proteine diente, die Abfolgen von Guaninen erkennen banden nur sehr wenige Proteine. Die meisten G-Quadruplex bindenden Proteine wurden bei KCl-Konzentrationen von 800 mM und 1200 mM identifiziert. Bei den geringeren Konzentrationen banden die Proteine tendenziell mehr unspezifisch an beide Sequenzen. Bei der höchsten KCl-Konzentration von 2600 mM war meist keine Proteinbande mehr zu erkennen.

Sowohl für MMP16, als auch für ARPC2 wurden die Pull-down-Assays dreimal mit Hek293 Proteinlysats und zweimal mit HeLa Proteinlysats wiederholt. Die weiteren Abbildungen der Pull-down-Assay SDS-PAGEs befinden sich im Anhang. Die ausgewählten Proteinbanden wurden mittels Massenspektrometrie identifiziert.

4.4.3. MASSENSPEKTROMETRISCHE ANALYSE DER PROTEINBANDEN

Proteinbanden, die von Interesse waren, wurden aus den SDS-PAGEs ausgeschnitten und die Proteine wurden im Gelstück mit Trypsin verdaut. Trypsin ist eine Serinprotease, die Peptidbindungen spezifisch nach Arginin- und Lysinresten spaltet. Dabei entstehen für jedes Protein spezifische Spaltprodukte. Die Massen dieser Spaltprodukte wurden im Massenspektrometer bestimmt². Anhand dieses Spektrums konnten die Proteine durch einen Datenbankvergleich mit Mascot (Matrix Science Ltd., <http://www.matrixscience.com>) identifiziert werden. Die Ergebnisse sind in Tabelle 13 zusammengefasst.

In Abbildung 20 und Tabelle 13 kann man erkennen, dass an die ARPC2 G-Quadruplex-Sequenz am meisten Proteine gebunden haben. Für die MMP16 G-Quadruplex-Sequenz wurden weniger Binder gefunden. Auffällig waren vier sehr prominente Proteinbanden, die in jedem Pull-down sehr stark an die mutierte MMP16 Sequenz banden. Diese wurden identifiziert als die heterogenen nukleären Ribonucleoproteine hnRNP A2/B1, hnRNP A3, hnRNP A2/B1 und hnRNP A1. Die Proteine hnRNP A2/B1, hnRNP A3 und hnRNP A2/B1 wurden auch als Bindungspartner der ARPC2 Kontrolle identifiziert, jedoch nicht so ausgeprägt und auch nicht mehr bei so hohen KCl-Konzentrationen wie für MMP16 mut.

Der Kontrollansatz, in dem das Proteinlysats zuerst mit der mutierten Sequenz und anschließend mit der G-Quadruplex-Sequenz inkubiert wurde (G-Quadruplex-M), zeigte vergleichbare Ergebnisse zu dem G-Quadruplex Ansatz, erleichterte es aber, unspezifische Binder auszusortieren.

Während die mutierten G-Quadruplex-Sequenzen (MMP16 mut und ARPC2 mut) vor allem dazu dienen sollten, zwischen Proteinen die G-Quadruplex-Strukturen erkennen und solchen die sequenzspezifisch binden, zu unterscheiden, sollte die Kontrolle ARPC2 2xG Proteine ausschließen, die Abfolgen von Guaninen erkennen. An diese Kontrolle haben nur sehr wenige Proteine gebunden. YB-1 und Aktin wurden hier als Binder identifiziert.

Aktin und Myosin banden an alle Oligonucleotide, sowohl an die G-Quadruplex-Sequenzen, als auch an die Kontrollen, und wurden daher als unspezifische Binder nicht weiter berücksichtigt.

² Die massenspektrometrische Analyse wurde von Dr. Christoph Weise an der FU Berlin durchgeführt.

Der größte Teil der Proteine, die mit den beiden G-Quadruplex-Sequenzen interagierten, gehörte zur Gruppe der ribosomalen Proteine. Sowohl Proteine der 60S als auch der 40S Untereinheit wurden gefunden. Eine weitere große Gruppe bildeten Proteine, die am Spleißen beteiligt sind. Hierzu zählen die hnRNPs F, H und U, SRSF1 und 9 sowie U2AF65. U2AF65 wurde als einziges Protein für beide G-Quadruplex-Sequenzen als spezifischer Binder identifiziert. YB-1 wurde zwar auch für beide Sequenzen gefunden, es zeigte sich aber in weiteren Kontrollexperimenten, dass es auch eine gewisse Affinität zu den Kontrollsequenzen zeigte (Abbildung 20, Banden 2, 5, f und g).

Weitere Proteine, die häufig gefunden wurden, waren für die ARPC2 G-Quadruplex-Sequenz das Calcium-bindende Protein *EF-hand-domain containing protein 2* (EFHD2) und Nucleolin (Nucl). Für die MMP16 G-Quadruplex-Sequenz wurde das mitochondriale Malatenzym *NAD-dependent malic enzyme* (ME2) identifiziert, das die oxidative Decarboxylierung von Malat zu Pyruvat in den Mitochondrien katalysiert.

In einem weiteren Ansatz wurde überprüft, ob die Interaktionen zelltypspezifisch sind. Dazu wurden die Pull-down Experimente mit HeLa Proteinlysate wiederholt. Generell haben in diesen Experimenten weniger Proteine an die G-Quadruplex-Oligonucleotide gebunden. Vor allem für MMP16 wurden hier nur zwei G-Quadruplex-bindende Proteine gefunden. Die Proteine, die für ARPC2 G-Quadruplex identifiziert wurden, wurden alle auch schon für Hek293 gefunden, bis auf das *Ras GTPase-activating protein-binding protein 2* (G3BP2).

Tabelle 13: Zusammenfassung der Ergebnisse der Pull-down-Assays. Aufgeführt sind nur die Proteine, die an die G-Quadruplex, aber nicht an die mutierten Oligonucleotide gebunden haben. Für jedes Protein ist die Anzahl angegeben, die es für das jeweilige G-Quadruplex identifiziert wurde. Die hochgestellten Zahlen verweisen auf die Banden in Abbildung 20. Die Pull-down-Assays wurden dreimal mit Hek293-Zelllysats und zweimal mit HeLa-Zelllysats durchgeführt.

	ARPC2 G- Quadruplex		MMP16 G- Quadruplex	
	Hek293	HeLa	Hek293	HeLa
Ribosomale Proteine				
<i>60S ribosomal protein L6</i> (RPL6 ¹⁰)	2	1		
<i>60S ribosomal protein L7</i> (RPL7 ¹¹)	2	1		
<i>60S ribosomal protein L10</i> (RPL10 ¹⁴)	2			
<i>60S ribosomal protein L12</i> (RPL12)		1		
<i>60S ribosomal protein L14</i> (RPL14 ¹³)	1			
<i>60S ribosomal protein L19</i> (RPL19)	1	1		
<i>60S ribosomal protein L26</i> (RPL26 ¹⁵)	1			
<i>60S ribosomal protein L27a</i> (RPL27a ¹⁶)	1			
<i>40S ribosomal protein S2</i> (RPS2 ⁶)	1			
<i>40S ribosomal protein S5</i> (RPS5)				1
<i>40S ribosomal protein S6</i> (RPS6 ¹²)	1			
<i>40S ribosomal protein S9</i> (RPS9 ^{14, 17})	2	1		
Spleiß Faktoren				
<i>Heterogeneous nuclear ribonucleoprotein F</i> (hnRNPF ⁵)	1			
<i>Heterogeneous nuclear ribonucleoprotein H</i> (hnRNPH ⁴)	2			
<i>Heterogeneous nuclear ribonucleoprotein U</i> (hnRNPU)			2	1
<i>Serine/arginine-rich splicing factor 1</i> (SRSF1 ^{9, 10})	3	1		
<i>Serine/arginine-rich splicing factor 9</i> (SRSF9 ^{12, 13})	2	1		
<i>Splicing factor U2AF</i> (U2AF65 ³)	1	1	2	
Andere				
<i>Nuclease-sensitive element-binding protein 1</i> (YB-1)	1	1	1	
<i>Y-box-binding protein 3</i> (YB-3)	1			
<i>EF-hand domain-containing protein D2</i> (EFHD2)	2	1		
<i>NAD-dependent malic enzyme, mitochondrial</i> (ME2 ¹)			3	
<i>Nucleolin</i> (Nucl ⁷)	3	2		
<i>Ras GTPase-activating protein-binding protein 2</i> (G3BP2)		1		
<i>RNA binding protein 14</i> (RBM14 ⁸)	2			
<i>Tropomyosin alpha 4</i> (TPM4)	1			

Zur weiteren Untersuchung der Interaktionen zwischen G-Quadruplex und Proteinen wurden SRSF1, SRSF9, U2AF65, EFHD2, Nucleolin, YB-1 und ME2 ausgewählt. Diese Proteine sollten in *E. coli* exprimiert und die Bindung zwischen Protein und G-Quadruplex sollte mittels Oberflächenplasmonresonanz bestätigt bzw. charakterisiert werden.

4.5. PROTEINEXPRESSION

Die ausgewählten Proteine wurden zur Produktion in *E. coli* in den prokaryotischen Expressionsvektor pET28a kloniert. Die Proteine wurden so kloniert, dass sie einen N-terminalen Polyhistidin-Tag (His-Tag) erhielten und so über eine Nickel-NTA (Nitrilotriessigsäure) Säule aufgereinigt werden konnten.

Tabelle 14: Proteine, die zur heterologen Expression in *E. coli* ausgewählt wurden.

Protein	Abkürzung	Molekulargewicht	Länge (Aminosäuren)
Nucleolin	Nucl	76 kDa	710
Splicing factor U2AF 65 kDa subunit	U2AF65	54 kDa	475
EF-hand domain-containing protein D2	EFHD2	27 kDa	240
Serine/arginine-rich splicing factor 1	SRSF1	28 kDa	248
Serine/arginine-rich splicing factor 9	SRSF9	26 kDa	221
Nuclease-sensitive element-binding protein 1	YB-1	36 kDa	324

Nucleolin ist nicht löslich, wenn es in *E. coli* exprimiert wird, daher wurde eine verkürzte Variante (Δ Nucleolin) verwendet, der die N-terminale Domäne (Aminosäuren 1-293) fehlte.

Hohe Ausbeuten wurden mit der Expression der Proteine in EnBase Medium erzielt. Die langsame Freisetzung von Glucose in diesem Medium führt zu einem langsameren Wachstum von *E. coli* und zu einer gemäßigten Proteinsynthese, die eine korrekte Faltung der Proteine ermöglicht.

Abbildung 21 zeigt eine SDS-PAGE einiger Proteine nach ihrer Reinigung. Für die markierten Banden wurde nach einem In-Gel Trypsinverdau massenspektrometrisch überprüft, ob die richtigen Produkte gebildet wurden. Alle gemessenen Massen der Verdauprodukte ließen sich den entsprechenden Proteinen zuordnen.

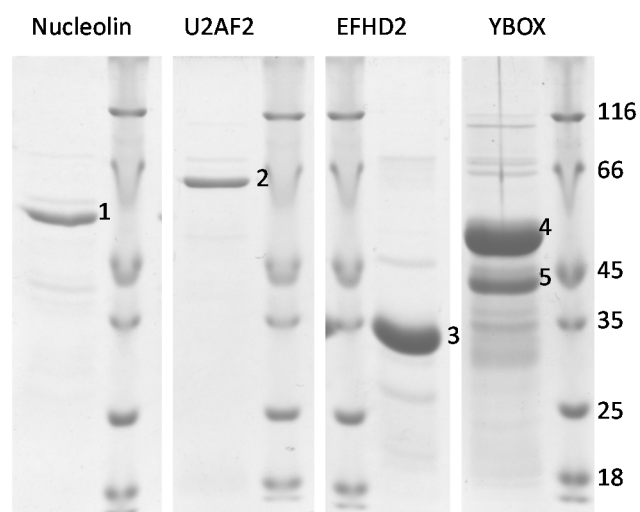


Abbildung 21: Die Proteine Δ Nucleolin, U2AF65, EFHD2 und YB-1 nach Expression in *E. coli* und Reinigung über Ni-NTA Säulen. Die markierten Banden wurden massenspektrometrisch validiert. 1. Δ Nucleolin, 2. U2AF65, 3. EFHD2, 4. YB-1, 5. YB-1 (C-terminal trunkiert)

4.6. INTERAKTIONSTUDIEN

4.6.1. NUCLEOLIN

Nucleolin wurde in den Pull-down-Assays als Binder des ARPC2 G-Quadruplex identifiziert. Es wurde in allen drei Experimenten mit Hek293 und in beiden Experimenten mit HeLa Zelllysate für das ARPC2 G-Quadruplex gefunden und war damit auch das am häufigsten identifizierte Protein. Nucleolin ist 76 kDa groß und besteht aus 6 Domänen, einem sauren N-Terminus, vier RNA-Binde-Domänen (RBDs) und einem *RNA recognition motif*. Der saure N-Terminus macht das Protein in *E. coli* unlöslich, weshalb es ohne diesen Teil exprimiert wurde.

Zur ersten Charakterisierung der Bindung von Nucleolin an die ARPC2 G-Quadruplex und mit Oligonucleotide, sowie zur Überprüfung der Funktionalität des Proteins, wurden Electrophoretic Mobility Shift Assays (EMSAs) durchgeführt. Dabei kann eine Retardation des Laufs einer Nucleinsäure im Gel durch Bindung eines Proteins beobachtet werden. Die Oligonucleotide wurden vor dem Assay zur Ausbildung der G-Quadruplex-Struktur erhitzt und langsam abgekühlt. Anschließend wurden zu je 20 pmol Oligonucleotid steigende Mengen an Protein gegeben. Die Ansätze wurden 30 min bei 30°C inkubiert, was den Bindungsbedingungen im Pull-down-Assay entsprach. Die Protein-Nucleinsäuregemische wurden dann auf 10%igen nativen Polyacrylamidgelen aufgetrennt und die Nucleinsäuren mit GelRed gefärbt.

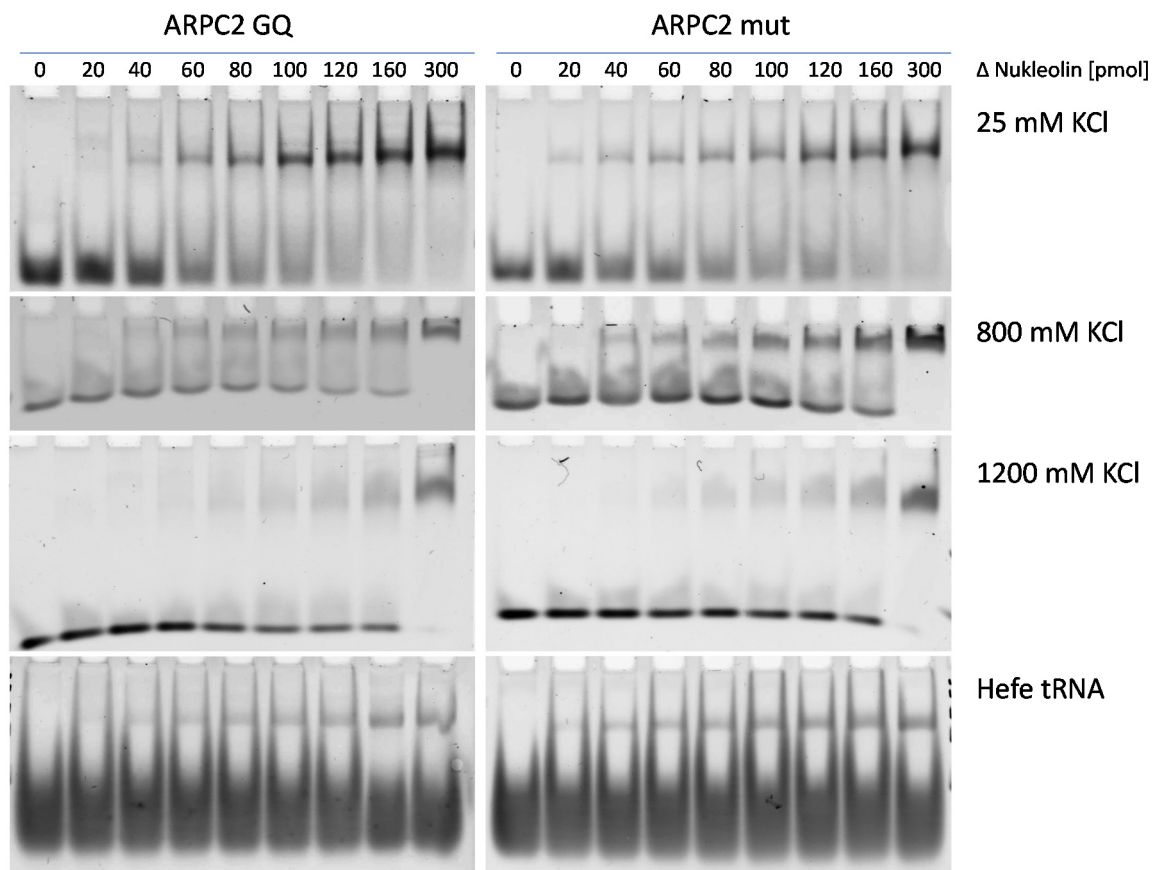


Abbildung 22: Electrophoretic mobility shift Assay mit Δ Nucleolin und den ARPC2 Oligonucleotiden. Die ARPC2 G-Quadruplex und mut Oligonucleotide wurden mit den angegebenen Mengen an Δ Nucleolin in TRIS-Puffer mit verschiedenen KCl-Konzentrationen sowie mit Hefe tRNA als Kompetitor 30 min inkubiert. Anschließend wurden die Ansätze auf 10%igen Polyacrylamidgelen aufgetrennt. Die Oligonucleotide wurden mit dem interkalierenden Farbstoff GelRed sichtbar gemacht.

Abbildung 22 zeigt verschiedene Gelretardationsexperimente mit Δ Nucleolin und den ARPC2 Oligonucleotiden. Die Experimente wurden mit verschiedenen KCl-Konzentrationen durchgeführt, da die Bindung zwischen Nucleinsäure und Protein von der Ionenstärke beeinflusst wird. Bei einer KCl-Konzentration von 25 mM kann man sehr gut eine Retardation der Oligonucleotide im Gel bei steigenden Proteinkonzentrationen beobachten. Schon bei einer Menge von 20 pmol Nucleolin kann man eine retardierte Bande im Gel erkennen, bei einer Menge von 300 pmol Protein, was einem 15fachen Überschuss entspricht, ist fast das komplette Oligonucleotid gebunden. Das Protein wurde also funktional exprimiert. Jedoch ließ sich kein Unterschied in der Gelretardation zwischen der G-Quadruplex-Sequenz und der mutierten Sequenz feststellen.

In den Pull-down-Assays zeigte sich bei dem 600 mM KCl Elutionsschritt für das mutierte Oligonucleotid auch noch eine Bande für Nucleolin, erst ab 800 mM KCl war ein Unterschied zwischen der Bindung von Nucleolin an das G-Quadruplex und das mutierte Oligonucleotid erkennbar. Um diesen Unterschied herauszuarbeiten, wurden EMSAs bei KCl-Konzentrationen von 800 mM und 1200 mM durchgeführt. Mit steigender Salzkonzentration trat die Verschiebung der Nucleinsäurebanden erst

später auf. Allerdings ließ sich auch hier zwischen den beiden Oligonucleotiden kein Unterschied in der Gelretardation beobachten.

Nucleolin besitzt fünf RNA Bindedomänen und ist für vielfältige Wechselwirkungen mit Nucleinsäuren bekannt, daher lag die Vermutung nahe, dass es nicht nur G-Quadruplexe erkennt, sondern auch andere RNA-Bindekapazitäten besitzt und so auch unspezifisch mit dem mutierten Oligonucleotid interagieren kann (95). Daher wurde ein Experiment mit Hefe tRNA als Kompetitor durchgeführt. Das Δ Nucleolin wurde zuerst 15 min bei 30°C mit 50 μ g Hefe tRNA inkubiert, anschließend wurden die ARPC2 Oligonucleotide zugegeben und die Ansätze für weitere 15 min bei 30°C inkubiert. Auch durch diesen Ansatz konnte keine erhöhte Spezifität des Nucleolins für G-Quadruplexe gezeigt werden.

Die Bindung zwischen Δ Nucleolin und den Oligonucleotiden wurde weiter durch Oberflächenplasmonresonanz (*surface plasmon resonance*, SPR) Spektroskopie untersucht (Abbildung 23). Δ Nucleolin wurde durch Ethanolamin-Kopplung der Aminogruppen kovalent immobilisiert. Als Analyten wurden die ARPC2 und MMP16 Oligonucleotide in steigenden Konzentrationen über eine definierte Zeit (90 s) in die Messkammer injiziert, aus dem Verlauf dieser Bindung lässt sich die Assoziationsgeschwindigkeitskonstante (k_a) ermitteln. Der Injektion des Analyten folgt die Dissoziationsphase, aus der die Dissoziationsgeschwindigkeitskonstante (k_d) ermittelt werden kann.

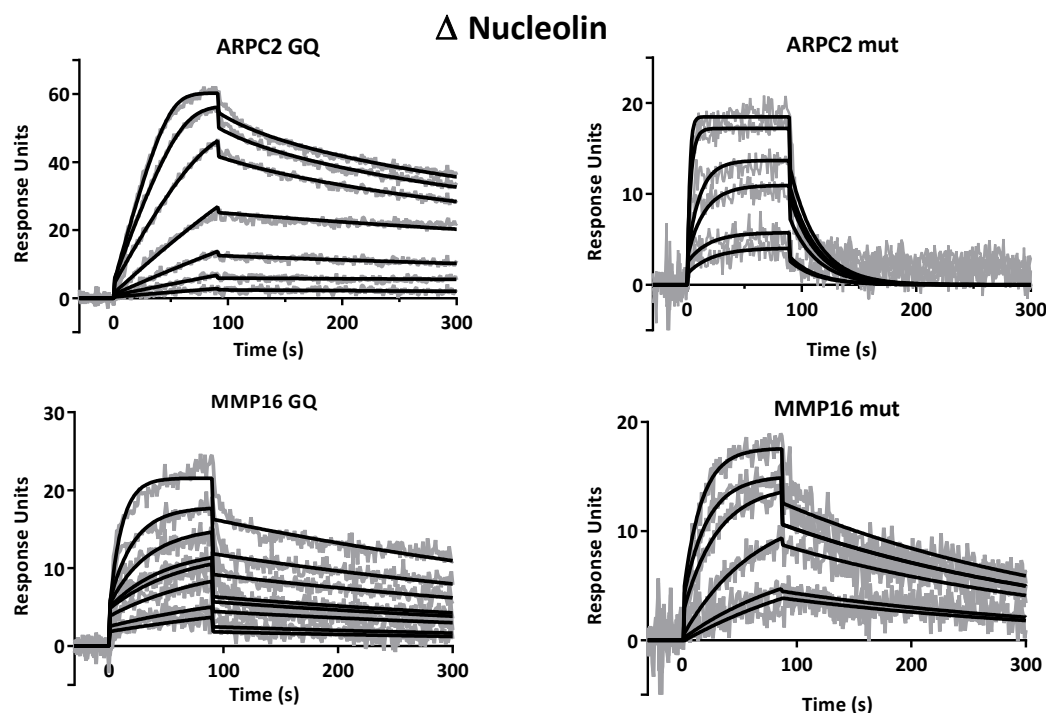


Abbildung 23: SPR-Spektroskopie der Interaktion von Δ Nucleolin mit den ARPC2 und MMP16 Oligonucleotiden. Die gemessenen SPR-Signale sind in grau dargestellt, die durch den 1:1 Langmuir-Fit erhaltenen Kurven sind in schwarz darübergelegt. 30 μ L je Oligonucleotid wurden mit einer Flussrate von 20 μ L/min in die Messkammer injiziert. Verwendete Oligonucleotidkonzentrationen: ARPC2 G-Quadruplex: 0,5, 1, 2,5, 5, 10, 15, 20 nM; ARPC2 mut: 10, 25, 50, 100, 500, 1000 nM; MMP16 G-Quadruplex: 2,5, 5, 10, 15, 20, 30, 50, 100 nM; MMP16 mut: 0,5, 1, 2, 5, 7,5, 10 nM. Zur Regeneration der Oberfläche wurde nach jeder Messung 1 M KCl injiziert.

Abbildung 23 zeigt die Sensorgramme (grau) und die Anpassungen an das Langmuir Modell (schwarz) für die Bindung der einzelnen Oligonucleotide an Δ Nucleolin. Für jedes Oligonucleotid wurde ein neuer Chip mit Protein beladen, da das Δ Nucleolin über die Zeit an Aktivität verlor, was man durch wiederholte Injektion derselben Oligonucleotidkonzentration feststellen kann. Die Immobilisierung exakt gleicher Proteinmengen war technisch nicht möglich. Daher kann hier die Größe des Signals für die verschiedenen Oligonucleotide nicht einfach miteinander verglichen werden. Die ermittelten kinetischen Parameter der Bindungen sind in Tabelle 15 aufgeführt. Die Güte der Anpassung an das Langmuir Modell ist durch den χ^2 Wert angegeben, je kleiner der Wert, desto besser beschreibt das Modell die Bindungsreaktion.

Tabelle 15: Bindungsparameter von Δ Nucleolin und den ARPC2 und MMP16 Oligonucleotiden. Die Bindungskonstanten wurden aus den Messungen ermittelt, die in Abbildung 23 gezeigt sind. k_a : Assoziationsgeschwindigkeitskonstante; k_d : Dissoziationsgeschwindigkeitskonstante; K_D : Dissoziationskonstante, χ^2 : Maß für die Güte des *Fits*.

Ligand	Analyt	k_a ($M^{-1}s^{-1}$) $\times 10^6$	k_d (s^{-1}) $\times 10^{-3}$	K_D (nM)	χ^2
Δ Nucleolin	ARPC2 GQ	8,7	5,1	0,6	0,96
	ARPC2 mut	0,5	41,4	83	1,61
	MMP16 GQ	0,9	1,9	2	0,79
	MMP16 mut	6,0	3,6	0,6	0,94

Für die Bindung zwischen Δ Nucleolin und ARPC2 G-Quadruplex wurde eine Dissoziationskonstante von 0,6 nM bestimmt. Für das mutierte ARPC2 Oligonucleotid wurde eine K_D von 83 nM bestimmt, also mehr als 10-mal größer als für das G-Quadruplex. Schon am Verlauf der Bindungskurven konnte man erkennen, dass die Bindung an das ARPC2 G-Quadruplex stärker war. Die Assoziation des mutierten Oligonucleotids erreichte schnell einen stationären Zustand, in dem Assoziation und Dissoziation gleich groß sind. Dieser Gleichgewichtszustand zeigt sich durch das Erreichen eines Plateaus in der Bindungskurve. Aus dem Signal, das in der stationären Phase erreicht wird (R_{eq}), lässt sich ebenfalls die Dissoziationskonstante ermitteln. Dazu werden die R_{eq} Werte über die Konzentrationen aufgetragen (Abbildung 24).

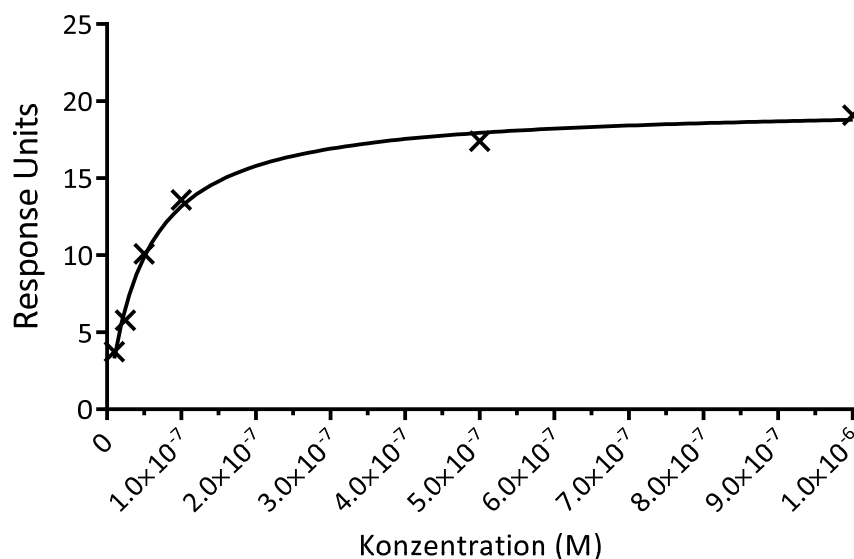


Abbildung 24: Gleichgewichtsanalyse zur Bestimmung der Dissoziationskonstanten von ARPC2 mut und Δ Nucleolin. Die Response Units im Gleichgewichtszustand wurden gegen die Konzentration aufgetragen um die Bindungsparameter zu ermitteln. Diese Methode ergab eine K_D von 50 nM ($\chi^2 = 0,367$).

Die so für ARPC2 mut ermittelte K_D betrug 50 nM und lag damit in derselben Größenordnung wie die Konstante, die durch das Langmuir Modell erhalten wurde. Durch dieses Modell lassen sich die Geschwindigkeitskonstanten für Bindung und Dissoziation (k_a und k_d) nicht ermitteln, jedoch liefert diese zusätzliche Auswertung der Daten einen Hinweis auf die Richtigkeit der durch das kompliziertere Langmuir Modell ermittelten Parameter.

Die Injektionszeit des ARPC2 G-Quadruplex Oligonucleotids reichte nicht aus um die stationäre Phase zu erreichen. Es konnte immer noch mehr Oligonucleotid an das Protein binden als abdissoziierte. Nach Beendigung der Injektion, in der Dissoziationsphase, erreichte das Signal nach der Bindung des mutierten Oligonucleotids schnell wieder den Ausgangswert, während die Dissoziation des ARPC2 G-Quadruplex Oligonucleotids nur sehr langsam war. Das ARPC2 Oligonucleotid konnte so stark an Δ Nucleolin binden, dass die Injektion einer Regenerationslösung nötig war um das Signal wieder auf den Ausgangswert zu bringen.

Dieses Resultat validierte die Ergebnisse aus den Pull-down-Assays, in denen Nucleolin mehrfach als Binder von ARPC2 G-Quadruplex identifiziert wurde.

Δ Nucleolin konnte ebenfalls stark an die mutierte MMP16 Kontrolle binden. Die bestimmte K_D betrug 0,6 nM und war damit genauso groß wie für ARPC2 G-Quadruplex, während die Bindung an das MMP16 G-Quadruplex mit einer K_D von 2 nM etwas geringer war.

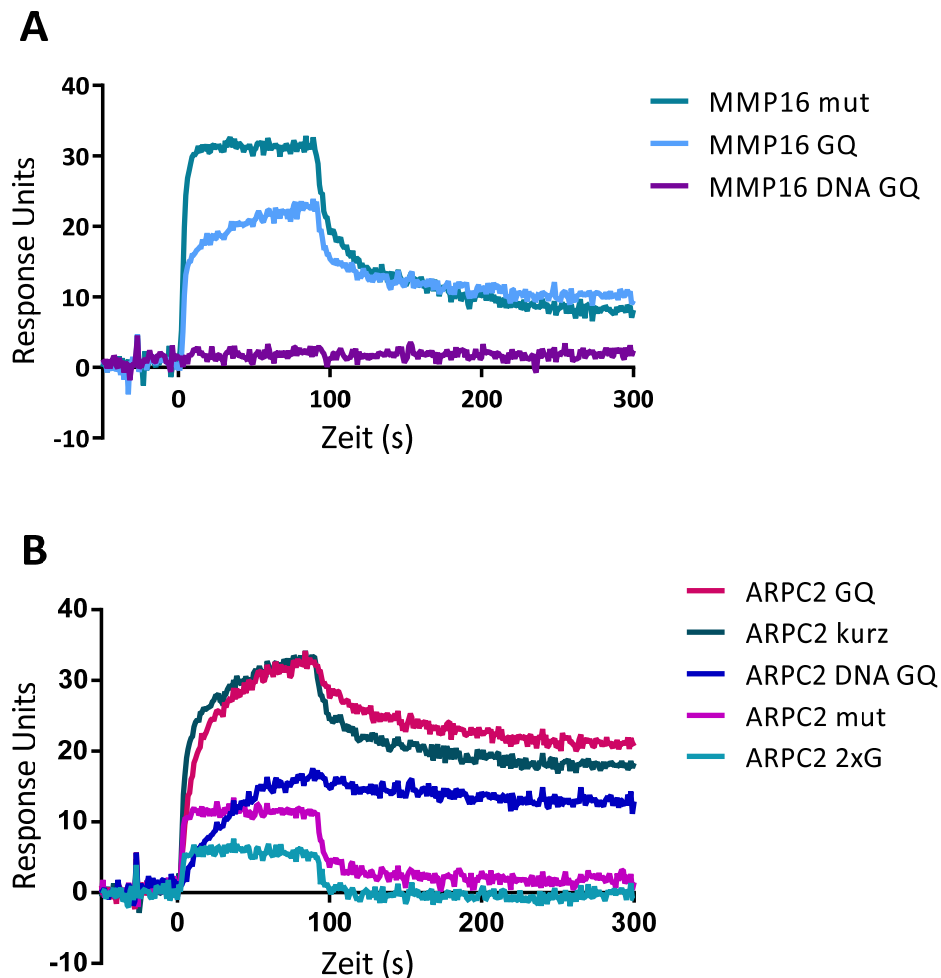


Abbildung 25: Oberflächenplasmonresonanzmessungen zum Vergleich der Bindung von verschiedenen Oligonucleotidsequenzen an Δ Nucleolin. (A) ARPC2-Oligonucleotide und (B) MMP16 Oligonucleotide wurden in einer Konzentration von 100 nM injiziert. Zur Regeneration der Oberfläche wurde 1 M KCl-Lösung verwendet.

Des Weiteren wurde die Bindung von Δ Nucleolin an verschiedene Kontrollsequenzen verglichen (Abbildung 25). Untersucht wurde die Bindung an die entsprechenden DNA Oligonucleotide der RNA G-Quadruplex-Sequenzen, um zu testen ob die Bindung spezifisch für DNA ist. Für das MMP16 G-Quadruplex DNA Oligonucleotid wurde keine Interaktion mit Δ Nucleolin beobachtet. Die Affinität von Δ Nucleolin für das ARPC2 DNA Oligonucleotid schien weniger stark zu sein als für das RNA Oligonucleotid. Um eine fundierte Aussage treffen zu können, müsste hier jedoch auch eine Bindungskonstante bestimmt werden. Für die ARPC2 G-Quadruplex-Sequenz wurde ein kürzeres Oligonucleotid designt, das nicht die flankierenden Basen enthält. Durch diese Basen wäre es möglich, dass sich eine Haarnadelstruktur ausbildet. Die Ausbildung dieser zusätzlichen Struktur könnte auch das Bindungsverhalten beeinflussen. In den SPR-Messungen konnte man keinen Unterschied im Bindungsverhalten zwischen den beiden Sequenzen erkennen. Eine weitere Kontrollsequenz enthielt zwei der vier Guanin Abfolgen. Mit dieser Kontrolle sollte ausgeschlossen werden, dass das Protein

nicht die G-Quadruplex-Struktur, sondern die G-reiche Sequenz erkennt. Die Affinität von Δ Nucleolin zu dieser Kontrolle war ähnlich gering wie die Affinität zu der ARPC2 mut Kontrolle.

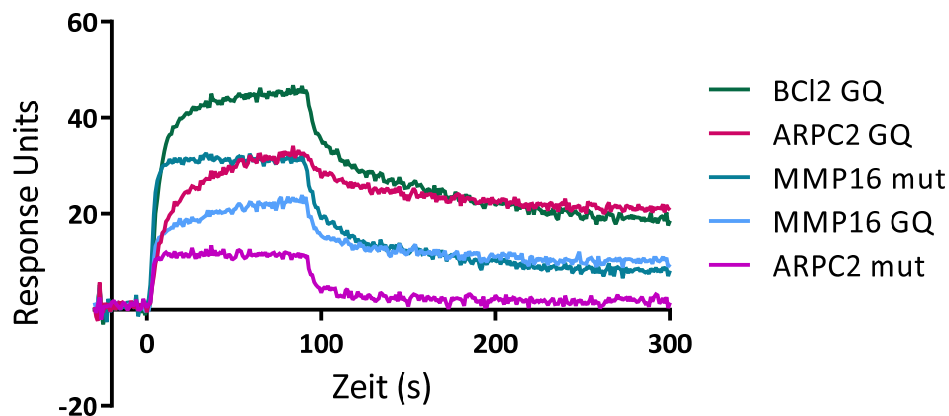


Abbildung 26: SPR-Messung zum Vergleich der Bindung der Bcl-2 G-Quadruplex-Sequenz an Δ Nucleolin. Die Oligonucleotide wurden in einer Konzentration von 100 nM in die Flusszelle injiziert. Zur Regeneration der Oberfläche wurde 1 M KCl-Lösung verwendet.

4.6.2. U2AF65

Ein besonders interessantes Protein war *U2 auxiliary factor 65* (U2AF65), da es als einziges sowohl für MMP16 G-Quadruplex, als auch für ARPC2 G-Quadruplex als Binder identifiziert wurde. Nach der heterologen Expression mit Hilfe des Expressionsvektors pET28a in *E. coli* und der Reinigung des Proteins wurden auch hier als erstes Electrophoretic Mobility Shift Assays durchgeführt, um die Funktionalität des Proteins zu überprüfen.

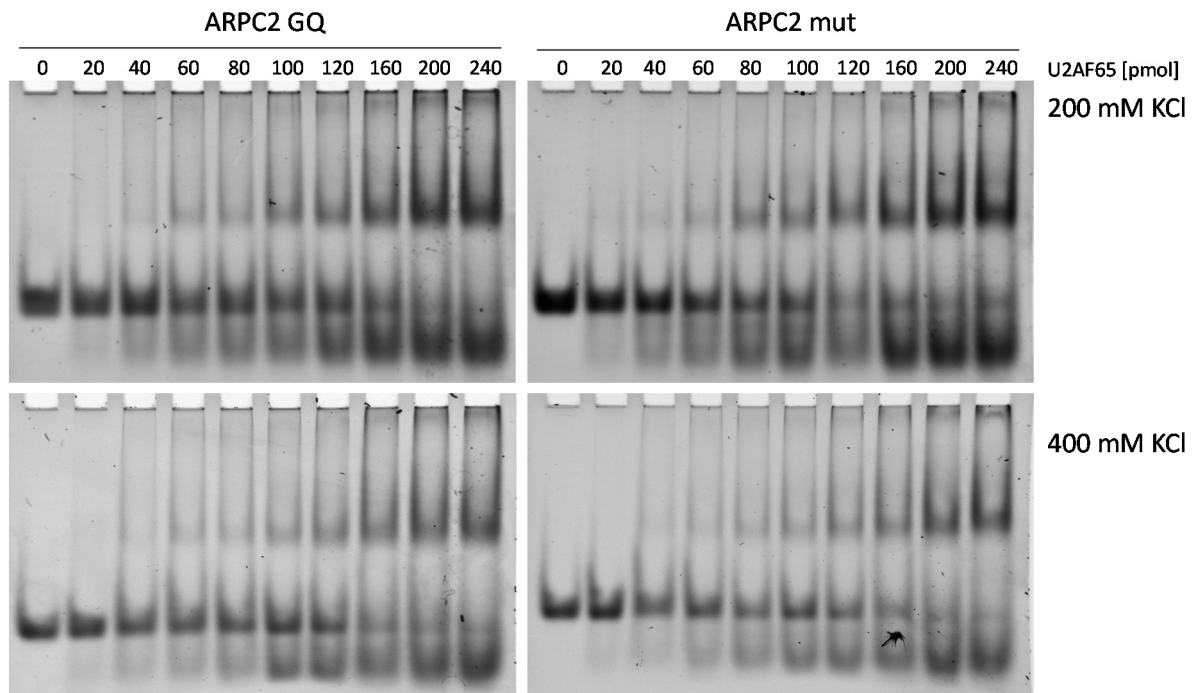


Abbildung 27: Electrophoretic Mobility Shift Assays mit U2AF65. Die ARPC2 G-Quadruplex und mut Oligonucleotide wurden mit den angegebenen Mengen an U2AF65 in TRIS-Puffer mit verschiedenen KCl-Konzentrationen 30 min inkubiert. Anschließend wurden die Ansätze auf 10%igen Polyacrylamidgelen aufgetrennt. Die Oligonucleotide wurden mit dem interkalierenden Farbstoff GelRed sichtbar gemacht.

Abbildung 27 zeigt zwei EMSAs mit U2AF65 und den ARPC2 G-Quadruplex und mut Oligonucleotiden bei 200 und 400 mM KCl. Interessanterweise kann man hier nach Zugabe des Proteins nicht nur eine retardierte Bande erkennen, sondern auch eine Bande, die schneller im Gel läuft als das freie Oligonucleotid. Dies ließ sich sowohl für das mutierte, als auch für das G-Quadruplex-Oligonucleotid beobachten. Bei beiden Sequenzen konnte man ab einer Proteinmenge von 40 pmol eine retardierte Bande erkennen, jedoch ist bei der höchsten Proteinmenge und der G-Quadruplex-Sequenz die freie RNA Bande verschwunden, während sie bei der mutierten Sequenz noch vorhanden ist. Das exprimierte Protein war also funktionell. Dieses Experiment deutete schon auf eine erhöhte Affinität von U2AF65 gegenüber der G-Quadruplex-Sequenz hin. Zur weiteren Charakterisierung dieser Bindungen wurden Oberflächenplasmonresonanzmessungen (SPR-Messungen) durchgeführt.

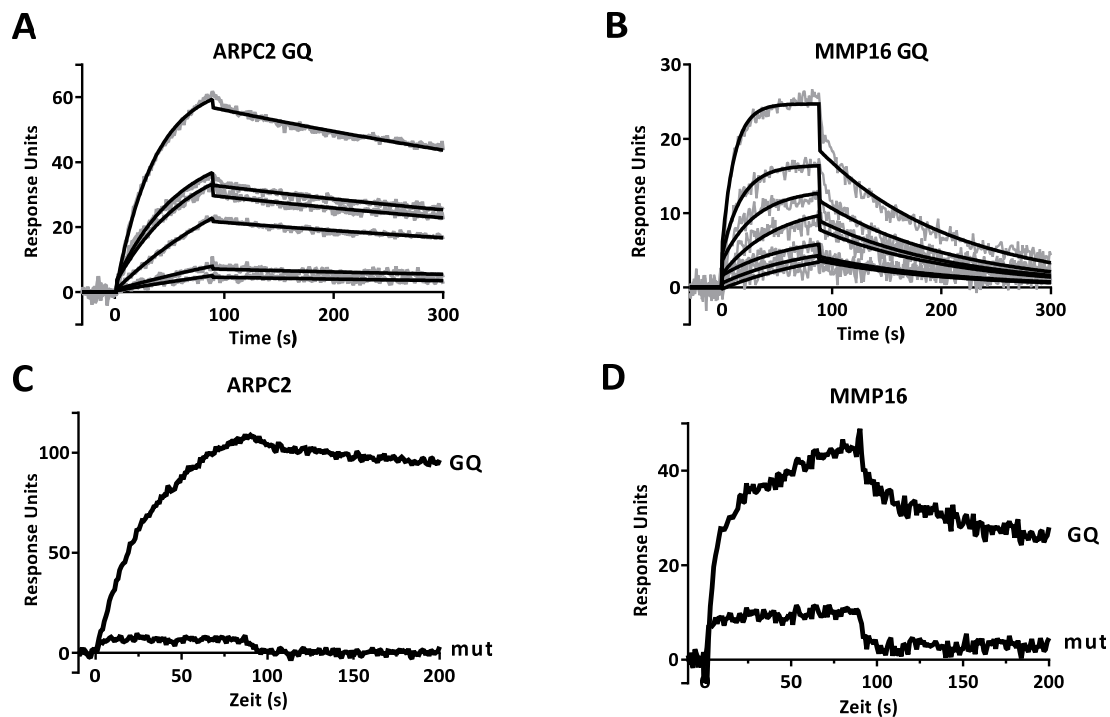


Abbildung 28: Oberflächenplasmonresonanzmessungen mit U2AF65 und den ARPC2 und MMP16 Oligonucleotiden. Die gemessenen SPR-Signale sind in grau dargestellt, die durch den Langmuir-Fit erhaltenen Kurven wurden in schwarz darübergelegt. (A) und (B) Bestimmung der kinetischen Parameter der Interaktion von U2AF65 mit den ARPC2 G-Quadruplex (1, 2, 5, 12,5, 15, 25 nM) und MMP16 G-Quadruplex Oligonucleotiden (1, 2,5, 5, 7,5, 15, 30, 50 nM). (C) und (D) Vergleich der Bindung an die G-Quadruplex- und die mutierten Sequenzen.

Die Ergebnisse der SPR-Messungen mit U2AF65 und den MMP16 und ARPC2 Oligonucleotiden sind in Abbildung 28 dargestellt. Die mutierten Sequenzen banden, im Vergleich zu den G-Quadruplex-Sequenzen, so wenig an U2AF65, dass mit dem verwendeten Versuchsaufbau keine Bindungskonstante bestimmt werden konnte (Abbildung 28 C, D). Man kann deutlich erkennen, dass die Affinität zu den beiden G-Quadruplex-Sequenzen viel größer war. Auch wenn man die Sensorgramme aufgrund der Verwendung neuer Sensorchips nicht direkt vergleichen kann, war auffällig, dass das Signal für das ARPC2 G-Quadruplex Oligonucleotid immer am größten war.

Tabelle 16: Bindungsparameter von U2AF65 und den ARPC2 und MMP16 G-Quadruplex Oligonucleotiden. Die Bindungskonstanten wurden aus den Messungen ermittelt, die in Abbildung 28 gezeigt sind. k_a : Assoziationsgeschwindigkeitskonstante; k_d : Dissoziationsgeschwindigkeitskonstante; K_D : Dissoziationskonstante, χ^2 : Maß für die Güte des Fits.

Ligand	Analyt	k_a ($M^{-1}s^{-1}$) $\times 10^6$	k_d (s^{-1}) $\times 10^{-3}$	K_D (nM)	χ^2
U2AF65	ARPC2 GQ	1.1	1.2	1	0.9
	ARPC2 mut			n. d.	
	MMP16 GQ	1.6	8.1	5	0.73
	MMP16 mut			n. d.	

Die Bindungskonstanten von U2AF65 und den beiden RNA G-Quadruplex Oligonucleotiden sind in Tabelle 16 aufgeführt. Die ermittelte Dissoziationskonstante für ARPC2 G-Quadruplex (1 nM) war etwa

fünffmal kleiner als für das MMP16 G-Quadruplex Oligonucleotid. Dieser Unterschied resultierte vor allem aus der langsameren Dissoziation von ARPC2 G-Quadruplex ($k_d = 1,2 \times 10^{-3} \text{s}^{-1}$) im Vergleich zu MMP16 G-Quadruplex ($k_d = 8,1 \times 10^{-3} \text{s}^{-1}$), während die Assoziationsgeschwindigkeitskonstante für beide in etwa gleich groß war. Die Bindung an beide G-Quadruplex-Sequenzen war jedoch sehr stark.

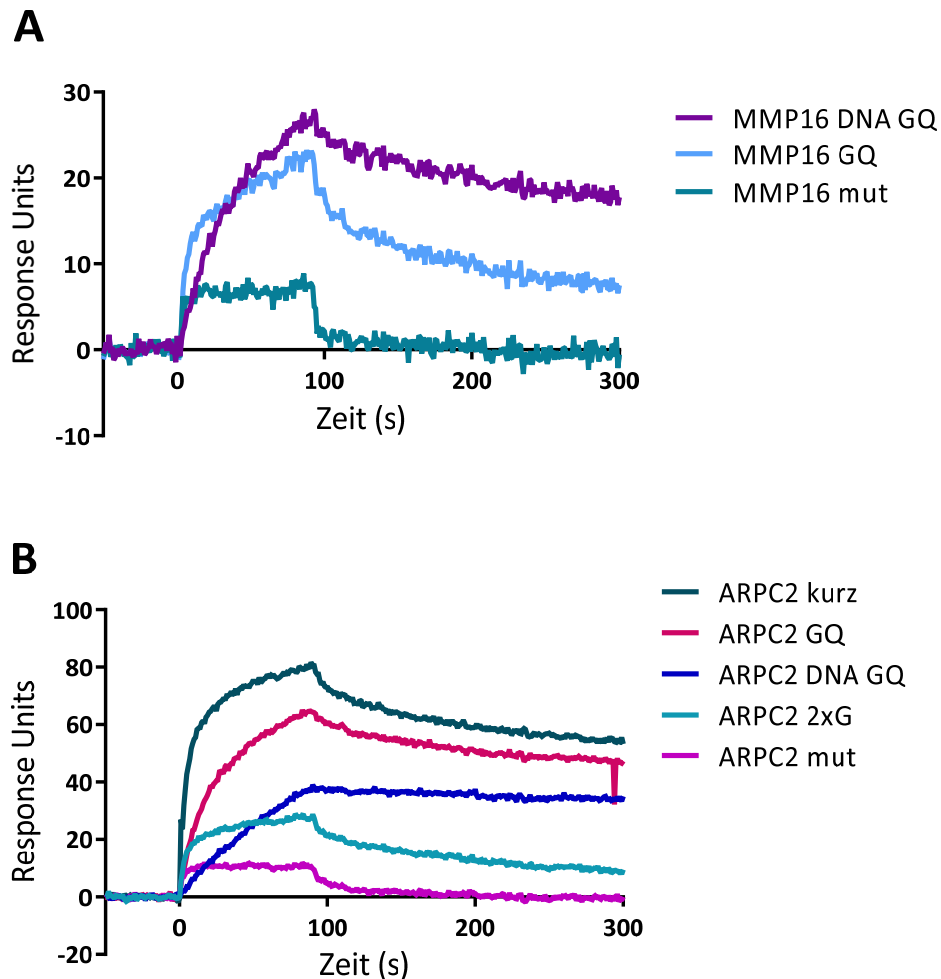


Abbildung 29: Oberflächenplasmonresonanzmessungen zum Vergleich der Bindung von verschiedenen Oligonucleotidsequenzen an U2AF65. (A) ARPC2-Oligonucleotide und (B) MMP16 Oligonucleotide wurden in einer Konzentration von 100 nM injiziert. Zur Regeneration der Oberfläche wurde 1 M KCl-Lösung benutzt.

Die Bindung verschiedener Kontrollsequenzen an U2AF65 wurde verglichen (Abbildung 29). Auch hier wurde die Bindung an die entsprechenden DNA G-Quadruplex-Sequenzen gemessen. Das MMP16 G-Quadruplex DNA Oligonucleotid zeigte eine ähnliche oder sogar stärkere Bindung an U2AF65 als an das RNA Oligonucleotid. Das ARPC2 G-Quadruplex DNA Oligonucleotid zeigte ein etwas weniger hohes Bindungssignal als das RNA Oligonucleotid. U2AF65 schien somit für das ARPC2 DNA G-Quadruplex eine etwas geringere Affinität zu haben. Um hier gesicherte Aussagen treffen zu können, müsste man jedoch die Bindungskonstante bestimmen. Das kürzere ARPC2 Oligonucleotid, dem die flankierenden Sequenzen fehlen, schien sogar etwas stärker an U2AF65 zu binden.

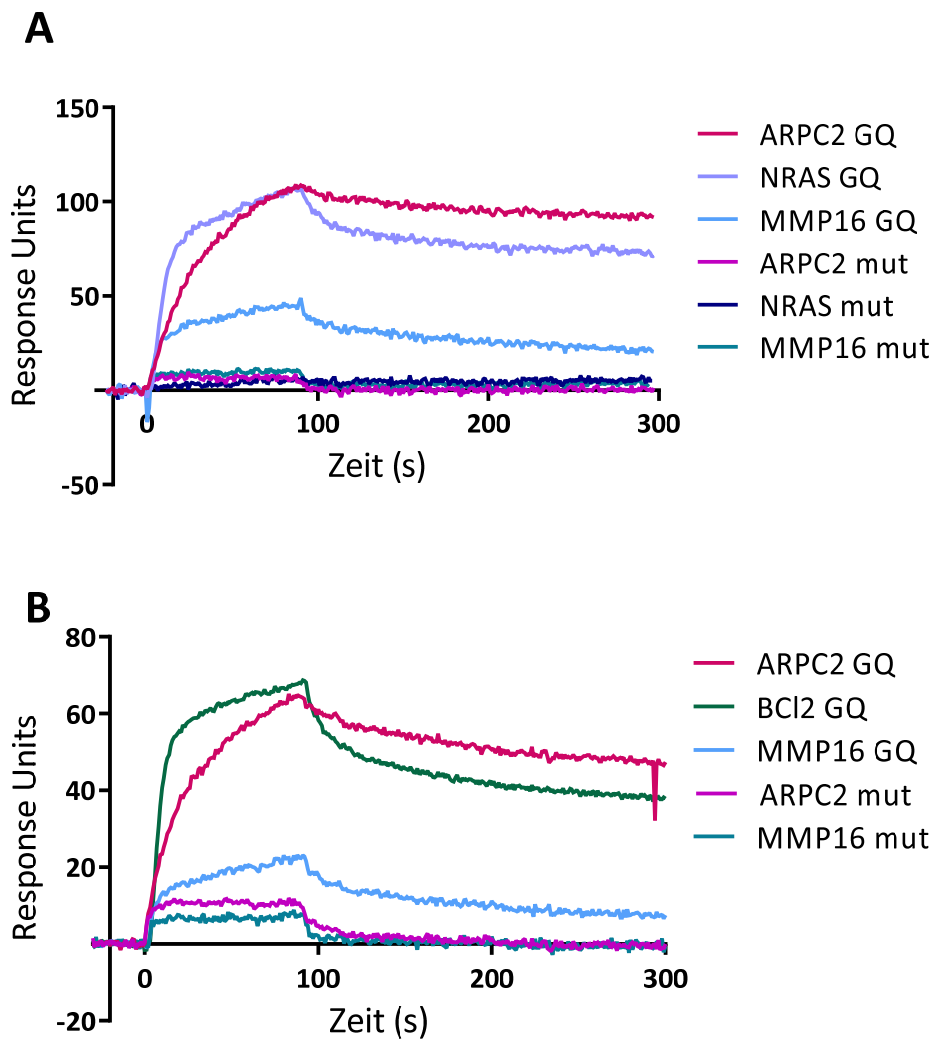


Abbildung 30: Oberflächenplasmonresonanzmessungen zum Vergleich der Bindung von U2AF65 an die G-Quadruplexe aus der 5'-UTR der NRAS und der Bcl-2 mRNA. (A) Vergleich der Bindung der NRAS G-Quadruplex- und mut Oligonucleotide mit der Bindung der ARPC2 und MMP16 Sequenzen an U2AF65 (B) Vergleich der Bindung der Bcl-2 G-Quadruplex- und mut Oligonucleotide mit der Bindung der ARPC2 und MMP16 Sequenzen an U2AF65.

Die Bindung anderer G-Quadruplexe an U2AF65 wurde ebenfalls durch SPR-Messungen untersucht. Die G-reichen Sequenzen aus der Bcl-2 und aus der NRAS 5'-UTR konnten auch sehr stark an U2AF65 binden. Beide Oligonucleotide zeigten ein sehr ähnliches Bindungssignal wie das ARPC2 G-Quadruplex und schienen sogar besser zu binden als das MMP16 G-Quadruplex.

Diese Ergebnisse deuten darauf hin, dass U2AF65 ein generelles G-Quadruplex bindendes Protein sein könnte.

4.6.3. EFHD2

Das Protein EFHD2 (*EF-hand domain family, member D2*) wurde in den Pull-Down-Assays dreimal als ARPC2 G-Quadruplex Binder identifiziert, zweimal in den Experimenten mit Hek293 Lysat und einmal mit HeLa Lysat. Das Gen des 27 kDa großen Proteins wurde zur Expression ebenfalls in pET28a kloniert und über einen N-terminalen Polyhistidin-Tag aufgereinigt. Im Gegensatz zu den beiden vorher beschriebenen Proteinen, Nucleolin und U2AF65, war eine nucleinsäurebindende Aktivität von EFHD2 bisher noch nicht bekannt.

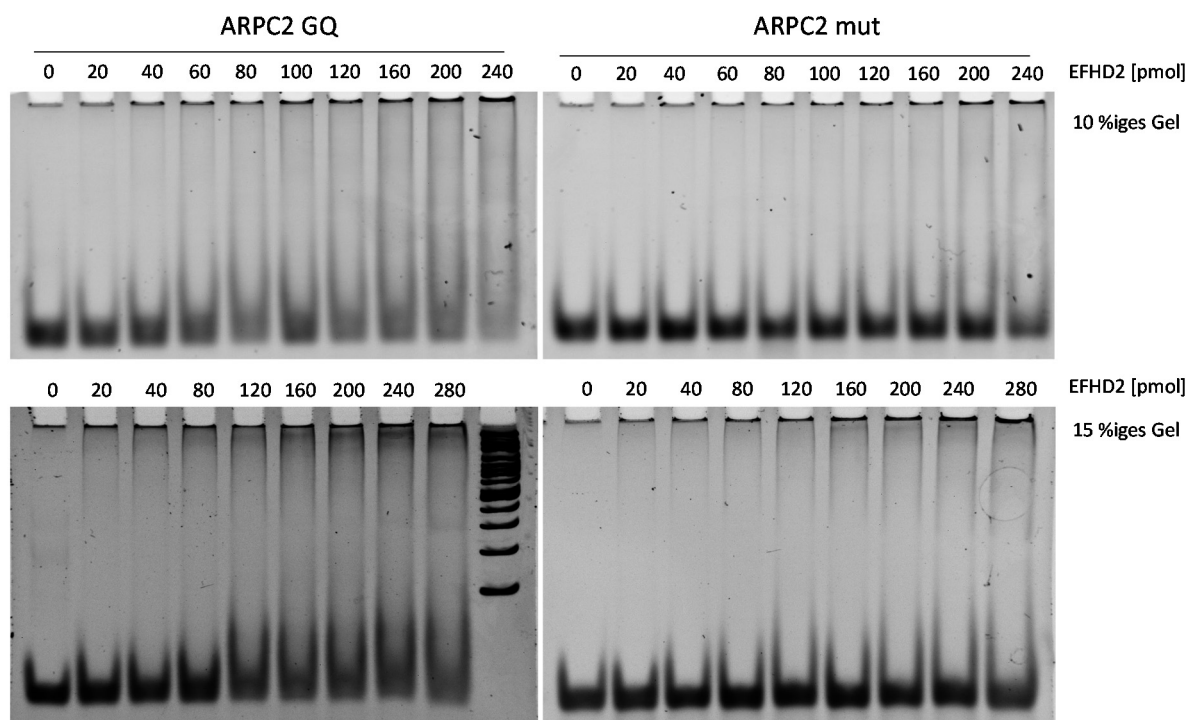


Abbildung 31: Electrophoretic Mobility Shift Assay mit EFHD2 und den Oligonucleotiden ARPC2 G-Quadruplex und mut. 20 pmol Oligonucleotid wurden mit steigenden Konzentrationen an EFHD2 bei einer KCl-Konzentration von 200 mM inkubiert. Anschließend wurden die Ansätze auf 10%igen Polyacrylamidgelen aufgetrennt. Die Oligonucleotide wurden mit dem interkalierenden Farbstoff GelRed sichtbar gemacht.

Die Electrophoretic Mobility Shift Assays mit EFHD2 und den ARPC2 G-Quadruplex und mut Oligonucleotiden sind in Abbildung 31 gezeigt. In diesen Assays konnte man gut erkennen, dass EFHD2 eine höhere Affinität für das ARPC2 G-Quadruplex Oligonucleotid hatte, als für die mutierte Kontrolle. Für ARPC2 G-Quadruplex ist spätestens ab 120 pmol Protein ein Shift zu erkennen, allerdings war keine retardierte Bande sichtbar. Man konnte nur das Verschwinden der freien Oligonucleotidbande beobachten. Um eine bessere Auflösung des Assays zu erreichen, wurde statt eines 10%igen Polyacrylamidgels ein 15%iges Gel verwendet. Auch hier konnte man ab etwa 120 pmol EFHD2 ein Verschwinden der ARPC2 G-Quadruplex Bande erkennen. Eine distinkte retardierte Oligonucleotidbande war jedoch auch auf dem höherprozentigen Gel nicht sichtbar. Das Laufverhalten des mutierten ARPC2 Oligonucleotids änderte sich nicht viel durch die Inkubation mit EFHD2. Nur bei

der höchsten Proteinkonzentration ist die Bande etwas schwächer geworden, dies war deutlicher auf dem 10%igen Polyacrylamidgel zu erkennen.

Auch die Bindung von EFHD2 an die RNA Oligonucleotide wurde mittels SPR-Messungen untersucht.

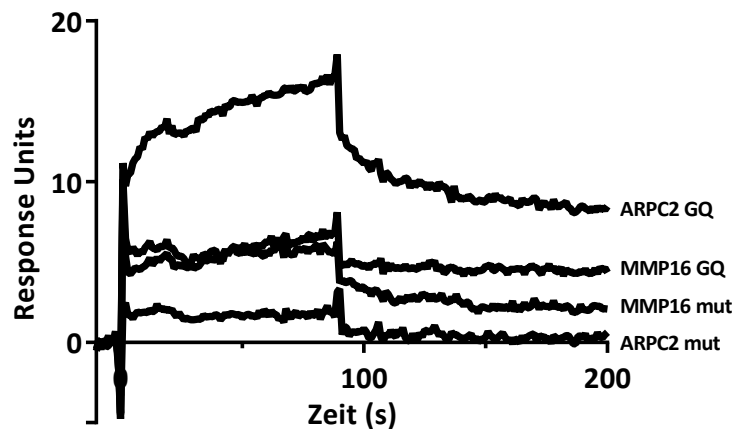


Abbildung 32: Oberflächenplasmonresonanzmessungen zur Untersuchung der Bindung der MMP16 und ARPC2 RNA Oligonucleotide an EFHD2. Die Oligonucleotide wurden als Analyten in einer Konzentration von 100 nM injiziert. Sie wurden mit steigender Affinität injiziert, da das immobilisierte Protein sehr schnell an Aktivität verlor.

Abbildung 32 zeigt die Ergebnisse der Oberflächenplasmonresonanzmessungen mit EFHD2 und den MMP16 und ARPC2 RNA Oligonucleotiden. Die Bestimmung von Bindungskonstanten war hier nicht möglich, da nach der Immobilisierung von EFHD2 an den Sensorchip die Aktivität des Proteins nach jeder Messung sichtbar abnahm. So wurden bei der mehrfachen Injektion der gleichen Konzentration eines Oligonucleotids jedes Mal geringere Response Units gemessen. Um die Bindungsparameter zu bestimmen, ist jedoch ein stabiles Signal über mehrere Injektionen notwendig, daher ist hier nur ein Vergleich der Bindung der vier RNA Oligonucleotide an EFHD2 gezeigt. Um das Ergebnis nicht zu verfälschen, wurden die am schlechtesten bindenden Oligonucleotide zuerst injiziert. Des Weiteren konnte man dadurch auf die Regeneration des Chips durch die Injektion von 1 M KCl verzichten, was die Aktivität des Proteins zusätzlich beeinträchtigt hätte.

In den SPR-Messungen konnten die vorherigen Ergebnisse bestätigt werden. Obwohl die ARPC2 G-Quadruplex-Sequenz als letztes injiziert wurde, zeigte sie das größte Bindungssignal. Das Signal war mit maximal 16 RU trotz der recht großen injizierten Konzentration von 100 nM im Vergleich zu den Messungen mit Nucleolin und U2AF65 eher klein. Das mutierte ARPC2 Oligonucleotid zeigte nur eine sehr geringe Bindung, während die beiden MMP16 Oligonucleotide etwas besser zu binden schienen.

Zusammengefasst konnten die Bindungsstudien mit ARPC2 die Ergebnisse aus dem Pull-down-Assay bestätigen. Sowohl im EMSA als auch in den SPR-Messungen wurde gezeigt, dass EFHD2 besser an die ARPC2 G-Quadruplex-Sequenz bindet als an die mutierte ARPC2 Kontrolle. Des Weiteren kann man aus

den SPR-Messungen schließen, dass EFHD2 nicht allgemein an G-Quadruplexe zu binden scheint, da die Bindung an die MMP16 G-Quadruplex-Sequenz sehr schwach war.

4.6.4. SRSF1

Der Spleißfaktor *Serine/arginine-rich splicing factor 1* (SRSF1) wurde dreimal mit Hek293 Zelllysate und einmal mit HeLa Zelllysate als Binder von ARPC2 G-Quadruplex identifiziert. Für die MMP16 Oligonucleotide wurde SRSF1 nicht gefunden. Da sich SRSF1 nur schwierig heterolog in *E. coli* exprimieren ließ, wurden mit diesem Protein keine Electrophoretic mobility shift Assays durchgeführt, für die man recht große Mengen an Protein benötigt. Bei der Expression in *E. coli* lag der größte Teil des Proteins in unlöslichen Aggregaten vor, war also nach der Zellyse und Zentrifugation im Pellet zu finden. Mit SRSF1 wurden direkt Oberflächenplasmonresonanz-Analysen durchgeführt; die Menge an Protein, die hier benötigt wird, ist nur sehr gering.

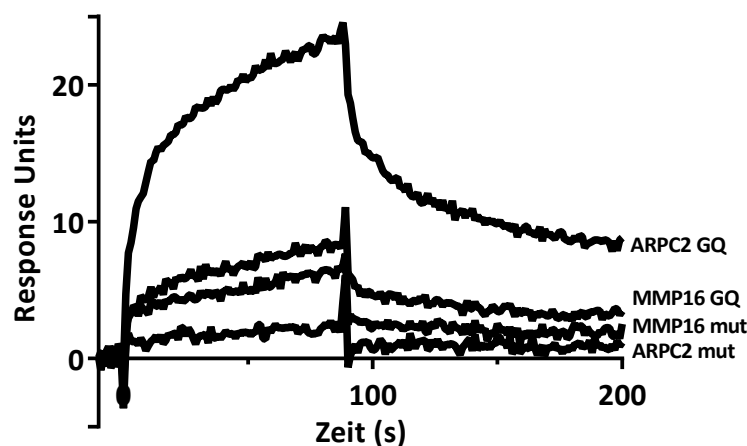


Abbildung 33: Oberflächenplasmonresonanzmessungen zum Vergleich der Bindung der MMP16 und ARPC2 RNA Oligonucleotide an SRSF1. SRSF1 wurde als Ligand auf der Chipoberfläche immobilisiert, und die Oligonucleotide wurden als Analyten in einer Konzentration von 100 nM injiziert.

Für die SPR-Messungen mit SRSF1 ergab sich das gleiche Problem wie für die Messungen mit EFHD2, auch hier nahm die Aktivität des Proteins nach wenigen Injektionen sichtbar ab. Aus diesem Grund wird für dieses Protein ebenfalls nur ein Vergleich der Bindungssignale gezeigt (Abbildung 33). Die am wenigsten bindenden Oligonucleotide wurden wieder zuerst injiziert. Das ARPC2 mut RNA Oligonucleotid schien am wenigsten mit SRSF1 zu interagieren. Die beiden MMP16 Oligonucleotide zeigten ein vergleichbares Bindungssignal, was für eine geringe Interaktion spricht. Mit dem ARPC2 G-Quadruplex Oligonucleotid konnte das größte Bindungssignal beobachtet werden. Der steile Kurvenverlauf deutet darauf hin, dass hier mit der Injektion von 100 nM noch keine Sättigung der Oberfläche bzw. noch kein Gleichgewichtszustand erreicht wurde. Die mehrfache Identifizierung von SRSF1 als ARPC2 G-Quadruplex bindendes Protein in den Pull-down-Assays und die Ergebnisse der SPR-Messungen deuten auf eine hohe Affinität des ARPC2 G-Quadruplex zu SRSF1 hin.

4.7. FUNKTIONELLE STUDIEN

Nachdem es gelungen war, Proteine zu identifizieren, die spezifisch an ein G-Quadruplex binden oder, wie U2AF65, sogar mehrere G-Quadruplexe erkennen konnten, stellte sich die Frage, welche Funktion diese Proteine in Bezug auf die G-Quadruplex-Strukturen erfüllen könnten. Da die beiden G-Quadruplexe aus den 5'-UTRs von MMP16 und ARPC2 aufgrund ihres hemmenden Einflusses auf die Translation ausgewählt wurden, sollte zunächst der Einfluss der bindenden Proteine auf die Translation untersucht werden.

4.7.1. LUCIFERASE CO-TRANSFEKTIONEN

Um den Einfluss der identifizierten G-Quadruplex bindenden Proteine auf die Translation zu untersuchen, wurden Co-Transfektionen mit Dual-Luciferase-Reporterplasmiden durchgeführt. Hierzu wurden die psiCHECK-2 Plasmide verwendet, in denen die isolierte ARPC2 G-Quadruplex-Sequenz, sowie die entsprechende mutierte Sequenz vor die Renilla-Luciferase kloniert wurden (Abbildung 18). Die Proteine wurden zur Co-Transfektion in den eukaryotischen Expressionsvektor pcDNA3.1 kloniert. Die proteincodierenden Plasmide und die Reporterplasmide wurden in Hek293 Zellen Co-Transfiziert. Von den proteincodierenden Plasmiden wurde in etwa die doppelte Menge eingesetzt wie von den Reporterplasmiden. Dadurch sollte erreicht werden, dass das jeweilige Protein in ausreichender Menge gebildet wurde. 24 h nach der Transfektion wurden die Zellen geerntet und die Luciferase-Aktivitäten wurden bestimmt. Die Renilla-Luciferase-Aktivität jedes Ansatzes wurde auf die Aktivität der Firefly-Luciferase normiert. Die Luciferase-Aktivitäten der Co-Transfektionen mit dem psiCHECK-2-ARPC2-G-Quadruplex Plasmid wurden auf die Ergebnisse der Co-Transfektionen mit dem psiCHECK-2-ARPC2-mut Plasmid normiert. Als Kontrolle dienten Ansätze, in denen kein zusätzliches Protein transfiziert wurde, sowie die Co-Transfektion eines pcDNA3.1 Plasmids, das das *green fluorescent protein* (GFP) exprimierte. Durch die Transfektion des GFP codierenden Plasmids konnte zusätzlich auch die Transfektionseffizienz der proteincodierenden Plasmide im Fluoreszenzmikroskop überprüft werden.

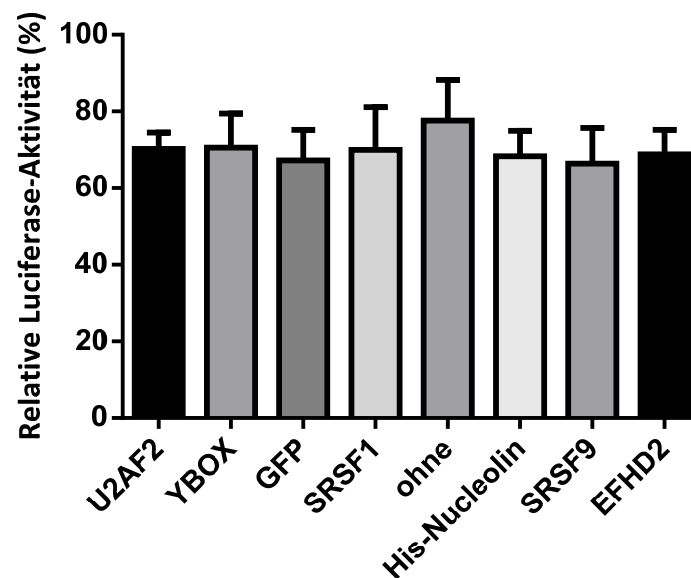


Abbildung 34: Co-Transfektion der Luciferase-Reporterplasmide und der identifizierten G-Quadruplex bindenden Proteine. Die Renilla-Luciferase-Aktivität wurde auf die Aktivität der Firefly-Luciferase normiert. Die Ergebnisse wurden auf die Aktivitäten der Co-Transfektion mit dem Plasmid psiCHECK-2-ARPC2mut bezogen. Das Experiment wurde dreimal in Duplikaten durchgeführt. Die Luciferase-Aktivitäten wurden mindestens zweimal bestimmt.

Die Ergebnisse des Luciferase Co-Transfektions-Experiments sind in Abbildung 34 gezeigt. In diesem Experiment konnte kein Einfluss der identifizierten G-Quadruplex bindenden Proteine auf die Translation nachgewiesen werden. Durch Co-Transfektion jedes der Proteine, auch der GFP Kontrolle, blieb die Luciferase-Aktivität in etwa gleich bei ca. 70 %. Die Kontrolle, in der kein zusätzliches Protein transfiziert wurde zeigte auch eine ähnliche Aktivität.

4.7.2. *IN VITRO* TRANSLATION

Mit der Co-Transfektion der als G-Quadruplex-Binder identifizierten Proteine und der G-Quadruplex Reporterplasmide konnte kein Effekt auf die Translation gezeigt werden. Bei den Co-Transfektionen besteht die Möglichkeit, dass die Proteine nicht richtig gebildet werden, oder dass Protein und Reporterkonstrukt in der Zelle nicht in Kontakt kommen. Daher wurde im nächsten Schritt ein einfacheres Modellsystem verwendet. Die *in vitro* Translation bietet auch den Vorteil, dass man den Translationsprozess getrennt von anderen zellulären Prozessen betrachten kann. Für die *in vitro* Translation wurden die Renilla-Luciferase Reporterkonstrukte zunächst durch *in vitro* Transkription in RNA umgeschrieben. Den Reporter RNAs wurde die Firefly Kontroll-RNA aus dem *in vitro* Translationssystem (*Rabbit Reticulocyte Lysate System*, Promega) als Standard zur Normierung der Ergebnisse beigemischt. Die Reporter RNAs wurden mit den Proteinen vorinkubiert, um eine Bindung auszubilden. Anschließend wurde die *in vitro* Translation mit Retikulozytenlysat durchgeführt und die Aktivitäten der beiden Luciferasen wurden bestimmt. Als Kontrolle wurde ein Translationsansatz mit

BSA statt mit einem der G-Quadruplex bindenden Proteine durchgeführt. Die Renilla-Luciferase-Aktivität wurde zunächst auf die Firefly-Luciferase-Aktivität normiert. Die Aktivitäten der ARPC2 G-Quadruplex und mut Ansätze wurden dann auf den jeweiligen Ansatz mit bovinem Serumalbumin (BSA) normiert.

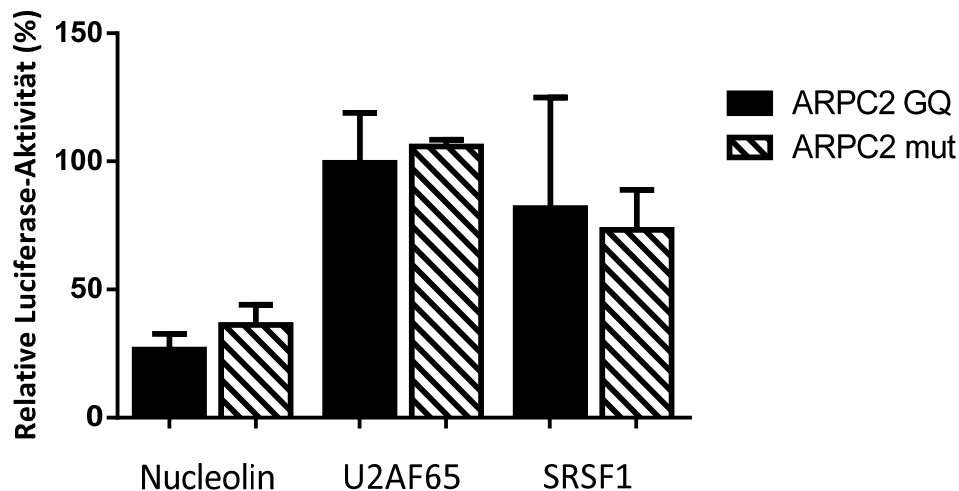


Abbildung 35: *In vitro* Translationsassay zur Bestimmung des Einflusses der G-Quadruplex bindenden Proteine auf die Translation. Zur Normierung dienten Translationsansätze der ARPC2 G-Quadruplex oder mut Luciferase-Konstrukte, zu denen BSA als Kontrolle, statt eines anderen Proteins gegeben wurde. Des Weiteren wurde jedem Ansatz Firefly-Luciferase-RNA beigemischt und die Renilla-Luciferase-Aktivität wurde auf die Firefly-Aktivität normiert. Das Experiment wurde dreimal in Duplikaten durchgeführt. Die Luciferase-Aktivitäten wurden mindestens zweimal bestimmt.

Abbildung 35 zeigt die Ergebnisse des *in vitro* Translationsassays. Durch die Normierung auf die BSA Ansätze sollte der Einfluss, den die Zugabe eines weiteren Proteins zu den Translationsansätzen hat, aus den Ergebnissen eliminiert werden. Die Ergebnisse zeigen, dass die G-Quadruplex bindenden Proteine zwar einen unterschiedlichen Einfluss auf die Translation haben, jedoch nicht speziell auf die Translation der mRNA mit der G-Quadruplex-Sequenz in der 5'-UTR, sondern genauso auf die mRNAs mit der Kontrollsequenz in der 5'-UTR. So hat Nucleolin die Translation beider Konstrukte um mehr als 60 % reduziert, U2AF65 hatte keinen Einfluss auf die Stärke der Translation und SRSF1 senkte die Translation um etwa 20 %.

Zusammengefasst ließ sich sowohl durch *in vivo* als auch durch *in vitro* Translationsversuche kein Einfluss der in dieser Arbeit identifizierten G-Quadruplex bindenden Proteine auf die Translation feststellen.

4.8. BESTÄTIGUNG DES PULL-DOWN EXPERIMENTS MITTELS SHOTGUN PROTEOMICS

Die Ergebnisse aus den Pull-down-Assays wurden zusätzlich über einen weiteren Ansatz überprüft. Hierfür wurde eine SILAC (*stable isotope labeling by amino acids in cell culture*) Markierung der Proteine mit einer Analyse mittels Ionenfallen-Massenspektrometer (Orbitrap) verknüpft. Der Vorteil

hierbei ist die Möglichkeit des quantitativen Vergleichs der Proteine, die an G-Quadruplex und mutierte Sequenz gebunden haben. Für dieses Experiment wurde eine Population Hek293 Zellen mit schweren Isotopen der Aminosäuren L-Lysin und L-Arginin (L-Lysin:2HCl, 13C6, 15N2, L-Arginin:HCl, 13C6) markiert, während eine zweite Population mit unmarkierten Aminosäuren kultiviert wurde. Aus den beiden Populationen wurde Proteinextrakt gewonnen. Der mit den schweren Aminosäuren markierte Extrakt wurde im Pull-down-Assay mit den G-Quadruplex Oligonucleotiden inkubiert, die mutierten Oligonucleotide wurden mit dem unmarkierten Extrakt inkubiert. In diesem Pull-down-Assay wurde nicht schrittweise mit steigenden KCl-Konzentrationen eluiert, es wurde nur eine Elutionsfraktion gesammelt. Die *Beads* wurden mit Waschpuffern mit 400 mM und 600 mM KCl gewaschen, um unspezifisch bindende Proteine zu entfernen, anschließend wurde mit 2600 mM KCl eluiert. Die Elutionsfraktion wurde durch Ultrafiltration aufkonzentriert und die hohe Salzkonzentration durch Waschen mit TRIS-Puffer ohne KCl reduziert. Die entsprechenden G-Quadruplex und mut Eluate wurden vereinigt und mittels Orbitrap analysiert³. Die Analyse mittels Orbitrap und vorgeschalteter HPLC ist wesentlich sensitiver als die gelbasierte Methode, die vorher verwendet wurde, so dass mit diesem Ansatz für ARPC2 172 Proteine identifiziert wurden, die mindestens doppelt so häufig an die G-Quadruplex-Sequenz als an die mutierte Sequenz gebunden haben. Für MMP16 wurden 90 Proteine gefunden, die dieses Kriterium erfüllten. Dieses Experiment wurde nur einmal durchgeführt. Es kann daher nicht zur quantitativen Auswertung verwendet werden und soll hier nur zur Bestätigung der vorherigen Ergebnisse herangezogen werden. In Tabelle 17 sind für jedes G-Quadruplex die 20 Proteine mit dem größten Verhältnis von schweren zu leichten Aminosäuren aufgeführt, d.h. die Proteine, die am häufigsten an die G-Quadruplex-Sequenz im Vergleich zur mutierten Sequenz gebunden haben. Die Liste der mit dieser Methode identifizierten Proteine, die mindestens doppelt so häufig an die G-Quadruplex-Sequenz als an die mutierte Sequenz gebunden haben, befindet sich im Anhang. Die Proteine, die auch im gelbasierten Pull-down gefunden wurden, sind fett gedruckt.

Mit diesem Ansatz wurden, wie schon in den Gel-basierten Pull-down-Assays, viel weniger Proteine mit einer Affinität für die MMP16 G-Quadruplex-Sequenz gefunden als für das ARPC2 G-Quadruplex. Nucleolin wurde für ARPC2 in jedem Pull-down Experiment (3-mal mit Hek-Lysat, 2-mal mit HeLa Lysat) als Binder identifiziert und band auch in diesem Experiment ca. 15-mal mehr an die Kontrolle als an die mutierte Sequenz. Das mitochondriale Protein ME2, das am häufigsten für MMP16 gefunden wurde, zeigte auch hier die größte Affinität für die MMP16 G-Quadruplex-Sequenz (Verhältnis H/L 17,4).

³ Die Orbitrap-Analyse fand im Labor von Prof. Juri Rappsilber an der University of Edinburgh, Schottland, statt.

Das 60S ribosomale Protein L35 (RPL35) hat 43-mal mehr an die ARPC2 G-Quadruplex-Sequenz als an die mutierte ARPC2 Sequenz gebunden. In den gelbasierten Pull-down-Assays wurde dieses Protein überhaupt nicht gefunden, wie auch die meisten der anderen ribosomalen Proteine, die hier, vor allem für ARPC2, als Binder identifiziert wurden. Der Grund hierfür liegt wahrscheinlich in der Größe der Proteine. RPL35 ist nur 14,5 kDa groß, die meisten der ribosomalen Proteine sind etwa zwischen 10 und 30 kDa groß. Dies macht die Identifizierung einzelner Banden auf der SDS-PAGE schwierig, da alle Proteine im selben Bereich des Gels laufen. Zusätzlich ist die Auflösung der verwendeten 10%igen Gele in diesem Bereich nicht optimal.

Generell konnten die Ergebnisse aus den gelbasierten Pull-down-Assays bestätigt werden. Alle Proteine, die hier identifiziert wurden, wurden auch im SILAC Pull-down-Assay gefunden.

Tabelle 17: Liste der 20 Proteine, die im SILAC Pull-down-Assay am meisten an die ARPC2 G-Quadruplex und MMP16 G-Quadruplex-Sequenzen im Vergleich zu den mutierten Sequenzen gebunden haben. Der Wert Heavy/Light gibt das Verhältnis an isotoopenmarkierten zu unmarkierten Proteinen in der Probe an. Die G-Quadruplex-Sequenzen wurden mit den isotoopenmarkierten Proteinen inkubiert. Die Proteine, die auch in den gelbasierten Pull-down-Assays identifiziert wurden, sind fett geschrieben.

	ARPC2	Heavy/Light	MMP16	Heavy/Light
1	<i>60S ribosomal protein L35</i>	42,64	<i>NAD-dependent malic enzyme, mitochondrial</i>	17,41
2	<i>60S ribosomal protein L17</i>	38,02	<i>RRP15-like protein</i>	17,36
3	<i>60S ribosomal protein L32</i>	26,79	<i>Putative methyltransferase NSUN5</i>	16,61
4	<i>60S ribosomal protein L37</i>	23,50	<i>Suppressor of SWI4 1 homolog</i>	15,55
5	<i>60S ribosomal protein L13</i>	17,55	<i>60S ribosomal protein L26</i>	12,90
6	<i>60S ribosomal protein L26</i>	17,10	<i>60S ribosomal protein L23a</i>	10,54
7	<i>40S ribosomal protein S6</i>	16,80	<i>Zinc finger protein 579</i>	9,77
8	<i>60S ribosomal protein L23a</i>	16,74	<i>Enhancer of rudimentary homolog</i>	9,44
9	<i>40S ribosomal protein S24</i>	16,29	<i>40S ribosomal protein S24</i>	8,90
10	<i>Nucleolar RNA helicase 2</i>	15,26	<i>7SK snRNA methylphosphate capping enzyme</i>	8,69
11	<i>RNA-binding protein 4</i>	14,90	<i>Bcl-2-associated transcription factor 1</i>	8,51
12	<i>Nucleolin</i>	14,88	<i>40S ribosomal protein S2</i>	8,45
13	<i>40S ribosomal protein S2</i>	14,76	<i>Glioma tumor suppressor candidate region gene 2 protein</i>	8,43
14	<i>40S ribosomal protein S9</i>	13,55	<i>ATP-dependent RNA helicase DHX8</i>	8,40
15	<i>60S ribosomal protein L27a</i>	13,42	<i>Methyl-CpG-binding domain protein 1</i>	8,22
16	<i>Peptidyl-prolyl cis-trans isomerase G</i>	13,00	<i>Heterogeneous nuclear ribonucleoprotein F</i>	7,80
17	<i>Heterogeneous nuclear ribonucleoprotein F</i>	12,94	<i>R3H and coiled-coil domain-containing protein 1</i>	7,73
18	<i>Ribosomal L1 domain-containing protein 1</i>	12,90	<i>ATP-dependent RNA helicase DDX50</i>	7,51
19	<i>PUF60_HUMAN Isoform 6 of Poly(U)-binding-splicing factor PUF60</i>	12,15	<i>Elongation factor 1-alpha 1</i>	7,44
20	<i>Serine/threonine-protein kinase PRP4 homolog</i>	11,80	<i>Thyroid hormone receptor-associated protein 3</i>	7,44

Die Liste der durch diesen Versuch erhaltenen, G-Quadruplex bindenden Proteine ist, wie oben bereits erwähnt, sehr umfangreich. Um zu versuchen über die bindenden Proteine, die Prozesse zu identifizieren, an denen die G-Quadruplexe beteiligt sind, wurde eine *Gene Ontology* (GO) Analyse durchgeführt. Dies ist eine Datenbank, in der jedem Gen Attribute, sogenannte GO Terms aus den Bereichen *cellular component*, *molecular function* und *biological process* zugeordnet werden. Hierdurch werden bioinformatische Clusteranalysen größerer biologischer Datensätze möglich. Die GO Term Analyse der G-Quadruplex bindenden Proteine ist in Abbildung 36 dargestellt. Sie wurde mit dem Webtool AmiGO 2 (96) durchgeführt. Die Analyse einer GO Term Anreicherung im Bereich „zelluläre Komponente“ ergab die Zuordnung der MMP16 und ARPC2 bindenden Proteine zum Ribosom, zu Ribonucleoproteinkomplexen und als Teile von Organellen. Diese Lokalisierungen sind bei RNA bindenden Proteinen zu erwarten. Ein Unterschied ist, dass für MMP16 eine Anreicherung nukleärer Proteine gefunden wurde. Auch in der Kategorie „molecular function“ wurden keine unerwarteten Gene Ontology Term Anreicherungen gefunden. Für beide Sequenzen wurden Anreicherungen von GO Terms wie „*RNA binding*“, „*nucleic acid binding*“ oder „*structural constituent of the ribosome*“ gefunden. Interessanterweise ließen sich in Bezug auf den biologischen Prozess sowohl für ARPC2 als auch für MMP16 einige Proteine dem Term „*viral gene expression*“ zuordnen.

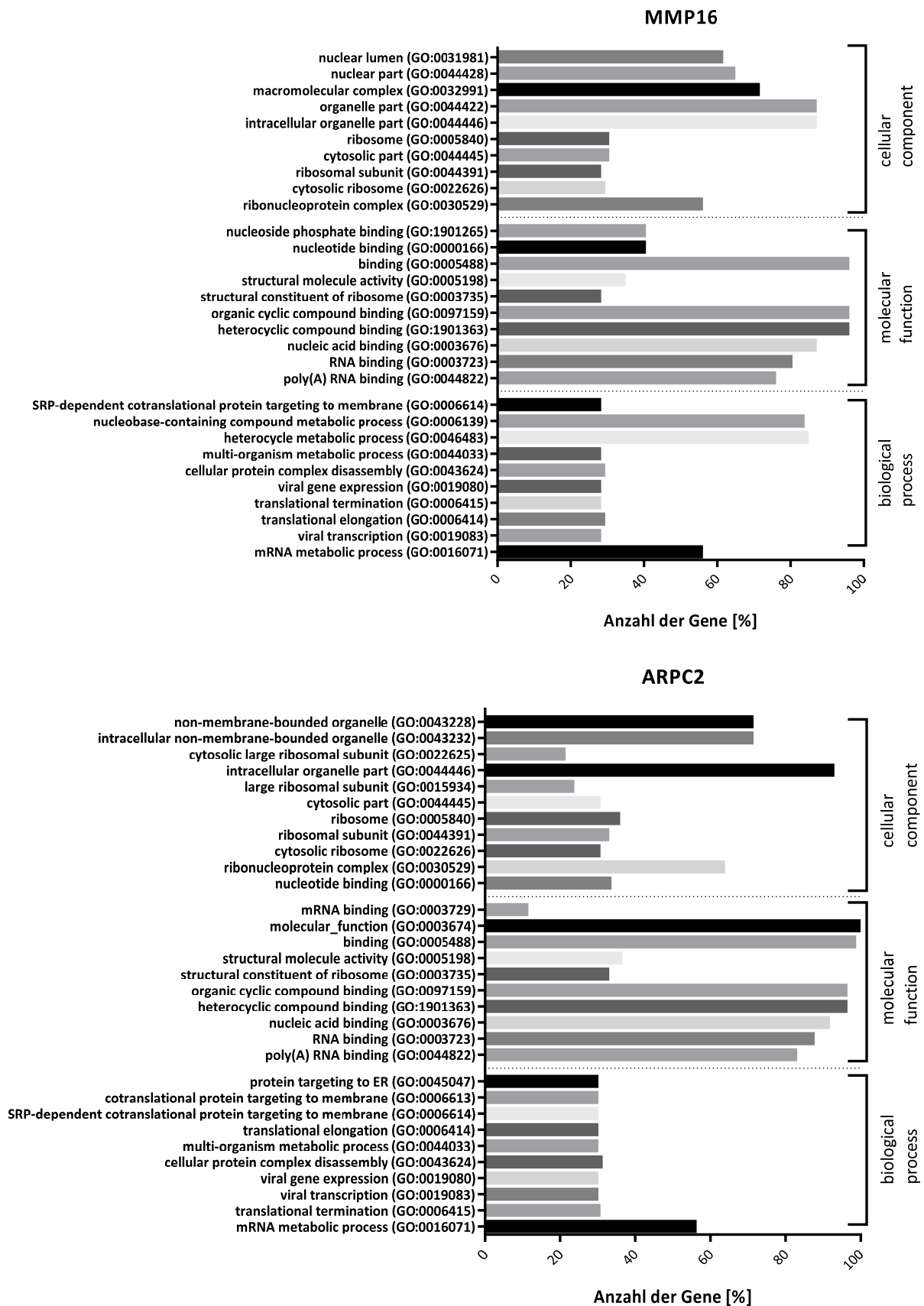


Abbildung 36: Gene Ontology (GO) Term Anreicherungen der MMP16 und ARPC2 G-Quadruplex bindenden Proteine. Die GO Term Analyse wurde mit dem Webtool AmiGO 2 durchgeführt (96). Auf der X-Achse ist die Anzahl der Gene in Prozent aufgeführt, die einer Kategorie zugeordnet werden konnten. Die gesamte Anzahl der eingegebenen Gene betrug für MMP16 90 und für ARPC2 172.

4.9. ARP2/3-KOMPLEX G-QUADRUPLEXE

4.9.1. *ARPC2-UTR VARIANTEN*

Für das Gen ARPC2 werden in der NCBI *Nucleotide* Datenbank (<http://www.ncbi.nlm.nih.gov/nucleotide>) zwei Spleißvarianten vorhergesagt, die sich lediglich in ihren 5'-UTR Sequenzen unterscheiden (Abbildung 7). Die eine UTR Variante enthält das G-Quadruplex, welches in dieser Arbeit untersucht wurde (NCBI Referenz Sequenz: NM_152862.2). Sie ist 169 Basen lang. Die andere UTR ist mit 70 Basen kürzer und enthält kein G-Quadruplex (NM_005731.3). Dies deutet auf eine regulatorische Rolle der beiden 5'-UTR Sequenzen hin. In den weiteren Versuchen sollte die Expression und die Rolle dieser beiden Transkriptvarianten untersucht werden.

Im ersten Schritt sollte die Expression der beiden in der Datenbank annotierten UTR Varianten in Hek293 Zellen nachgewiesen werden. Hierzu wurde Hek293 RNA durch Reverse Transkription (RT) in DNA umgeschrieben. Anschließend wurde eine PCR mit spezifischen Primern für die jeweilige UTR Sequenz durchgeführt (Abbildung 37).

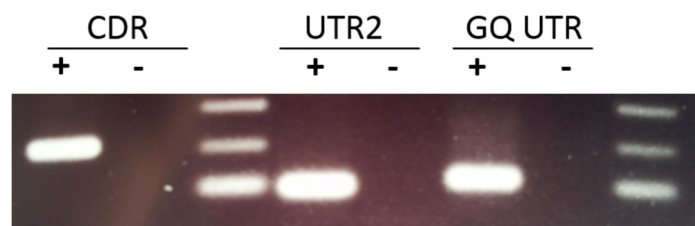


Abbildung 37: Nachweis der Expression der beiden ARPC2 UTR Varianten (UTR2 und G-Quadruplex UTR) sowie der codierenden Sequenz (CDR) mittels RT-PCR. Die Markierungen + und – beziehen sich auf die Reverse Transkription. Ansätze, mit denen vor der PCR keine Reverse Transkription durchgeführt wurden, dienten als Kontrolle für Kontaminationen mit genomischer DNA.

Die Expression beider UTR Varianten konnte in den RT PCRs mit Hek293 RNA nachgewiesen werden. Beide PCR Produkte wurden sequenziert, und die vorhergesagte Sequenz konnte bestätigt werden.

Der Einfluss der kürzeren UTR2 auf die Translation sollte, wie schon für die G-Quadruplex UTR, durch Luciferase-Reporter-Assays untersucht werden. Dazu wurde die Sequenz vor die Renilla-Luciferase in den psiCHECK-2 Vektor kloniert. Das Reporterplasmid wurde in Hek293 Zellen transfiziert, und die Luciferase-Aktivitäten wurden bestimmt (Abbildung 38).

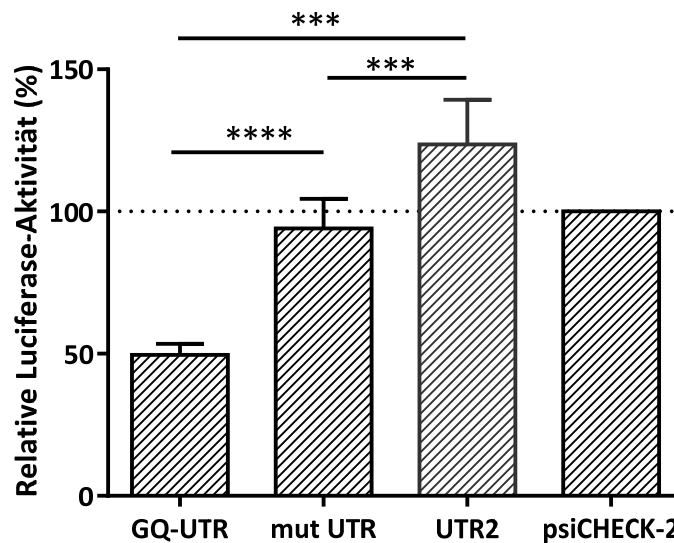


Abbildung 38: Dual Luciferase-Assay zur Bestimmung des Einflusses der zwei ARPC2 5'-UTR Spleißvarianten auf die Translation. Das Experiment wurde zweimal mit jeweils drei biologischen Replikaten durchgeführt. Jede Luciferase-Aktivität wurde dreimal bestimmt. Die Aktivität der Renilla-Luciferase wurde auf die der Firefly-Luciferase normiert. Dargestellt sind die Mittelwerte mit den Standardabweichungen. Die Signifikanzanalyse von Messergebnissen erfolgte mittels einfaktorieller ANOVA. Das Signifikanzniveau wurde mit dem Sidak-Test bestimmt (***) = $P \leq 0.001$; **** = $P \leq 0.0001$).

In den Luciferase-Assays mit den Vollängen-UTR-Varianten konnte gezeigt werden, dass die UTR2 keinen, oder eher einen verstärkenden Einfluss auf die Translation hat. Die gemessenen Luciferase-Aktivitäten wurden auf die des psiCHECK-2 Vektors normiert. Mit der UTR2 vor der Renilla-Luciferase waren die gemessenen Aktivitäten höher als bei der mutierten Variante. Die Insertion der G-Quadruplex-Sequenz führte zu einer Reduktion der Renilla-Luciferase-Aktivität um etwa 50 %, wie auch schon in Abbildung 19 B gezeigt wurde. Die mutierte G-Quadruplex-UTR hatte so gut wie keinen Einfluss auf die Translation.

Es konnte also gezeigt werden, dass von ARPC2 zwei Spleißvarianten existieren, die sich nur in ihrer 5'-UTR unterscheiden. Diese beiden UTR Sequenzen hatten einen unterschiedlichen Einfluss auf die Translation der nachfolgenden codierenden Sequenz. Als nächstes wurde die Expression der beiden Varianten in verschiedenen Zelllinien mittels quantitativer real time PCR (qPCR) untersucht.

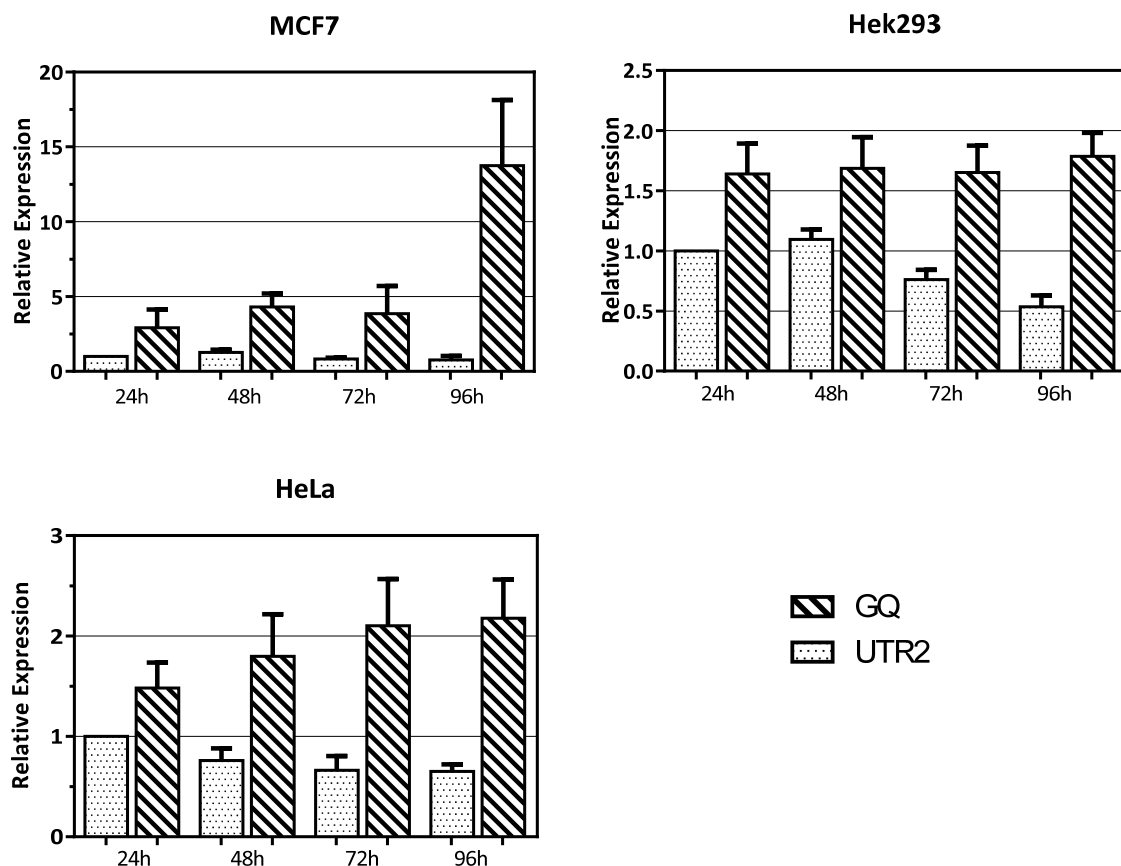


Abbildung 39: qPCR Quantifizierung der beiden UTR Varianten (G-Quadruplex und kurz) in MCF7, Hek und HeLa Zellen. Die Ct-Werte der einzelnen UTRs wurden auf die Gesamt-RNA Konzentration sowie auf die Ct-Werte der ARPC2 CDR normiert. Die Effizienz für jedes Primerpaar wurde nach der Methode von Pfaffl et al. miteinbezogen (97). Jedes Experiment wurde dreimal in Duplikaten durchgeführt. Die Ct-Werte wurden in Triplikaten bestimmt. Die Mittelwerte mit ihren Standardabweichungen sind angegeben.

Die Ergebnisse der Quantifizierung der Expression der beiden ARPC2 UTR Varianten sind in Abbildung 39 dargestellt. Die Zellen wurden jeweils in 12-Well Platten ausgesät und nach verschiedenen Zeitpunkten geerntet. Die Gesamt-RNA der Zellen wurde isoliert und mittels reverser Transkription in cDNA umgeschrieben. Anschließend wurden qPRCs mit spezifischen Primern für die codierende Region (CDR), die kurze und die G-Quadruplex-UTR durchgeführt. Die in der Quantifizierung erhaltenen Ct-Werte für die beiden UTRs wurden auf die Ct-Werte der CDR normiert und im Verhältnis zur Expression der kurzen UTR zum Zeitpunkt 24 h dargestellt. Für jedes verwendete Primerpaar wurde die Effizienz bestimmt und in die Auswertung miteinbezogen. Man kann erkennen, dass das Verhältnis von G-Quadruplex-UTR zur kurzen UTR in allen drei Zelllinien mit der Zeit immer größer wurde. Während dies in Hek und HeLa Zellen jedoch eher auf die Abnahme der kurzen Variante zurückzuführen war und die Expression der G-Quadruplex-UTR recht konstant blieb, steigerte sich die Expression der G-Quadruplex-UTR in MCF7 Zellen nach 96 h um fast das Fünffache. Interessanterweise blieb die Expression der UTR bis 72 h recht konstant und stieg dann sprunghaft an.

4.9.2. G-QUADRUPLIXE DES ARP2/3-KOMPLEXES⁴

Das Protein ARPC2 ist Teil eines aus sieben Untereinheiten bestehenden Komplexes, welcher für die Nukleation und Verzweigung von Aktinfilamenten zuständig ist. Von diesen sieben Untereinheiten besitzen sechs eine potentiell G-Quadruplex bildende Sequenz in entweder der 5'-UTR oder in der 3'-UTR (Tabelle 18). Diese G-Quadruplex-Sequenzen wurden mit dem QGRS (*Quadruplex forming G-rich sequences*) Mapper (<http://bioinformatics.ramapo.edu/QGRS/index.php> (93)) identifiziert. Betrachtet man die Gesamtzahl der G-Quadruplex-Sequenzen in den untranslatierten Regionen proteinkodierender Transkripte (Abbildung 4) ist dieses Vorkommen überproportional hoch. Außer dem in dieser Arbeit untersuchten ARPC2 besitzen noch zwei weitere Untereinheiten, ARP3 und ARPC5, G-Quadruplex-Sequenzen in der 5'-UTR. Die Untereinheiten ARP2, ARPC1B und ARPC4 besitzen potentielle G-Quadruplex-Sequenzen in ihrer 3'-UTR. Nur das Protein ARPC3 besitzt keine G-Quadruplex-Sequenz in seinen untranslatierten Regionen. Der G-score der meisten der G-Quadruplex-Sequenzen ist vergleichbar mit dem von ARPC2, welches in der Lage ist ein stabiles G-Quadruplex zu bilden. Der G-score des ARP3 G-Quadruplexes ist mit 31 am geringsten, da es einen 13 Nucleotide langen Loop besitzt.

Tabelle 18: Potentielle G-Quadruplex bildende Sequenzen im ARP2/3-Komplex. Die potentiellen G-Quadruplex bildenden Nucleotide sind unterstrichen. Der G-score wurde mit dem QGRS-Mapper ermittelt und ist ein Maß für die Wahrscheinlichkeit der Ausbildung des G-Quadruplexes (93). Die Sequenzen sind in 5'→3' Richtung angegeben.

Name	5'-UTR	3'-UTR	G-score
ARP2(a)	-	<u>GGG</u> ATGGG <u>GGG</u> TGGTTC <u>GGG</u> ATGGGT <u>GGG</u>	41
ARP2(b)	-	<u>GGGGGGGAGGGTAACAATGGG</u>	36
ARP3	<u>GGGAGGAGGCCGCGGCGGGGCAGGGCGGGG</u>	-	31
ARPC1B		<u>GGG</u> AAGC <u>GGG</u> GAGA <u>GGGGTCAGGG</u>	42
ARPC2	<u>GGGCTGGGCGGGGACCGGG</u>	-	41
ARPC3	-	-	
ARPC4	-	<u>GGG</u> AGTT <u>GGG</u> TT <u>GGGGTGGG</u>	40
ARPC5	<u>GGGTAGTGGGTTGCTGGGCTGGG</u>	-	39

Um den Einfluss der G-Quadruplex-Sequenzen des ARP2/3-Komplexes auf die Translation zu bestimmen, wurden sie in den psiCHECK-2 Reportervektor kloniert. Je nach ihrer natürlichen Position wurden sie in die 5'-UTR oder in die 3'-UTR der Renilla-Luciferase kloniert. Diese Plasmide wurden in Hek293 Zellen transfiziert und nach 24 h wurde die Luciferase-Aktivität bestimmt. Das Expressionslevel der Reporter mRNAs wurde mittels qPCR überprüft (Abbildung 40).

⁴ Die Ergebnisse dieses Abschnittes wurden von der studentischen Mitarbeiterin Delia Yang, unter meiner Anleitung, erzielt.

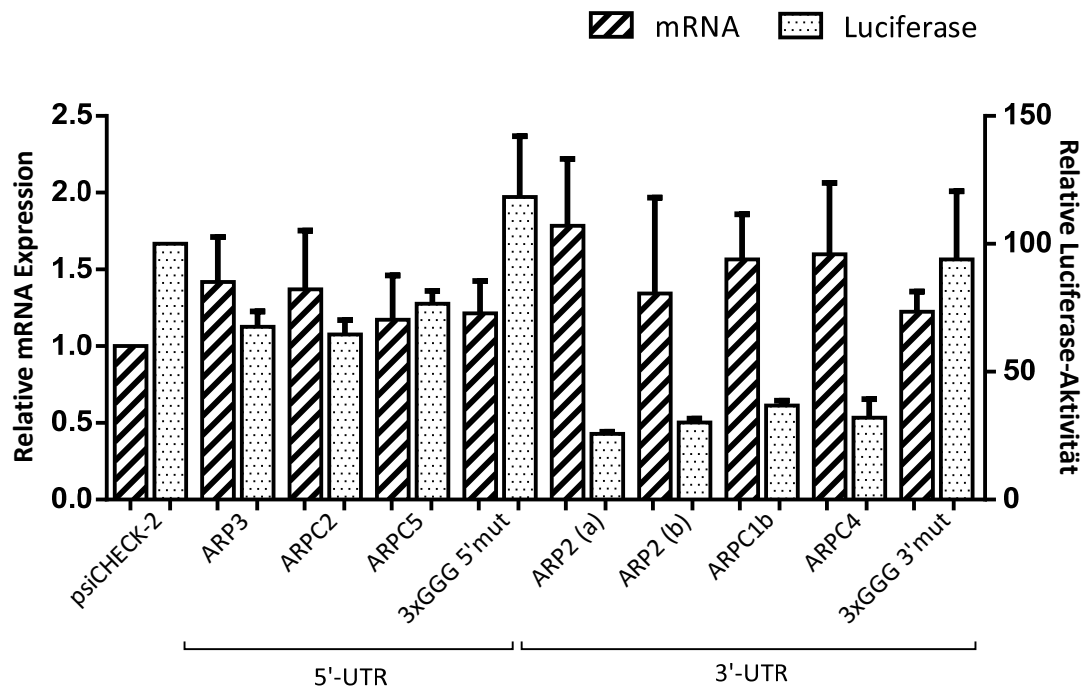


Abbildung 40: Einfluss der ARP2/3 G-Quadruplex-Sequenzen auf die Transkription (mRNA Expression) und auf die Translation (Luciferase-Aktivität). Die G-Quadruplex-Sequenzen wurden in die 5' oder in die 3'-UTR der Renilla-Luciferase des psiCHECK-2 Vektors inseriert. Der Gehalt an Reporter-mRNA wurde mittels qPCR bestimmt und das Expressionslevel des Reporterproteins wurde anhand der Luciferase-Aktivität ermittelt.

Die 3'-UTR G-Quadruplex-Sequenzen des ARP2/3-Komplexes zeigten eine stärkere Inhibition der Translation als die 5'-UTR G-Quadruplexe. Das ARP2 (a) G-Quadruplex reduzierte die Expression der Renilla-Luciferase auf ca. 25 % im Vergleich zum psiCHECK-2 Vektor. Der Gehalt an mRNA war hier sogar tendenziell höher als bei der Kontrolle. Die 5'-UTR G-Quadruplexe inhibierten die Translation ebenfalls, jedoch nur um etwa 20-40 %.

5. DISKUSSION

Die Rolle von G-Quadruplexen als regulatorische Elemente in Nucleinsäuren wird zunehmend erforscht. Die Ausbildung dieser ungewöhnlichen Sekundärstruktur nimmt Einfluss auf die verschiedensten Prozesse, an denen Nucleinsäuren beteiligt sind. So tragen G-Quadruplexe zur Telomerstruktur bei und haben in DNA Promotoren Einfluss auf die Transkription (siehe 1.2). Besonders vielfältig sind die beschriebenen G-Quadruplex-Funktionen in *messenger* RNAs, hier reguliert deren Ausbildung Prozesse wie die Translation, das alternative Spleißen, das mRNA *targeting* und die alternative Polyadenylierung. Die Ausbildung eines G-Quadruplexes in der 5'-UTR einer mRNA kann deren Translation um bis zu 80 % inhibieren (34). Dieser starke Effekt legt nahe, dass die Ausbildung eines G-Quadruplexes in der mRNA durch weitere Faktoren, wie Proteine, reguliert wird und so eine Feinabstimmung oder auch eine schnelle Regulation der Proteinexpression erlaubt. Der energetische Aufwand, den die Zelle betreibt, um eine mRNA zu synthetisieren, die dann nur zu 20 % translatiert wird, wäre sonst schwer zu erklären. Ziel dieser Arbeit war es daher, Proteine zu identifizieren, die spezifisch mit 5'-UTR G-Quadruplexen interagieren.

5.1. CHARAKTERISIERUNG DER G-QUADRUPLEXE

Die beiden in dieser Arbeit untersuchten G-Quadruplexe stammten aus den 5'-UTRs der *matrix metalloproteinase 16* und des *actin related protein 2/3 complex subunit 2* (ARPC2). Die MMP16 5'-UTR war mit 285 Basen etwas länger als die ARPC2 UTR mit 169 Basen. Beide G-Quadruplex-Sequenzen waren stark konserviert, was auf eine wichtige regulatorische Funktion der Sequenzen hindeutet. Für die Auswahl der G-Quadruplex-Sequenzen wurde zunächst die Expression der entsprechenden mRNAs nachgewiesen. Die Expression der mRNAs in der untersuchten Hek293 Zelllinie sollte sicherstellen, dass entsprechende regulatorische Proteine in dieser Zelllinie auch gebildet werden. Die Expression der beiden G-Quadruplexe aus der 5'-UTR von MMP16 und ARPC2 in Hek293 Zellen konnte mittels RT-PCR nachgewiesen werden. In HeLa Zellen wurde nur ARPC2 exprimiert. MMP16 wird normalerweise nur in Gehirn und Herz gebildet, wird jedoch in vielen Tumorgeweben überexprimiert (98).

Die Ausbildung eines G-Quadruplexes in der 5'-UTR der MMP16 mRNA wurde bereits beschrieben (37, 2.4.1). Das G-Quadruplex-Motiv ist 16 Nucleotide lang und besteht aus vier Abschnitten mit je drei Guaninen, daher ist hier nur die Ausbildung einer Struktur möglich. Die potentiellen ARPC2 G-Quadruplex-Strukturen sind vielfältiger, da die Sequenz längere Guanin-Abfolgen enthält. So werden bioinformatisch auch sechs mögliche G-Quadruplex-Strukturen vorhergesagt, bei denen jeweils andere Guanine am zentralen Stapel beteiligt sind. Für weitere Experimente, wie den Nachweis der G-Quadruplex-Struktur, wurden Oligonucleotide mit ein paar zusätzlichen Basen designt, die die

natürliche Umgebung der Sequenz simulieren sollten. Durch diese zusätzlichen Basen wäre es auch möglich, dass sich die Sequenz in eine Haarnadelstruktur faltet (Abbildung 11).

Die Faltung der ausgewählten Sequenzen in G-Quadruplexe wurde mit spektroskopischen Methoden untersucht. In der CD-Spektroskopie zeigen parallele G-Quadruplex-Strukturen ein charakteristisches Spektrum mit einem positiven Peak bei 260 nm und einem negativen Peak bei 240 nm. Dies sind typische Charakteristika von G-Quadruplex-Strukturen (12). Beide G-Quadruplex-Oligonucleotide zeigten das entsprechende Spektrum, die beiden mutierten ARPC2 Kontrollen wiesen jedoch ein recht ähnliches Spektrum auf, bei dem die beiden Peaks nur etwas weniger ausgeprägt und leicht verschoben waren. Vor allem bei der ARPC2 mut Kontrolle ist die Sequenz so stark mutiert, dass die Ausbildung eines G-Quadruplexes nicht mehr möglich ist, nur die ARPC2 2xG Kontrolle könnte durch Zusammenlagerung mehrerer Stränge ein intermolekulares G-Quadruplex bilden. Interessanterweise ist in den CD-Spektren der kürzeren ARPC2 Oligonucleotide, in denen die flankierenden Sequenzen fehlen, ein eindeutiger Unterschied zwischen der G-Quadruplex und der mutierten Sequenz zu erkennen (Abbildung 15). Daher liefert das CD Spektrum zwar einen Hinweis auf die Ausbildung einer G-Quadruplex-Struktur, reicht aber nicht als eindeutiger Beweis der Ausbildung.

G-Quadruplexe zeigen ein typisches sigmoidales Schmelzprofil mit einer maximalen Absorptionsdifferenz bei einer Wellenlänge von 295 nm, im Gegensatz zu anderen Nucleinsäuresekundärstrukturen, deren Ausbildung bei 260 nm verfolgt wird (13). Diese Schmelzprofile waren für beide G-Quadruplex-Sequenzen klar von ihren Kontrollen unterscheidbar. Die ARPC2 G-Quadruplex Schmelzkurve steigt leicht an, bis der Schmelzpunkt erreicht ist, dies könnte auf die Ausbildung einer nicht-G-Quadruplex Sekundärstruktur hindeuten, die sich hier auflöst. Die Schmelzkurve des kurzen ARPC2 Oligonucleotids stützt diese Theorie, sie zeigt keinen solchen Anstieg. Durch die fehlenden flankierenden Nucleotide kann sich hier keine Haarnadelstruktur ausbilden. Das von ARPC2 gebildete G-Quadruplex war intramolekular und wurde nicht zwischen mehreren Strängen gebildet, da die Schmelztemperatur nicht konzentrationsabhängig war (Abbildung 14).

Das MMP16 G-Quadruplex Oligonucleotid bildete eine sehr stabile Sekundärstruktur, die bei einer KCl-Konzentration von 100 mM nicht aufgeschmolzen werden konnte. Die Schmelztemperatur bei 10 mM KCl betrug 69°C. Morris et al. bestimmten für die MMP16 G-Quadruplex-Sequenz bei 1 mM KCl eine Schmelztemperatur von 72°C (39). Das verwendete Oligonucleotid war etwas kürzer, statt fünf flankierender Oligonucleotide enthielt es nur zwei auf jeder Seite der G-Quadruplex-Sequenz. Vielleicht wirkten die zusätzlichen Nucleotide destabilisierend auf die Struktur. Das kurze MMP16 G-Quadruplex Oligonucleotid, das ausschließlich aus der G-Quadruplex-Sequenz bestand, besaß eine höhere Schmelztemperatur. Allerdings war diese bei 1mM KCl mit 63°C immer noch etwas geringer als beschrieben.

Zusammenfassend konnte gezeigt werden, dass beide Sequenzen stabile G-Quadruplexe ausbilden, wobei das ARPC2 G-Quadruplex Oligonucleotid vielfältige Strukturen einnehmen konnte.

5.2. EINFLUSS DER G-QUADRUPLIXE AUF DIE TRANSLATION

Beide G-Quadruplex-Sequenzen waren in der Lage, die Translation im Dual-Luciferase-Reporter-Assay zu inhibieren. In der 5'-UTR der Renilla-Luciferase reduzierten sie die Expression um etwa 40-50 % im Vergleich zum Ausgangsvektor. Da das ARPC2 G-Quadruplex noch nicht beschrieben war, wurde hier auch die 5'-UTR Volllängensequenz untersucht. Ihre Insertion vor die Renilla-Luciferase hatte einen ähnlichen Einfluss wie die der isolierten G-Quadruplex-Sequenz. Die Insertion der mutierten UTR führte zu einer auf 84 % verringerten Luciferase-Aktivität im Vergleich zum Ausgangsvektor. Der Einfluss der MMP16 Volllängen-UTR wurde bereits von Morris et al. beschrieben. Sie fanden eine Reduktion der Translation um ca. 60 %, wobei der mRNA Gehalt stabil blieb.

Der Einfluss der Position des G-Quadruplexes auf die Inhibition der Translation ist strittig. Kumari et al. beschrieben, dass die NRAS G-Quadruplex-Sequenz die Translation nur inhibierte, wenn sie sich innerhalb der ersten 50 Basen nach dem 5'-Cap befindet (43). Im Gegensatz dazu fanden Halder et al. für eine künstliche G-Quadruplex-Sequenz keinen Einfluss der Position in der UTR auf die Inhibition der Translation (44). Eine Erklärung für diese widersprüchlichen Ergebnisse könnte eine Studie liefern, in der gezeigt wird, dass die G-Quadruplexe am 5'-Ende der mRNA sowohl aktivierend als auch inhibierend auf die Translation wirken können (99, 100). So wurde hier die Expression der Proteine *AKT interacting protein* und Cathepsin B durch ein 5'-UTR G-Quadruplex gehemmt, während die Proteinsynthese von FOXE3 gesteigert wurde. Die Autoren schlagen vor, dass der Einfluss des G-Quadruplexes auf die Translation maßgeblich von weiteren Faktoren wie der flankierenden Sequenz des G-Quadruplexes, Trans-aktivierenden Faktoren (z.B. Proteine) und der zellulären Umgebung beeinflusst wird.

Bei der Verwendung der Renilla-Luciferase aus dem Vektor psiCHECK-2 als Reporter werden vor der inserierten Sequenz noch ca. 70 Basen transkribiert, so dass man aus den Ergebnissen nicht auf den Einfluss der G-Quadruplex-Struktur nahe am 5'-Cap schließen kann. Die Position relativ zum Start-Codon verändert sich jedoch, wenn man die isolierte Sequenz und die Volllängen-UTR betrachtet. Die isolierten Sequenzen befinden sich direkt vor dem AUG, in beiden Volllängen-UTRs ist der Abstand zwischen G-Quadruplex und Start-Codon größer. Da für beide untersuchten Sequenzen (ARPC2 und MMP16) die Inhibition der Translation für die isolierte Sequenz und für die gesamte UTR recht ähnlich ist, scheint das G-Quadruplex keinen Einfluss auf das Erkennen des Start-Codons zu haben. Es wäre möglich, dass es das Scannen der kleinen Untereinheit entlang der UTR inhibiert.

5.3. RNA G-QUADRUPLIX BINDENDE PROTEINE

Ein Ziel dieser Arbeit war es Proteine zu identifizieren, die spezifisch an 5'-UTR RNA G-Quadruplexe binden. Durch Pull-down-Assays mit den MMP16 und ARPC2 Oligonucleotiden und Hek293 sowie HeLa Zelllysate konnten mehrere Proteine identifiziert werden, die die G-Quadruplex-Strukturen erkennen. Unspezifisch bindende Proteine oder Proteine, die sequenzspezifisch binden, sollten über die Kontroll-Oligonucleotide ausgeschlossen werden. Die meisten der Proteine waren spezifisch für ein G-Quadruplex, manche konnten auch beide Strukturen erkennen. Proteine könnten entweder die zentrale G-Quadruplex-Struktur binden, oder über die Loops mit dem G-Quadruplex interagieren. Für das ARPC2 G-Quadruplex wurden wesentlich mehr Proteine gefunden als für MMP16. Dies zeigte sich sowohl in den gelbasierten Pull-down-Assays mit Hek293 und HeLa Zelllysate, als auch bei der Bestätigung der Experimente über den kombinierten Ansatz aus SILAC und Orbitrap. Eine Erklärung hierfür könnten die vielfältigeren Strukturen sein, die die ARPC2 G-Quadruplex-Sequenz einnehmen kann. Obwohl die G-Quadruplexe nach ihrem Einfluss auf die Translation ausgewählt wurden, waren viele der identifizierten Proteine in den Spleißprozess involviert. Eine weitere große Gruppe bildeten die ribosomalen Proteine. Im Pull-down-Assay kann man nicht unterscheiden, ob G-Quadruplex und Protein direkt interagieren oder ob Proteine als Teil eines Proteinkomplexes an das G-Quadruplex binden. Für einige der gefundenen Proteine wurde beschrieben, dass sie in einem Komplex vorkommen. So besteht z. B. der Nucleolin-bindende Ribonucleoproteinkomplex unter anderem aus den Proteinen Nucleolin, RPL6, RPL7, RPL10, RPL27a, RPS6, RPS9, YB-1 und hnRNP, die alle als ARPC2 Binder identifiziert wurden (101). Es wäre also möglich, dass diese Proteine eine gewisse Affinität zueinander besitzen und so auch *in vitro* interagieren. Ausgewählte Kandidaten wurden heterolog in *E. coli* exprimiert, um die Bindung an die G-Quadruplex-Oligonucleotide weiter zu charakterisieren.

5.3.1. NUCLEOLIN

Nucleolin wurde als einziges Protein in allen fünf gelbasierten Pull-down-Assays als ARPC2 G-Quadruplex-Binder identifiziert und auch in dem SILAC Ansatz band Nucleolin an das ARPC2 G-Quadruplex. Die Bindungskonstanten von Nucleolin und den ARPC2 G-Quadruplex und mut, sowie den MMP16 G-Quadruplex und mut Oligonucleotiden wurden mittels SPR-Messungen bestimmt. Dabei zeigte Nucleolin eine große Affinität für das ARPC2 G-Quadruplex mit einer Dissoziationskonstante von 0,6 nM. Die Bindung an das mutierte ARPC2 Oligonucleotid war mit 83 nM sehr viel geringer. Allerdings zeigte Nucleolin eine genauso hohe Affinität (0,6 nM) für die MMP16 mut Kontrolle wie für das ARPC2 G-Quadruplex, was es als spezifisch G-Quadruplex bindendes Protein ausschließt. Die Affinität für das MMP16 G-Quadruplex war mit einer K_D von 2 nM auch sehr hoch.

Erstaunlich war die Diskrepanz zwischen den Ergebnissen aus den Elektrophoretic mobility shift Assays (EMSA) und den SPR-Messungen. Der offensichtlichste Unterschied zwischen den beiden Methoden ist der experimentelle Aufbau. Im EMSA befinden sich beide Bindungspartner zunächst in Lösung und anschließend in einem Gel. Im Gegensatz dazu ist bei den SPR-Messungen ein Partner auf dem Chip immobilisiert, während der zweite in einem Flüssigkeitsstrom an ihm vorbeigeleitet wird. Dies gibt den Komplexen im EMSA mehr Gelegenheit sich auszubilden, während die Bedingungen in den SPR-Messungen stringenter sind. So wurde beschrieben, dass viele Komplexe auf dem Gel wesentlich stabiler sind als in Lösung (102–104). Dies kann die Bewertung von Bindungsaffinitäten im EMSA erheblich beeinflussen. Ein weiterer Faktor, der zu der hohen Affinität von Nucleolin für das ARPC2 mut Oligonucleotid im EMSA beitragen kann, ist die schnelle Assoziationskinetik des Komplexes. Die Geschwindigkeitskonstante der Assoziation (k_a) betrug $0,5 \text{ M}^{-1}\text{s}^{-1} \times 10^6$, was eine schnelle Assoziation der Bindungspartner bedeutet. Die Assoziation war hier sogar schneller als für die anderen drei RNA-Sequenzen. Die größere Stabilität der anderen Komplexe ergibt sich nur aus ihrer langsameren Dissoziation. Wird der gebildete Komplex im Gel stabilisiert, kann das die unterschiedlichen Ergebnisse der beiden Methoden erklären.

Nucleolin ist ein multifunktionales Protein und Hauptbestandteil der Nucleoli von eukaryotischen Zellen. Es besteht aus einer sauren N-terminalen Domäne, vier zentralen RNA Bindedomänen (RBDs) und einer C-terminalen RNA Bindedomäne (RGG-Domäne) (105). Dieser Aufbau aus mehreren Domänen erlaubt die Interaktion mit verschiedenen Proteinen und Nucleinsäuren. Es ist unter anderem beteiligt an der Chromatinstruktur, der rDNA Transkription, der rRNA Reifung, dem Zusammenbau der Ribosomen und dem Nucleo-Cytoplasmatischen Transport (95). Es wurden bereits mehrere Nucleolin-Bindemotive beschrieben. So wurde in einer Studie, die global Nucleolin-Ziel-RNAs identifizierte, ein G-reiches Konsensus-Motiv gefunden, welches wahrscheinlich eine Haarnadelstruktur bildet (106). Die Autoren dieser Studie spekulieren, dass sich die Affinität Nucleolins für RNA-Sequenzen nach Stimulation durch Wachstumsfaktoren, Mitogene, Stressoren oder andere Stimuli ändern könnte. Auf prä-ribosomaler RNA bindet Nucleolin an ein Element, welches ebenfalls eine Haarnadelstruktur bildet und die die Sequenz UCCCGA im Loop besitzt (107). Hierbei bilden die RBDs 1 und 2 eine Klammer und stabilisieren so die Haarnadelstruktur. Auch eine Interaktion mit einem AU-reichen Element in der 3'-UTR der Bcl-2 mRNA wurde gefunden (108). Für das DNA G-Quadruplex im MYC Promotor wurde Nucleolin ebenfalls als Binder beschrieben, an dieser Bindung sind die RNA Bindedomänen 3 und 4 und die RGG-Domäne beteiligt (62, 109). Hier wirkte es stabilisierend auf die Struktur. Es ist also möglich, dass die einzelnen Domänen für verschiedene Bindungsspezifitäten verantwortlich sind. Auch als Teil des Heterodimers LR1, bestehend aus hnRNP D und Nucleolin, wurde Nucleolin als G-Quadruplex-Binder beschrieben (63). Für diese Interaktion mit einem DNA G-Quadruplex wurde eine Dissoziationskonstante von 0,4 nM bestimmt, was gut mit den hier

bestimmten Werten übereinstimmt. In den vergleichenden SPR-Messungen konnte gezeigt werden, dass Nucleolin die kürzere Version des ARPC2 G-Quadruplex Oligonucleotids genauso stark bindet. Die Interaktion scheint also nicht über die Haarnadelstruktur, die über die flankierenden Oligonucleotide gebildet werden kann, zustande zu kommen (Abbildung 25). Die entsprechende ARPC2 G-Quadruplex DNA Sequenz wurde sehr schlecht gebunden, ebenso wie die MMP16 G-Quadruplex DNA Sequenz. Eine Erkennung der Guanin-Abfolgen ist auch ausgeschlossen, da die Affinität zu der ARPC2 2xG Kontrolle nicht sehr groß war. Interessanterweise war die Affinität zur Bcl-2 RNA G-Quadruplex-Sequenz auch sehr hoch (Abbildung 26). Nucleolin zeigte also zu allen untersuchten RNA G-Quadruplexen eine hohe Affinität, eine Interaktion über eine andere Bindedomäne könnte die ebenfalls gute Bindung an das MMP16 mut Kontrolloligonucleotid erklären.

Auch wenn Nucleolin primär in den Nucleoli der Zellen vorliegt, ist es in Krebszellen oder unter Stressbedingungen auch im Cytoplasma, Nucleoplasma und auf der Zellmembran lokalisiert (110). Des Weiteren wurde beschrieben, dass Nucleolin zwischen dem Nucleolus und dem Cytoplasma pendelt, und so am nucleo-cytoplasmatischen Transport beteiligt ist (111). Die cytoplasmatische Lokalisierung Nucleolins wurde auch mit dem oben beschriebenen G-reichen RNA-Bindemotiv bestätigt. In Pull-down-Assays mit den entsprechenden RNA-Oligonucleotiden und cytoplasmatischem Zellextrakt konnte die Bindung von Nucleolin an diese Sequenzen gezeigt werden (106). Es ist also durchaus möglich, dass Nucleolin in der Zelle mit mRNA in Kontakt kommt und Einfluss auf die Translation nimmt.

5.3.2. SPLEIßFAKTOREN

U2AF65 (*U2 auxiliary factor 65*) bildet zusammen mit U2AF35 das Heterodimer U2AF (112). U2AF65 konnte als einziges Protein an beide G-Quadruplex-Sequenzen binden. Es wurden Bindungskonstanten von 1 nM für ARPC2 G-Quadruplex und 5 nM für MMP16 G-Quadruplex bestimmt. Zu den beiden mutierten Kontrollen hatte es in den SPR-Messungen eine so geringe Affinität, dass mit dem verwendeten experimentellen Aufbau keine Bindungskonstanten bestimmt werden konnten. Wie schon für Nucleolin, stehen auch hier die Ergebnisse der EMSAs im Kontrast zu denen der SPR-Messungen. Man konnte zwar eine etwas höhere Affinität für das G-Quadruplex-Oligonucleotid als für die Kontrollsequenz erkennen, der Unterschied in der Retardation der Banden war jedoch nur sehr gering. Hier wäre es ebenfalls möglich, dass ein sonst schnell dissoziierender Komplex auf dem Gel stabilisiert wird. Eine weitere Ursache für die unterschiedlichen Ergebnisse könnte in der Immobilisierung des Proteins liegen. Für die SPR-Messungen wurden die Proteine über Aminkopplung immobilisiert. Je nach Verteilung oder Zugänglichkeit der Aminogruppen kann es dabei zu einer nicht

zufälligen Orientierung des Proteins auf dem Sensorchip kommen. Wenn dadurch z.B. Bindedomänen von Proteinen blockiert werden, kann dies die Ergebnisse verfälschen (113).

Zu der ARPC2 2xG Kontrolle zeigte U2AF65 ebenfalls eine niedrige Affinität. In den vergleichenden SPR-Messungen interagierten die entsprechenden DNA-G-Quadruplex Oligonucleotide recht stark mit U2AF65. Für die beiden G-Quadruplex-Sequenzen aus der Bcl-2 und NRAS 5'-UTR zeigte es ebenfalls eine hohe Affinität (Abbildung 30). Zusammengenommen zeigten diese Ergebnisse, dass U2AF65 die G-Quadruplex-Struktur erkennt und nicht Abfolgen von Guaninen. U2AF65 ist beteiligt am Spleißen von Prä-mRNA. Es bindet an den Polypyrimidin-Trakt zwischen der Verzweigungsstelle und der Spleiß-Akzeptorstelle, die durch ein AG-Dinucleotid markiert wird, und dirigiert so das Spleißosom an die 3' Spleiß-Stelle. U2AF65 ist aufgebaut aus einer N-terminalen Arginin-Serin-reichen Domäne, einer U2AF35 Interaktionsdomäne und aus drei C-terminalen RNA Erkennungsmotiven (*RNA recognition motif*, RRM) (114). Der Polypyrimidin-Trakt wird von den RRM 1 und 2 erkannt. Der Polypyrimidin-Trakt ist eine degenerierte Erkennungssequenz, die je nach Cytidin- und Uridin-Gehalt verschiedene Aktivität zeigt. Ein höherer Gehalt an Uridin erhöht in der Regel die Nutzung der benachbarten 3' Spleiß-Stelle (115). Das dritte RRM interagiert mit dem *splicing factor 1*, der an die Verzweigungsstelle bindet, und könnte sowohl mit Proteinen als auch mit RNA interagieren (116). U2AF65 wurde auch als Triplex-DNA bindendes Protein beschrieben (117). Hier wurden Proteine aus Tumorgewebe und Zelllysate analysiert, die an eine Triplex-DNA Sonde binden konnten. U2AF65 wurde als das mengenmäßig am meisten gebundene Protein aus cytoplasmatischem Extrakt gefunden. Die Autoren spekulieren, dass die verwendete DNA Sequenz nicht nur Triplex-Strukturen annehmen kann, sondern auch andere mehrsträngige Strukturen wie D-Loops, R-Loops und G-Quadruplexe und dass U2AF65 daher auch diese Strukturen erkennen könnte.

Weitere hier identifizierte Proteine, die am Spleißprozess beteiligt sind, gehören zu den SR- (*serine-arginine rich*) Proteinen und den hnRNPs. SR-Proteine fördern im Allgemeinen das Spleißen, indem sie an exonische Sequenzen binden und Komponenten des Spleißosoms rekrutieren (118). Dies geschieht nicht nur beim alternativen, sondern auch beim konstitutiven Spleißen. Im Gegensatz dazu binden hnRNPs an Exon- oder Intron-Sequenzen und wirken meist hemmend auf den Spleißprozess durch das Verdecken von Spleiß-Stellen (119). Die Erkennungssequenzen beider Proteingruppen befinden sich in der Nähe der Exon/Intron Übergänge, daher nimmt man an, dass das Gleichgewicht aus aktivierenden und inaktivierenden Faktoren über die Inklusion eines Exons entscheidet. Aus der Gruppe der SR-Proteine wurden hier SRSF1 und SRSF9 als Binder von ARPC2 identifiziert. Mitglieder der SR-Proteinfamilie zeichnen sich durch eine Arginin-Serin reiche Domäne, mit vielen Wiederholungen dieses Dipeptids, und mindestens einer RNA-bindenden RRM Domäne aus. Die Anzahl der SR-Proteine in einem Organismus hängt mit dem Ausmaß des alternativen Spleißens, das in ihm stattfindet, zusammen. So gibt es in der Spalthefe *Schizosaccharomyces pombe* zwei Vertreter der SR-Proteine, im

Menschen 12 (119). Die 12 SR-Proteine des Menschen können phylogenetisch in 6 Familien eingeteilt werden, interessanterweise bilden die hier identifizierten Proteine SRSF1 und SRSF9 eine Familie (120). Beide Proteine zeigten eine hohe Affinität für das ARPC2 G-Quadruplex, SRSF1 wurde in vier und SRSF9 in drei von insgesamt fünf Experimenten gefunden. Die Bindung zwischen SRSF1 und den G-Quadruplex bildenden sowie den mutierten Oligonucleotiden wurde weiter untersucht. Da SRSF1 in den Oberflächenplasmonresonanz-Messungen sehr schnell an Aktivität verlor, ließen sich für die Interaktionen keine Bindungskonstanten bestimmen. Die Bindung der einzelnen Sequenzen an das Protein ist als Vergleich gezeigt (Abbildung 33). In der SPR-Messung konnte bestätigt werden, dass SRSF1 eine größere Affinität für die ARPC2 G-Quadruplex-Sequenz als für die anderen untersuchten Sequenzen besitzt. Auch in dem zweiten Pull-down Ansatz über SILAC und Orbitrap wurde SRSF1 als Binder identifiziert. Zusammengefasst deuten diese Ergebnisse auf eine biologisch relevante Interaktion hin.

SR-Proteine können in allen Schritten des mRNA Metabolismus mitwirken, von der Synthese bis zum Abbau. SRSF1 kann Einfluss auf die Translation nehmen, dieser kann indirekter oder direkter Art sein. So beeinflusst die Überexpression von SRSF1, wie sie in verschiedenen humanen Tumoren auftritt, das Spleißen der Prä-mRNA der Proteinkinase MNK2 (*mitogen activated protein kinase*) (121). Diese Kinase reguliert die Translationsinitiation. Ein hoher Gehalt an SRSF1 fördert die Bildung einer MNK2 Spleißvariante, welche durch Phosphorylierung von eIF4E die Cap-abhängige Translation verstärkt. SRSF1 kommt nicht nur nukleär als Spleißfaktor vor, sondern wird auch assoziiert mit der Translationsmaschinerie im Cytoplasma gefunden (122). Der Großteil der SR-Proteine befindet sich im Nucleus, ein Teil von ihnen pendelt jedoch ständig zwischen Nucleus und Cytoplasma (123). Dieses *shuttling* von SRSF1 in das Cytoplasma kann die Translation unabhängig von seiner RNA-Bindeaktivität stimulieren. Des Weiteren wurde beschrieben, dass die SRSF1 Bindung an eine 3'-UTR deren Degradation bewirkt (124). Eine purinreiche Konsensussequenz in Exonen wurde als RNA Bindemotif für SRSF1 identifiziert (125). Da SRSF1 an so vielen Schritten des RNA Metabolismus beteiligt ist, könnte die Bindung an das ARPC2 G-Quadruplex verschiedene Prozesse beeinflussen. Ein direkter Einfluss auf die Translation konnte weder *in cellulo* durch Co-Transfektionen, noch im *in vitro* Translationsassay gezeigt werden.

Auch SRSF9 ist ein Protoonkogen (126). Es wird in vielen Tumoren überexprimiert. Eine Überexpression führte zu maligner Transformation von Fibroblasten und induzierte so Tumorbildung in Mäusen. Für SRSF9 wurde bisher nur eine Beteiligung an konstitutiven und alternativen Spleißprozessen beschrieben und es scheint ausschließlich im Zellkern vorzukommen (123).

Die heterogenen nukleären Ribonucleoproteine (hnRNPs) sind eine Gruppe multifunktionaler Proteine, die an allen Schritten des mRNA Metabolismus, vom Spleißen über Export und Lokalisierung bis zu Translation und Stabilität, beteiligt sind. Hier wurden die hnRNPs F und H als Binder der ARPC2

G-Quadruplex-Sequenz gefunden und hnRNP U wurde dreimal als MMP16 G-Quadruplex Binder identifiziert. hnRNP F und H unterscheiden sich von den anderen hnRNPs in ihren RNA Bindedomänen (127). Diese werden daher quasi RRM (qRRMs) genannt und erkennen ihr Substrat durch eine Interaktion mit drei hochkonservierten Loops und nicht durch Interaktion mit einer β -Faltblattoberfläche. Daher besitzen sie auch eine andere Substratspezifität und binden spezifisch an poly(G)-Abschnitte (128). Beide hnRNPs sind bekannt für ihre Rolle im alternativen Spleißen, wo mehrere Proteine an die Zielsequenzen binden und das Intron dann als Schleife herausragt. Sie sind sich sehr ähnlich in ihrer Sequenz, Struktur und in ihren Bindepräferenzen, haben aber antagonistische Funktionen in der Polyadenylierung und binden an verschiedene regulatorische Elemente in Genen (129, 130). hnRNP F liegt vor allem cytoplasmatisch vor, während hnRNP H sich hauptsächlich im Zellkern befindet. Es wurde gezeigt, dass die qRRM3 von hnRNP F selektiv den einzelsträngigen Guanin-Abschnitt und nicht das gefaltete G-Quadruplex einer G-reichen RNA Sequenz bindet (131). Auf diese Weise könnte hnRNP F die Ausbildung von G-Quadruplexen verhindern. In einer anderen Studie wird jedoch beschrieben, dass hnRNP F an die G-Quadruplex-Struktur in der 3'-UTR der p53 mRNA bindet, was zur Polyadenylierung dieser mRNA führt (132). Hier wurde hnRNP F nur als Binder der G-Quadruplex-Sequenz gefunden. In dem Pull-down-Assay mit dem G-reichen Kontrolloligonukletid wurde es nicht identifiziert. Auch für die MMP16 G-Quadruplex-Sequenz wurde es nicht gefunden, was gegen eine einfache Erkennung der Guanin-Abschnitte spricht, die für diese Sequenz genauso stattfinden müsste. Andererseits hinge die Bindung an einzelsträngige Sequenzen vom Gleichgewicht zwischen gefalteter und ungefalteter Form ab. Da MMP16 die stabilere G-Quadruplex-Struktur ausbilden konnte, würde das Gleichgewicht auf Seite der gefalteten Sequenz liegen und deshalb auch weniger hnRNP F binden.

hnRNP U ist mit 120 kDa das größte Protein aus der Familie der hnRNPs. Es bindet an A/T-reiche Doppelstrang-DNA und an G/U-reiche RNA Motive (133, 134). Es ist, neben vielen weiteren beschriebenen Funktionen, wie die anderen hnRNPs hauptsächlich am Spleißprozess beteiligt. Eine genomweite Analyse der hnRNP U gebundenen RNAs ergab, dass vor allem intronische Sequenzen gebunden werden, aber auch viele nicht-codierende regulatorische RNAs (135). Eine Interaktion mit G-Quadruplexen wurde bis jetzt noch nicht beschrieben.

Zusammengenommen wurden, vor allem für das ARPC2 G-Quadruplex, viele Proteine gefunden, die am konstitutiven oder alternativen Spleißen beteiligt sind. Dass G-Quadruplexe an der Erkennung von Spleißstellen beteiligt sein können, wurde bereits beschrieben (1.2.3). Es wäre möglich, dass ein G-Quadruplex als Bindungsstelle für Spleißfaktoren dient und so das Spleißen beeinflusst. Da für die 5'-UTR der ARPC2 mRNA zwei Spleißvarianten existieren, von denen nur eine die G-Quadruplex-Sequenz enthält, wäre es möglich, dass über das G-Quadruplex Einfluss darauf genommen wird, welche Spleißvariante entsteht.

Die in dieser Arbeit untersuchten G-Quadruplexe wurden ursprünglich wegen ihrer Rolle in der Translation ausgewählt, um Bindungspartner zu identifizieren die auf diesen Prozess Einfluss nehmen. Die in den Pull-down Assays identifizierten Spleißfaktoren deuten jedoch auf eine zusätzliche Rolle der G-Quadruplexe in diesem Prozess hin. Dies sollte noch näher untersucht werden.

5.3.3. RIBOSOMALE PROTEINE

Sehr auffällig war, dass ein sehr großer Anteil der ARPC2 G-Quadruplex-Binder ribosomale Proteine waren. Dies konnte in beiden Ansätzen beobachtet werden. Die meisten der gefundenen ribosomalen Proteine gehörten zur großen 60S Untereinheit des Ribosoms, aber auch viele Proteine der 40S Untereinheit wurden identifiziert. Eine Theorie besagt, dass die Hemmung der Translation darauf beruht, dass das Scannen nach dem Start-Codon durch die Sekundärstruktur behindert wird (41). Der Präinitiationskomplex, der an das 5'-Cap der mRNA bindet und dann nach dem Start-Codon sucht, besteht aus der 40S Untereinheit des Ribosoms und weiteren Initiationsfaktoren. Wird das Start-Codon erkannt, bindet die 60S Untereinheit an die 40S Untereinheit und die Translation wird gestartet (42). Da nun jedoch die G-Quadruplex bindenden Proteine in einem *in vitro* Assay mit G-Quadruplex bildenden Oligonucleotiden identifiziert wurden, sollte man hier nicht erwarten, ribosomale Untereinheiten zu identifizieren, die beim Scannen nach dem Start-Codon auf ein Hindernis gestoßen sind. Die Inhibition der Translation würde eher auf einer unspezifischen, sterischen Hinderung des Scanvorgangs durch den Präinitiationskomplex beruhen, als auf einer spezifischen Bindung einer ribosomalen Untereinheit an das G-Quadruplex. Ein inhibitorischer Effekt auf die Translation wurde nicht nur für G-Quadruplexe, sondern auch für anderweitig strukturierte 5'-UTRs beschrieben. Diesem Effekt muss also ein allgemeiner Mechanismus zugrunde liegen (136, 137). Gegen eine spezifische Bindung durch eine ribosomale Untereinheit spricht auch, dass für das MMP16 G-Quadruplex viel weniger ribosomale Proteine gefunden wurden, obwohl dieses G-Quadruplex die Translation mindestens genauso stark inhibierte wie das ARPC2 G-Quadruplex. Es muss also einen anderen Grund geben, warum so viele ribosomale Proteine spezifisch an das ARPC2 G-Quadruplex gebunden haben. Es werden immer mehr extraribosomale Funktionen ribosomaler Proteine bekannt, dies erklärt jedoch auch nicht das gehäufte Binden so vieler ribosomaler Proteine (138). Eine Möglichkeit wäre, dass das G-Quadruplex Teil einer *internal ribosome entry site* (IRES) ist, wie es zum Beispiel für das VEGF G-Quadruplex bereits beschrieben ist (47). Diese Elemente sind vor allem in der viralen Translation gut untersucht (139). Über sie kann die Proteinsynthese ohne das 5'-Cap der mRNA eingeleitet werden, das normalerweise die Erkennungsstelle für die 40S Untereinheit der Ribosomen ist. Da virale Transkripte keine Cap-Struktur besitzen, wird hier die Translation über IRES-Elemente eingeleitet. Bei Eukaryoten ist diese Form der Translationsinitiation weniger gut definiert (45). Es sind bereits zahlreiche zelluläre IRES-Elemente bekannt, eine Konsensussequenz oder -struktur scheint es jedoch

nicht zu geben. Die Translation über zelluläre IRESs wird vor allem mit stressinduzierten Reaktionen der Zelle in Verbindung gebracht, bei denen die Cap-abhängige Translation beeinträchtigt ist, wie z. B. bei Nährstoffmangel oder Apoptose (45, 140). Für die G-Quadruplexe aus der VEGF und aus der FGF2 5'-UTR wurde eine solche Funktion als Teil einer IRES bereits gezeigt (46, 47). Auch die Sequenz um das ARPC2 G-Quadruplex ist sehr GC-reich und hätte das Potential ausgeprägte Sekundärstrukturen einzunehmen, so wäre es durchaus möglich, dass das G-Quadruplex Teil einer IRES ist.

5.3.4. *PROTEINE OHNE BEKANNTE RNA-BINDEAKTIVITÄT*

Für zwei der wiederholt identifizierten G-Quadruplex bindenden Proteine war bisher nicht bekannt, dass sie RNA-Bindeaktivitäten besitzen. Das Protein EF-hand domain family, member D2 (EFHD2) wurde dreimal als ARPC2 G-Quadruplex-Binder gefunden. Das mitochondriale NAD-abhängige Malat-Enzym (ME2) wurde ebenfalls dreimal als MMP16 G-Quadruplex-Binder identifiziert. Auch im SILAC Pull-down Experiment war dies das Protein, das, verglichen mit der Kontrolle, am meisten an MMP16 G-Quadruplex band (Ratio Heavy/Light = 17,4).

EFHD2 wurde nach seiner calciumbindenden Domäne, der EF-Hand benannt. Diese Domäne bildet ein Helix-Loop-Helix Motiv und wurde in vielen Ca^{2+} -bindenden Proteinen gefunden. EFHD2 ist beteiligt am B-Zell-Rezeptor vermittelten Ca^{2+} Einstrom in unreife B-Lymphozyten (141). Eine Überexpression von EFHD2 verstärkte die B-Zell-Rezeptor vermittelte Apoptose, während eine Herunterregulation die Apoptose unterdrückte. Bemerkenswert ist, dass EFHD2, wie ARPC2, in Membranausstülpungen und Lamellipodien lokalisiert ist (142). Eine Überexpression förderte die Bildung von Lamellipodien und die Ausbreitung von Zellen. EFHD2 besitzt mindestens drei Bindestellen, über die es direkt mit Aktin interagieren kann. Während ARPC2 an der Verzweigung und Nukleation von Aktinfilamenten beteiligt ist, induziert EFHD2 die Bildung von Aktinbündeln aus den einzelnen Filamenten. Beide Prozesse sind wichtig für die Ausbildung von Membranausstülpungen wie den Lamellipodien. Durch den Calcium-Chelator EGTA wurde die Bildung von Aktinbündeln durch EFHD2 reduziert, die Bindung an Aktin jedoch nicht verringert. Möglicherweise steuert Ca^{2+} die Dimerisierung von EFHD2, die für die Aktivität zur Bildung von Aktinbündeln notwendig ist. Eine weitere Verbindung zwischen EFHD2 und dem ARP2/3-Komplex stellt die Phosphorylierung nach Aktivierung des EGF-Rezeptors (epidermal growth factor-Rezeptor) dar (91). Alle Untereinheiten des ARP2/3-Komplexes werden nach Stimulierung von HeLa Zellen durch EGF phosphoryliert. Diese Phosphorylierung zeigt die Signalweitergabe des Rezeptors an seine Effektormoleküle an. EFHD2 zeigte das gleiche Aktivierungsmuster nach EGF-Stimulierung wie die ARP2/3 Untereinheiten, mit einem Peak 5 min nach Stimulation und einem weiteren Anstieg der Aktivierung nach 20 min. Die Entdeckung, dass EFHD2 eine Affinität für die G-Quadruplex-Sequenz der ARPC2 mRNA zeigt, deutet auf eine neue Art der Interaktion zwischen den

beiden Proteinen hin. Da die ARPC2 mRNA zur Translation in die Membranausstülpungen transportiert wird und EFHD2 ebenfalls dort lokalisiert ist, könnte diese Interaktion auch *in vivo* stattfinden. Eine Dissoziationskonstante für die Bindung zwischen EFHD2 und ARPC2 G-Quadruplex zu bestimmen, war nicht möglich, da das Protein nach der Immobilisierung zu schnell an Aktivität verlor. Man kann jedoch sowohl an den Gelretardationsexperimenten als auch an den Oberflächenplasmonresonanzmessungen erkennen, dass EFHD2 an die ARPC2 G-Quadruplex-Sequenz band, während es nur eine sehr geringe Affinität für die anderen Sequenzen zeigte. Durch Co-Transfektionen mit dem Luciferase-Reportervektor konnte kein direkter Einfluss auf die Translation gezeigt werden, daher ist die Funktion der Bindung an das G-Quadruplex noch unbekannt.

Ein weiteres identifiziertes Protein, für das vorher keine RNA Bindeaktivität beschrieben wurde, war das mitochondriale NAD(P)⁺ abhängige Malatenzym. Das Protein konnte nicht heterolog in *E. coli* hergestellt werden, da es nicht löslich exprimiert wurde. Malatenzyme katalysieren die oxidative Decarboxylierung von Malat zu Pyruvat mit NAD⁺ oder NADP⁺ als Cofaktor. Im Menschen gibt es drei Isoformen des Malatenzyms, die verschiedene Cofaktoren nutzen und unterschiedlich lokalisiert sind, das cytosolische NADP⁺ abhängige Malatenzym (ME1), das mitochondriale NAD(P)⁺ abhängige Malatenzym (ME2) und das mitochondriale NADP⁺ abhängige Malatenzym (ME3). ME1 und ME2 sind die Hauptformen. Malat ist ein Intermediat des Citratcyclus, die oxidative Decarboxylierung zu Pyruvat könnte einen regulatorischen Einfluss auf den Citratcyclus haben und ihn so dem Energiestatus der Zelle anpassen. Des Weiteren können Vorstufen und Reduktionsequivalente bereitgestellt werden. Auch aufgrund der Beteiligung am Energiehaushalt der Zelle wird ME2 mit dem Tumorstoffwechsel assoziiert. Es wurde gezeigt, dass der Tumorsuppressor p53 die Expression aller drei Malatenzyme inhibiert, indem er direkt an *response elemente* in deren Genen bindet (143). Eine Herunterregulation von ME1 oder ME2 induzierte p53 Aktivierung und Seneszenz, während eine Überexpression die entgegengesetzte Wirkung hatte. So bilden p53 und die Malatenzyme einen negativen Feedback-Loop. Die Expression von ME2 nimmt während der Progression von Melanomen zu (144). Ein *knockdown* konnte die Proliferation von Melanomzelllinien hemmen. Des Weiteren wurde ME2 mit der Migration und Invasion von Tumorzellen assoziiert, so führte eine Herunterregulation von ME2 zu reduzierter Invasivität. Dies ist besonders interessant, da MMP16 direkt an der Invasivität von Tumorzellen beteiligt ist. Zu klären wäre, ob ME2 und die MMP16 mRNA in der Zelle überhaupt in Kontakt treten können, da ME2 ein mitochondriales Enzym ist und, soweit bekannt, in der mitochondrialen Matrix vorliegt. Das Protein ist zwar im Zellkern codiert und wird dann im Cytoplasma translatiert, mitochondriale Proteine werden jedoch in der Regel ungefaltete in das Mitochondrium importiert und erst dort zum funktionalen Protein gefaltet (145).

In den Pull-down-Assays wurden für die beiden untersuchten G-Quadruplexe Proteine identifiziert, die spezifisch und mit hoher Affinität an die Strukturen binden konnten. Das Protein U2AF65 wurde sogar

als Binder beider G-Quadruplexe gefunden und schien auch für weitere G-Quadruplex-Strukturen eine hohe Affinität zu haben. Dieses Protein könnte somit ein allgemeines G-Quadruplex bindendes Protein sein. Es wurden Bindungskonstanten im nanomolaren Bereich bestimmt, was auf biologisch relevante Interaktionen hindeutet. Die G-Quadruplex-Sequenzen wurden nach ihrem Einfluss auf die Translation ausgewählt, um Proteine zu identifizieren, die hier eine regulatorische Rolle spielen. Die identifizierten Binder sind jedoch an verschiedenen Prozessen des mRNA Metabolismus beteiligt, wobei kaum eines der Proteine direkt mit der Translationsregulation assoziiert ist. Diese Ergebnisse deuten auf die vielfältigen Funktionen der identifizierten Proteine hin, die durch ein G-Quadruplex ausgeübt werden können. Für diese Funktionen sind aus der Literatur schon viele Beispiele bekannt (Abschnitt 1.2.3, (41, 100)). G-Quadruplexe scheinen in Nucleinsäuren ein Erkennungsmotiv für Proteine zu bilden. Diese Interaktion ist bereits aus den Wechselwirkungen zwischen Aptameren und Proteinen bekannt. Aptamere sind kurze Nucleinsäurestränge, die eine definierte dreidimensionale Struktur einnehmen und spezifisch an ein Zielmolekül binden. Sowohl DNA als auch RNA Aptamere, die durch *in vitro* Selektion gewonnen werden, bilden oft G-Quadruplex-Strukturen (74, 146). In den Translationsassays, sowohl *in vitro* als auch *in cellulo*, konnte für die untersuchten Proteine kein Einfluss auf die Translation festgestellt werden. Dies deutet darauf hin, dass die Bindung entweder eine andere Funktion hat, oder die Funktion indirekt über Aktivierung weiterer Proteine ausgeübt wird.

Ein Nachteil der verwendeten gelbasierten Methode ist, dass nur Proteine selektiert werden, die auch in großer Menge in der Zelle vorkommen. Proteine, die nur wenig vorkommen, sind auf dem Gel schlecht zu erkennen oder werden durch andere Proteine überlagert. Auch Proteine wie Helikasen, von denen man erwarten würde, dass sie G-Quadruplexe binden und auflösen, um so z. B. die Translation möglich zu machen, könnten hier übersehen werden. Es ist vorstellbar, dass sie auch in den Pull-down-Assays die G-Quadruplex-Struktur auflösen und anschließend abdissoziieren und deshalb nicht in großen Mengen gefunden werden. Die Kombination aus Isotopen-markierten Aminosäuren (SILAC) und der Analyse mittels Orbitrap ist in dieser Hinsicht klar überlegen. Hier waren für das MMP16 G-Quadruplex unter den ersten 20 Proteinen zwei Helikasen zu finden (Tabelle 17). Ein weiterer Vorteil der Methode ist, dass die Messung quantitativ ist und man direkt die Verhältnisse für die Bindung an Zielsequenz und an Kontrolle bilden kann. Die hohe Sensitivität kann jedoch auch schnell zu falschen Ergebnissen führen und die erhaltenen Datensätze sind umfangreich und verlangen eine gewissenhafte Auswertung. Des Weiteren sind noch Wiederholungen des Experiments nötig.

5.4. ARPC2 SPLEIßVARIANTEN

Von der ARPC2 mRNA gibt es zwei Spleißvarianten, die sich nur in ihrer 5'-UTR unterscheiden (4.9.1). Die eine ist etwas länger und enthält das G-Quadruplex, die kürzere enthält kein G-Quadruplex. Das Vorkommen beider Spleißvarianten in Hek293 Zellen konnte mittels RT-PCR gezeigt werden. Die 5'-

UTRs hatten einen unterschiedlichen Einfluss auf die Translation. Die G-Quadruplex-Variante inhibierte die Translation, wie bereits beschrieben, um ca. 50 %, während die kürzere Variante auf die Translation des Reporterproteins im Vergleich zur Kontrolle keinen Einfluss hatte. Mittels RT-qPCR wurden die Verhältnisse der beiden UTRs in Zellen bestimmt, die ohne weitere Behandlung verschieden lang kultiviert wurden. In der Mammakarzinom Zelllinie MCF7 zeigte sich dabei, dass die G-Quadruplex-Variante nach vier Tagen Kultivierung um ca. das Fünffache hochreguliert wurde. Auch in den anderen Zelllinien wurde das Verhältnis von G-Quadruplex- zu kurzer UTR größer, jedoch weil die kurze Variante zurückging. Dieses Ergebnis zu erklären ist schwierig, da die Zellen keinen wirklich definierten Bedingungen ausgesetzt waren. Vier Tage Kultivierung ohne Wechsel des Mediums lösen wahrscheinlich eine Stressantwort in den Zellen aus. Dass eine GC-reiche 5'-UTR durch alternatives Spleißen verkürzt wird, um erhöhte Proteinlevel zu erhalten, ist bereits beschrieben worden (147). Dass die Zelle jedoch unter diesen Mangelbedingungen eine mRNA synthetisiert, die wegen des G-Quadruplexes nur zu 50 % translatiert wird, ist unwahrscheinlich. Es könnte auch einfach weniger mRNA gebildet werden. Diese Hochregulation könnte jedoch erklärt werden, wenn die UTR-Variante mit G-Quadruplex tatsächlich ein IRES-Element enthielte. So werden unter Stressbedingungen nur bestimmte mRNAs translatiert, die in der entsprechenden Situation wichtig für das Überleben der Zelle sind (45). Wie bereits in Abschnitt 1.4.2 beschrieben, ist der ARP2/3-Komplex an vielen Prozessen beteiligt, die je nach Umgebungsbedingungen essentiell für die Zelle sein könnten (84). So wäre es denkbar, dass für die hier untersuchte Tumorzelllinie die Fähigkeit zur Migration von Bedeutung für ihr Überleben ist.

5.5. G-QUADRUPELLE IM ARP2/3-KOMPLEX

Das Protein ARPC2, dessen 5'-UTR G-Quadruplex hier untersucht wurde, ist Teil eines Proteinkomplexes, der aus 7 Untereinheiten besteht. Von diesen 7 Untereinheiten besitzen 6 eine G-Quadruplex-Sequenz in ihrer 5'- oder in ihrer 3'-UTR. Dies ist eine erstaunliche Anhäufung, betrachtet man das Vorkommen von G-Quadruplex-Sequenzen in den untranslatierten Regionen von mRNAs, von denen ca. 17 % ein 5'- oder 3'-UTR G-Quadruplex besitzen (32). Es wäre möglich, dass diese Häufung mit der gemeinsamen Funktion der Proteine zusammenhängt. Interessant ist, dass drei Proteine eine G-Quadruplex-Sequenz in ihrer 5'-UTR besitzen und drei in der 3'-UTR. Es stellte sich die Frage, ob diese G-Quadruplex-Sequenzen unterschiedliche Funktionen erfüllen. Um sich dieser Frage zu nähern, wurde der Einfluss der Sequenzen auf die Translation untersucht. Jede der Sequenzen wurde an seine ursprüngliche Position, also in die 5'- oder 3'-UTR, eines Reportergens kloniert und auf diese Weise wurde ihr Einfluss auf die Translation bestimmt (Abbildung 40). Der Einfluss auf die Transkription wurde durch RT-qPCR ermittelt. Alle G-Quadruplex-Sequenzen führten zu einer Hemmung der Translation. Der mRNA Gehalt blieb konstant oder war eher noch etwas höher als bei der Kontrolle.

Auffällig war, dass die Inhibition der Translation durch die 3'-UTR G-Quadruplexe insgesamt stärker war als durch die 5'-UTR G-Quadruplexe. Es wurde bereits gezeigt, dass ein 3'-UTR G-Quadruplex die Translation inhibieren kann (1.2.3). Das G-Quadruplex in der 3'-UTR des Protoonkogens Pim-1 hemmte die Translation um etwa 50 % (48). Bedenkt man, dass die mRNA, während sie abgelesen wird, durch Translationsfaktoren zirkularisiert werden kann, ist es vorstellbar, dass auch ein stabiles Element in der 3'-UTR inhibierend wirkt, indem es das Ribosom blockiert. Hier wäre es interessant zu testen, ob dieselbe G-Quadruplex-Sequenz in der 5'-UTR auch den gleichen Einfluss auf die Translation hat oder ob die Inhibition von der Position abhängt. Die eigentliche Funktion dieser Elemente auf die ARP2/3-Komplex mRNAs bleibt dennoch unklar. Eine denkbare Erklärung wäre, dass sie einen schnellen Eingriff in die Translation erlauben und damit eine weitere Ebene zur Regulation der Genexpression bilden. Nötig wären hierzu wieder Proteine, die durch ihr Binden einen aktivierenden oder hemmenden Einfluss auf die Translation ausüben. In diesem Zusammenhang ist auch interessant, dass alle 7 ARP2/3 mRNAs an den Wirkungsort des Komplexes transportiert und dort lokal translatiert werden (89). Hier könnte der Zeitfaktor eine Rolle spielen. Wird mehr Protein gebraucht, wäre es wahrscheinlich schneller, die Translation einer mRNA, die bereits am Wirkungsort vorliegt, hochzuregulieren, als neue mRNAs zu synthetisieren und diese dann zu transportieren. Umgekehrt könnte die Proteinsynthese gehemmt werden, wenn der Komplex gerade nicht gebraucht wird. Eine schnelle Regulation könnte gerade bei Prozessen, die von extrazellulären Stimuli abhängen, wie der Migration, der Endozytose oder der Ausbildung von Zell-Zell-Kontakten, nützlich sein. In diesem Hinblick sollte untersucht werden, welche Proteine an die ARP2/3 G-Quadruplexe binden und ob sie eventuell von denselben Proteinen erkannt werden. Dies würde auf einen gemeinsamen Regulationsmechanismus hindeuten, der die Expression der einzelnen Untereinheiten beeinflusst.

5.6. AUSBLICK

Ein offener Punkt bleibt die Funktion der identifizierten G-Quadruplex bindenden Proteine. Die untersuchten Proteine wirkten nicht direkt auf die Translation, sie schienen also nicht die Ausbildung oder Auflösung der G-Quadruplex-Struktur zu beeinflussen. Daher wäre es interessant zu untersuchen, ob durch die Interaktion andere Prozesse beeinflusst werden. Viele der identifizierten Proteine waren am Spleißen beteiligt, es wäre also möglich, dass die G-Quadruplexe und die daran bindenden Proteine hier eine Rolle spielen. Um dies zu untersuchen könnten Minigen-Spleiß-Assays durchgeführt werden, in denen das Spleißen eines Reportergens durch eine inserierte Sequenz beeinflusst wird (148). Besonders in Bezug auf die beiden ARPC2 5'-UTR-Varianten könnten die Spleißfaktoren eine Rolle spielen. Möglicherweise sind sie an der Regulation des Spleißens der beiden Varianten beteiligt. Interessant wäre, ob eine Überexpression oder ein *Knockdown* eines dieser Proteine das Verhältnis der beiden UTR-Varianten beeinflussen kann.

In Bezug auf die Translation könnte man noch weitere Kandidaten untersuchen. Vor allem im *in vitro* Translationsassay wurden nicht alle identifizierten Proteine getestet, was an Schwierigkeiten bei der heterologen Proteinexpression in *E. coli* lag. Auch die Auswirkungen eines *Knockdowns* der Proteine mittels RNA Interferenz auf die Luciferase-Reporter-Plasmide könnten untersucht werden. Dies wäre vor allem in Kombination mit dem SILAC-Pull-down-Assay nützlich, da hier Proteinkandidaten in großem Maßstab parallel untersucht werden könnten. Der SILAC-Pull-down-Assay könnte auch sehr hilfreich sein, um weitere G-Quadruplex interagierende Proteine zu identifizieren, auch solche, die nicht in großen Mengen in der Zelle vorhanden sind. Diese Ergebnisse sollten unbedingt bestätigt werden.

Für die Expression der beiden ARPC2 5'-UTR-Varianten wurde in MCF7 Zellen eine Regulation gefunden. Für diese Regulation müsste der genaue Auslöser und damit die beteiligte Signalkaskade gefunden werden. Am vielversprechendsten wäre es auch hier einen proteomischen Ansatz zu wählen. So könnte mittels quantitativer Massenspektroskopie die Proteinexpressionen der beiden Zellpopulationen verglichen werden. Über Unterschiede in der Proteinexpression könnte man auf die beteiligten Signalwege schließen und so den Auslöser ausmachen.

Des Weiteren fanden sich für das ARPC2 G-Quadruplex, viele ribosomale Proteine als Binder. Eine Erklärung hierfür wäre, dass das G-Quadruplex Teil eines IRES Elements sein könnte und deshalb vom Ribosom erkannt wird. Dies könnte man durch ein bicistronisches Transkript überprüfen, in dem die potentielle IRES vor der zweiten codierenden Sequenz inseriert wird. Wird das zweite Protein exprimiert, muss es sich um eine IRES Sequenz handeln.

Bei der Untersuchung der G-Quadruplexe des Arp2/3 Komplexes wurde für alle Sequenzen eine Inhibition der Translation gefunden. Die G-Quadruplexe der 3'-UTR hatten insgesamt jedoch einen etwas höheren Effekt. Hier wäre es interessant herauszufinden, ob dieser Effekt an der Position oder an der Stabilität der Sequenzen liegt, wozu der Einfluss derselben G-Quadruplexe an beiden Positionen untersucht werden müsste. Des Weiteren stellt sich die Frage, ob alle G-Quadruplexe, in der 5'-UTR und in der 3'-UTR, die gleiche Funktion erfüllen oder ob sie verschiedene Aufgaben haben. Hierfür könnte eine Identifikation der Bindungspartner nützlich sein. Binden dieselben Proteine an die unterschiedlichen Sequenzen, könnte dies auf eine gemeinsame Funktion hindeuten.

6. ZUSAMMENFASSUNG

Ziel dieser Arbeit war es, Proteine zu identifizieren, die spezifisch an RNA G-Quadruplexe binden. G-Quadruplexe sind ungewöhnliche Sekundärstrukturen, die von guaninreichen Nucleinsäuren gebildet werden. In der 5'-UTR von mRNAs können sie die Translation inhibieren. Als Modellsequenzen wurden die RNA G-Quadruplexe aus der 5'-UTR von ARPC2 (*actin related protein 2/3 complex subunit 2*) und MMP16 (*matrix metalloproteinase 16*) ausgewählt. Diese Sequenzen waren hoch konserviert und ihre Expression konnte in der humanen Nierenzelllinie Hek293 nachgewiesen werden. Das ARPC2 G-Quadruplex wurde in dieser Arbeit erstmals beschrieben. Die G-reichen Sequenzen wurden zunächst mit spektroskopischen Methoden hinsichtlich der Ausbildung eines G-Quadruplexes, sowie auf dessen Stabilität untersucht. Der Einfluss der G-Quadruplexe auf die Translation wurde mittels Dual-Luciferase-Assays bestimmt. Das ARPC2 G-Quadruplex hemmte die Translation um etwa 40 %, das MMP16 G-Quadruplex sogar um ca. 50 %. Um G-Quadruplex bindende Proteine zu identifizieren, wurden Pull-down-Assays mit Hek293 und HeLa Zelllysats durchgeführt. Es wurden eine Reihe spezifisch bindender Proteine gefunden, von denen viele zu den ribosomalen Proteinen und zu den Spleißfaktoren gehörten. Für das ARPC2 G-Quadruplex wurden deutlich mehr Proteine identifiziert als für das MMP16 G-Quadruplex. Die Bindungen zwischen einigen dieser Proteine und den G-Quadruplexen wurden genauer charakterisiert. Mittels Oberflächenplasmonresonanz-Spektroskopie konnten Bindungskonstanten im nanomolaren Bereich bestimmt werden, die ein Indiz für biologisch relevante Interaktionen liefern. Ein direkter Einfluss der Proteine auf die Translation konnte weder in Co-Transfektionsexperimenten, noch in *in vitro* Translationsassays gezeigt werden. Entweder ist der Einfluss der Proteine auf die Translation indirekt oder deren Bindung an die G-Quadruplex-Sequenzen hat eine andere Funktion. Durch das Verwenden eines sensitiveren Orbitrap Massenspektrometers war es möglich, die vorherigen Ergebnisse zu bestätigen und noch weitere Proteine zu detektieren. Von der ARPC2 5'-UTR wurden zwei Spleißvarianten gefunden, von denen nur eine das G-Quadruplex enthält. Es konnte gezeigt werden, dass sie einen unterschiedlichen Einfluss auf die Translation eines Reportergens haben. Die G-Quadruplex-Variante inhibierte die Translation, während die zweite, kürzere Variante keinen Einfluss hatte. Beide Varianten wurden in MCF7, Hek293 und HeLa Zellen exprimiert. Durch Induktion von Stress auf die MCF7 Zellen wurde die G-Quadruplex-Variante deutlich hochreguliert.

ARPC2 ist eine von 7 Untereinheiten des ARP2/3-Komplexes. Erstaunlicherweise besitzen sechs der Untereinheiten ein G-Quadruplex in ihrer 5'- oder ihrer 3'-UTR. Diese Überrepräsentation von G-Quadruplex-Sequenzen deutet auf eine gemeinsame Regulation der Untereinheiten hin. Der Einfluss dieser G-Quadruplexe auf die Translation wurde in Dual-Luciferase-Reporter-Assays untersucht. Alle G-Quadruplex-Sequenzen hemmten die Translation, wobei die 3'-UTR G-Quadruplexe einen viel stärkeren Einfluss hatten als die 5'-UTR G-Quadruplexe.

7. SUMMARY

G-quadruplexes are noncanonical secondary structures formed by Guanine-rich nucleic acids. They can serve as regulatory elements affecting different levels of gene expression. The present study focused on the identification of proteins binding RNA G-quadruplexes in 5'-UTRs of mRNAs, which are able to inhibit translation of the respective mRNA. So far, little is known about trans-acting factors, interacting with RNA G-quadruplexes and thereby influencing their formation or function. To identify G-quadruplex binding proteins, the G-rich sequences of the 5'-UTRs of matrix metalloproteinase 16 (MMP16) and actin related protein 2/3 complex subunit 2 (ARPC2) were chosen. Both sequences were highly conserved between different species. Their expression in the investigated Hek293 cell line was shown by RT-PCR. Formation of a G-quadruplex was already described for the MMP16 sequence, while the ARPC2 G-quadruplex was for the first time described in this work. Biophysical characterization of the G-quadruplexes was performed using CD-spectroscopy and UV-melting experiments. Both sequences were able to inhibit translation of a reporter gene. To identify G-quadruplex binding proteins, pull-down assays followed by MALDI-TOF mass spectroscopy were performed. Many of the identified proteins belonged to the groups of ribosomal proteins and splicing factors. In total, more proteins were binding to the ARPC2 than to the MMP16 G-quadruplex. The interaction of selected proteins with the G-quadruplexes was further analysed by surface plasmon resonance spectroscopy. Binding constants at nanomolar level were determined, indicating a biological relevance of the interactions. A direct influence of the identified proteins on translation could neither be shown in co-transfections with reporter plasmids nor in *in vitro* translation assays. This suggests they may influence translation indirectly, via activation of other factors, or they do not function on translation but rather on another process. The results have been confirmed using a more sensitive mass spectrometry method. In addition, this method identified a lot more potential G-quadruplex binding proteins.

There are two variants of the ARPC2 5'-UTR. A longer form containing the G-quadruplex and a shortened isoform lacking it. Their influence on translation was investigated in a luciferase reporter assay. While the G-quadruplex UTR inhibited translation of the reporter protein, the shorter variant had no influence on protein expression compared to the control. Expression of both variants could be shown in Hek293, MCF7 and HeLa cell lines. Interestingly, in MCF7 cells expression of the G-quadruplex variant was highly upregulated after cultivation for 96 h, without changing cultivation media, which is likely to represent stress conditions.

ARPC2 is a subunit of the ARP2/3 complex, consisting of seven proteins. Six of this subunits contain a G-quadruplex in either their 5'- or their 3'-UTR. This accumulation suggests a common regulatory mechanism for the proteins. Influence on translation of this sequences was analysed by dual-luciferase assays. While all G-quadruplex sequences were able to inhibit translation, influence of the 3'-UTR sequences on translation was higher.

8. LITERATURVERZEICHNIS

1. Bang, I. (1910) Untersuchungen über die Guanylsäure. *Biochem Z.*
2. Gellert, M., Lipsett, M.N. and Davies, D.R. (1962) HELIX FORMATION BY GUANYLIC ACID. *Proc. Natl. Acad. Sci. U. S. A.*, **48**, 2013–2018.
3. Henderson, E., Hardin, C.C., Walk, S.K., Tinoco, I. and Blackburn, E.H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell*, **51**, 899–908.
4. Huppert, J.L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.*, **37**, 1375–84.
5. Williamson, J.R., Raghuraman, M.K. and Cech, T.R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, **59**, 871–80.
6. Chen, Y. and Yang, D. (2012) Sequence, stability, and structure of G-Quadruplexes and their interactions with drugs. *Curr. Protoc. Nucleic Acid Chem.*, 10.1002/0471142700.nc1705s50.
7. Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. and Neidle, S. (2006) Quadruplex DNA: Sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
8. Guédin, A., Gros, J., Alberti, P. and Mergny, J.L. (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.*, **38**, 7858–7868.
9. Joachimi, A., Benz, A. and Hartig, J.S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem.*, **17**, 6811–6815.
10. Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L. and Wu, J.R. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.*, **85**, 6622–6.
11. Phan, a. T. (2010) Human telomeric G-quadruplex: Structures of DNA and RNA sequences. *FEBS J.*, **277**, 1107–1117.
12. Vorlíčková, M., Kejnovská, I., Sagi, J., Renčiuk, D., Bednářová, K., Motlová, J. and Kypr, J. (2012) Circular dichroism and guanine quadruplexes. *Methods*, **57**, 64–75.
13. Mergny, J.L., Phan, a T. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74–8.
14. Hsu, S.-T.D., Varnai, P., Bugaut, A., Reszka, A.P., Neidle, S. and Balasubramanian, S. (2009) A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc.*, **131**, 13399–409.
15. Dai, J., Chen, D., Jones, R.A., Hurley, L.H. and Yang, D. (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.*, **34**, 5133–44.

16. Ambrus, A., Chen, D., Dai, J., Jones, R.A. and Yang, D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry*, **44**, 2048–58.
17. Agrawal, P., Hatzakis, E., Guo, K., Carver, M. and Yang, D. (2013) Solution structure of the major G-quadruplex formed in the human VEGF promoter in K⁺: insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res.*, **41**, 10584–92.
18. Parkinson, G.N., Ghosh, R. and Neidle, S. (2007) Structural basis for binding of porphyrin to human telomeres. *Biochemistry*, **46**, 2390–7.
19. Parkinson, G.N., Lee, M.P.H. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–80.
20. Collie, G.W., Haider, S.M., Neidle, S. and Parkinson, G.N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.*, **38**, 5569–80.
21. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H.J. and Plückthun, A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylomychia lemnae* macronuclei. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 8572–7.
22. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–5.
23. Gros, J., Guédin, A., Mergny, J.-L. and Lacroix, L. (2008) G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC. *Chembiochem*, **9**, 2075–9.
24. Lattmann, S., Stadler, M.B., Vaughn, J.P., Akman, S.A. and Nagamine, Y. (2011) The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic Acids Res.*, **39**, 9390–404.
25. Redon, S., Reichenbach, P. and Lingner, J. (2010) The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.*, **38**, 5797–5806.
26. Hirashima, K. and Seimiya, H. (2015) Telomeric repeat-containing RNA/G-quadruplex-forming sequences cause genome-wide alteration of gene expression in human cancer cells in vivo. *Nucleic Acids Res.*, 10.1093/nar/gkv063.
27. Biffi, G., Tannahill, D., McCafferty, J. and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–6.
28. Biffi, G., Di Antonio, M., Tannahill, D. and Balasubramanian, S. (2013) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.*, **6**, 75–80.
29. Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, **35**, 406–413.
30. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 11593–11598.

31. Eddy, J. and Maizels, N. (2006) Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.*, **34**, 3887–3896.
32. Huppert, J.L., Bugaut, A., Kumari, S. and Balasubramanian, S. (2008) G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.*, **36**, 6260–8.
33. Kumari, S., Bugaut, A., Huppert, J.L. and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.*, **3**, 218–21.
34. Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S. and Kurreck, J. (2008) Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA*, **14**, 1290–6.
35. Lammich, S., Kamp, F., Wagner, J., Nuscher, B., Zilow, S., Ludwig, A.-K., Willem, M. and Haass, C. (2011) Translational repression of the Disintegrin and Metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5'-untranslated region. *J. Biol. Chem.*, **286**, 45063–45072.
36. Shahid, R., Bugaut, A. and Balasubramanian, S. (2010) The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression. *Biochemistry*, **49**, 8300–6.
37. Weng, H.-Y., Huang, H.-L., Zhao, P.-P., Zhou, H. and Qu, L.-H. (2012) Translational repression of cyclin D3 by a stable G-quadruplex in its 5' UTR: implications for cell cycle regulation. *RNA Biol.*, **9**, 1099–109.
38. Balkwill, G.D., Derecka, K., Garner, T.P., Hodgman, C., Flint, A.P.F. and Searle, M.S. (2009) Repression of translation of human estrogen receptor alpha by G-quadruplex formation. *Biochemistry*, **48**, 11487–95.
39. Morris, M.J. and Basu, S. (2009) An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry*, **48**, 5313–5319.
40. Preiss, T. and Hentze, M.W. (2003) Starting the protein synthesis machine: Eukaryotic translation initiation. *BioEssays*, **25**, 1201–1211.
41. Millevoi, S., Moine, H. and Vagner, S. (2012) G-quadruplexes in RNA biology. *Wiley Interdiscip. Rev. RNA*, **3**, 495–507.
42. Gebauer, F. and Hentze, M.W. (2004) Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell Biol.*, **5**, 827–35.
43. Kumari, S., Bugaut, A. and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene. *Biochemistry*, **47**, 12664–9.
44. Halder, K., Wieland, M. and Hartig, J.S. (2009) Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes. *Nucleic Acids Res.*, **37**, 6811–7.
45. Komar, A. a. and Hatzoglou, M. (2011) Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle*, **10**, 229–240.
46. Bonnal, S., Schaeffer, C., Créancier, L., Clamens, S., Moine, H., Prats, A.-C. and Vagner, S. (2003) A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth

- factor 2 gene expression at four alternative translation initiation codons. *J. Biol. Chem.*, **278**, 39330–6.
47. Morris, M.J., Negishi, Y., Pazsint, C., Schonhoft, J.D. and Basu, S. (2010) An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *J. Am. Chem. Soc.*, **132**, 17831–9.
48. Arora, A. and Suess, B. (2011) An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation. *RNA Biol.*, **8**, 802–805.
49. Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J.-L. and Moine, H. (2011) G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.*, **12**, 697–704.
50. Andreassi, C. and Riccio, A. (2009) To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol.*, **19**, 465–74.
51. Besse, F. and Ephrussi, A. (2008) Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nat. Rev. Mol. Cell Biol.*, **9**, 971–80.
52. Martin, K.C. and Ephrussi, A. (2009) mRNA localization: gene expression in the spatial dimension. *Cell*, **136**, 719–30.
53. Beaudoin, J.-D. and Perreault, J.-P. (2013) Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. *Nucleic Acids Res.*, **41**, 5898–911.
54. Murat, P., Zhong, J., Lekieffre, L., Cowieson, N.P., Clancy, J.L., Preiss, T., Balasubramanian, S., Khanna, R. and Tellam, J. (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.*, **10**, 358–64.
55. Metifiot, M., Amrane, S., Litvak, S. and Andreola, M.-L. (2014) G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res.*, **42**, 12352–12366.
56. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–5.
57. Nilsen, T.W. and Graveley, B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
58. Shepard, P.J. and Hertel, K.J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA*, **14**, 1463–1469.
59. Didiot, M.C., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J.L. and Moine, H. (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Res.*, **36**, 4902–4912.
60. Gomez, D., Lemarteleur, T., Lacroix, L., Mailliet, P., Mergny, J.-L. and Riou, J.-F. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.*, **32**, 371–9.
61. Fisette, J.F., Montagna, D.R., Mihailescu, M.R. and Wolfe, M.S. (2012) A G-Rich element forms a G-quadruplex and regulates BACE1 mRNA alternative splicing. *J. Neurochem.*, **121**, 763–773.

62. González, V., Guo, K., Hurley, L. and Sun, D. (2009) Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *J. Biol. Chem.*, **284**, 23622–35.
63. Dempsey, L. a., Sun, H., Hanakahi, L. a. and Maizels, N. (1999) G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, a role for G-G pairing in immunoglobulin switch recombination. *J. Biol. Chem.*, **274**, 1066–1071.
64. Etzioni, S., Yafe, A., Khateb, S., Weisman-Shomer, P., Bengal, E. and Fry, M. (2005) Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. *J. Biol. Chem.*, **280**, 26805–12.
65. Walsh, K. and Gualberto, A. (1992) MyoD binds to the guanine tetrad nucleic acid structure. *J. Biol. Chem.*, **267**, 13714–8.
66. Schaeffer, C., Bardoni, B., Mandel, J.L., Ehresmann, B., Ehresmann, C. and Moine, H. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J.*, **20**, 4803–13.
67. Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T. and Darnell, R.B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell*, **107**, 489–99.
68. Paramasivam, M., Membrino, A., Cogoi, S., Fukuda, H., Nakagama, H. and Xodo, L.E. (2009) Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription. *Nucleic Acids Res.*, **37**, 2841–53.
69. Khateb, S., Weisman-Shomer, P., Hershco, I., Loeb, L. a and Fry, M. (2004) Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)_n is mediated by homolog-conserved domains in three members of the hnRNP family. *Nucleic Acids Res.*, **32**, 4145–54.
70. Zaug, A.J., Podell, E.R. and Cech, T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci.*, **102**, 10864–10869.
71. Creacy, S.D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A. and Vaughn, J.P. (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.*, **283**, 34626–34.
72. Kamath-Loeb, A.S., Loeb, L.A., Johansson, E., Burgers, P.M. and Fry, M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)_n trinucleotide repeat sequence. *J. Biol. Chem.*, **276**, 16439–46.
73. Sun, H., Karow, J.K., Hickson, I.D. and Maizels, N. (1998) The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.*, **273**, 27587–27592.
74. Sissi, C., Gatto, B. and Palumbo, M. (2011) The evolving world of protein-G-quadruplex recognition: A medicinal chemist's perspective. *Biochimie*, **93**, 1219–1230.
75. Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., et al. (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell*, **107**, 477–87.

76. Lu, R., Wang, H., Liang, Z., Ku, L., O'donnell, W.T., Li, W., Warren, S.T. and Feng, Y. (2004) The fragile X protein controls microtubule-associated protein 1B translation and microtubule stability in brain neuron development. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 15201–15206.
77. Castets, M., Schaeffer, C., Bechara, E., Schenck, A., Khandjian, E.W., Luche, S., Moine, H., Rabilloud, T., Mandel, J.L. and Bardoni, B. (2005) FMRP interferes with the Rac1 pathway and controls actin cytoskeleton dynamics in murine fibroblasts. *Hum. Mol. Genet.*, **14**, 835–844.
78. Rodríguez, D., Morrison, C.J. and Overall, C.M. (2010) Matrix metalloproteinases: What do they not do? New substrates and biological roles identified by murine models and proteomics. *Biochim. Biophys. Acta - Mol. Cell Res.*, **1803**, 39–54.
79. Kitagawa, Y., Kunimi, K., Uchibayashi, T., Sato, H. and Namiki, M. (1999) Expression of messenger RNAs for membrane-type 1, 2, and 3 matrix metalloproteinases in human renal cell carcinomas. *J. Urol.*, **162**, 905–9.
80. Jung, M., Römer, A., Keyszer, G., Lein, M., Kristiansen, G., Schnorr, D., Loening, S.A. and Jung, K. (2003) mRNA expression of the five membrane-type matrix metalloproteinases MT1-MT5 in human prostatic cell lines and their down-regulation in human malignant prostatic tissue. *Prostate*, **55**, 89–98.
81. Tatti, O., Arjama, M., Ranki, A., Weiss, S.J., Keski-Oja, J. and Lehti, K. (2011) Membrane-type-3 matrix metalloproteinase (MT3-MMP) functions as a matrix composition-dependent effector of melanoma cell invasion. *PLoS One*, **6**, 19–22.
82. Hotary, K., Li, X.Y., Allen, E., Stevens, S.L. and Weiss, S.J. (2006) A cancer cell metalloprotease triad regulates the basement membrane transmigration program. *Genes Dev.*, **20**, 2673–2686.
83. Morris, M.J., Wingate, K.L., Silwal, J., Leeper, T.C. and Basu, S. (2012) The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells. *Nucleic Acids Res.*, **40**, 4137–4145.
84. Rotty, J.D., Wu, C. and Bear, J.E. (2013) New insights into the regulation and cellular functions of the ARP2/3 complex. *Nat. Rev. Mol. Cell Biol.*, **14**, 7–12.
85. Bailly, M., Ichetovkin, I., Grant, W., Zebda, N., Machesky, L.M., Segall, J.E. and Condeelis, J. (2001) The F-actin side binding activity of the Arp2/3 complex is essential for actin nucleation and lamellipod extension. *Curr. Biol.*, **11**, 620–5.
86. Rauhala, H.E., Teppo, S., Niemelä, S. and Kallioniemi, A. (2013) Silencing of the ARP2/3 complex disturbs pancreatic cancer cell migration. *Anticancer Res.*, **33**, 45–52.
87. Steffen, A., Faix, J., Resch, G.P., Linkner, J., Wehland, J., Small, J.V., Rottner, K. and Stradal, T.E.B. (2006) Filopodia formation in the absence of functional WAVE- and Arp2/3-complexes. *Mol. Biol. Cell*, **17**, 2581–91.
88. Di Nardo, A., Cicchetti, G., Falet, H., Hartwig, J.H., Stossel, T.P. and Kwiatkowski, D.J. (2005) Arp2/3 complex-deficient mouse fibroblasts are viable and have normal leading-edge actin structure and function. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 16263–8.

89. Mingle, L. a, Okuhama, N.N., Shi, J., Singer, R.H., Condeelis, J. and Liu, G. (2005) Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. *J. Cell Sci.*, **118**, 2425–33.
90. Liao, G., Simone, B. and Liu, G. (2011) Mis-localization of Arp2 mRNA impairs persistence of directional cell migration. *Exp. Cell Res.*, **317**, 812–22.
91. Blagoev, B., Ong, S.-E., Kratchmarova, I. and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.*, **22**, 1139–1145.
92. Šidák, Z. (1967) Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *J. Am. Stat. Assoc.*, **62**, 626–633.
93. Kikin, O., D’Antonio, L. and Bagga, P.S. (2006) QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–82.
94. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
95. Ginisty, H., Amalric, F. and Bouvet, P. (1998) Nucleolin functions in the first step of ribosomal RNA processing. *EMBO J.*, **17**, 1476–1486.
96. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–9.
97. Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, e45.
98. Lang, R., Braun, M., Sounni, N.E., Noel, a., Frankenne, F., Foidart, J.M., Bode, W. and Maskos, K. (2004) Crystal Structure of the Catalytic Domain of MMP-16/MT3-MMP: Characterization of MT-MMP Specific Features. *J. Mol. Biol.*, **336**, 213–225.
99. Agarwala, P., Pandey, S. and Maiti, S. (2014) Role of G-quadruplex located at 5′ end of mRNAs. *Biochim. Biophys. Acta - Gen. Subj.*, **1840**, 3503–3510.
100. Agarwala, P., Pandey, S. and Maiti, S. (2015) The tale of RNA G-quadruplex. *Org. Biomol. Chem.*, 10.1039/C4OB02681K.
101. Yanagida, M., Shimamoto, a, Nishikawa, K., Furuichi, Y., Isobe, T. and Takahashi, N. (2001) Isolation and proteomic characterization of the major proteins of the nucleolin-binding ribonucleoprotein complexes. *Proteomics*, **1**, 1390–404.
102. Vossen, K.M. and Fried, M.G. (1997) Sequestration stabilizes lac repressor-DNA complexes during gel electrophoresis. *Anal. Biochem.*, **245**, 85–92.
103. Fried, M.G. and Bromberg, J.L. (1997) Factors that affect the stability of protein-DNA complexes during gel electrophoresis. *Electrophoresis*, **18**, 6–11.
104. Fried, M.G. and Liu, G. (1994) Molecular sequestration stabilizes CAP-DNA complexes during polyacrylamide gel electrophoresis. *Nucleic Acids Res.*, **22**, 5054–9.

105. Tajrishi, M.M., Tuteja, R. and Tuteja, N. (2011) Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun. Integr. Biol.*, **4**, 267–275.
106. Abdelmohsen, K., Tominaga, K., Lee, E.K., Srikantan, S., Kang, M.J., Kim, M.M., Selimyan, R., Martindale, J.L., Yang, X., Carrier, F., et al. (2011) Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs. *Nucleic Acids Res.*, **39**, 8513–8530.
107. Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F. and Bouvet, P. (1996) Nucleolin is a sequence-specific RNA-binding protein: characterization of targets on pre-ribosomal RNA. *J. Mol. Biol.*, **260**, 34–53.
108. Sengupta, T.K., Bandyopadhyay, S., Fernandes, D.J. and Spicer, E.K. (2004) Identification of nucleolin as an AU-rich element binding protein involved in bcl-2 mRNA stabilization. *J. Biol. Chem.*, **279**, 10855–63.
109. González, V. and Hurley, L.H. (2010) The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry*, **49**, 9706–14.
110. Abdelmohsen, K. and Gorospe, M. (2012) RNA-binding protein nucleolin in disease. *RNA Biol.*, **9**, 799–808.
111. Borer, R. a, Lehner, C.F., Eppenberger, H.M. and Nigg, E. a (1989) Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell*, **56**, 379–390.
112. Graveley, B.R., Hertel, K.J. and Maniatis, T. (2001) The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA*, **7**, 806–18.
113. Kortt, A.A., Oddie, G.W., Iliades, P., Gruen, L.C. and Hudson, P.J. (1997) Nonspecific amine immobilization of ligand can Be a potential source of error in BIAcore binding experiments and may reduce binding affinities. *Anal. Biochem.*, **253**, 103–11.
114. Zamore, P.D., Patton, J.G. and Green, M.R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, **355**, 609–14.
115. Coolidge, C.J., Seely, R.J. and Patton, J.G. (1997) Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Res.*, **25**, 888–96.
116. Berglund, J. a, Abovich, N. and Rosbash, M. (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.*, **12**, 858–67.
117. Nelson, L.D., Bender, C., Mannsperger, H., Buergy, D., Kambakamba, P., Mudduluru, G., Korf, U., Hughes, D., Van Dyke, M.W. and Allgayer, H. (2012) Triplex DNA-binding proteins are associated with clinical outcomes revealed by proteomic measurements in patients with colorectal cancer. *Mol. Cancer*, **11**, 38.
118. Long, J.C. and Cáceres, J.F. (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.*, **417**, 15–27.
119. Busch, A. and Hertel, K.J. (2012) Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA*, **3**, 1–12.
120. Shepard, P.J. and Hertel, K.J. (2009) The SR protein family. *Genome Biol.*, **10**, 242.

121. Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D. and Krainer, A.R. (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.*, **14**, 185–193.
122. Sanford, J.R., Gray, N.K., Beckmann, K. and Cáceres, J.F. (2004) A novel role for shuttling SR proteins in mRNA translation. *Genes Dev.*, **18**, 755–768.
123. Twyffels, L., Gueydan, C. and Kruys, V. (2011) Shuttling SR proteins: More than splicing factors. *FEBS J.*, **278**, 3246–3255.
124. Lemaire, R., Prasad, J., Kashima, T., Gustafson, J., Manley, J.L. and Lafyatis, R. (2002) Stability of a PKC δ -1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins. *Genes Dev.*, **16**, 594–607.
125. Sanford, J.R., Coutinho, P., Hackett, J. a., Wang, X., Ranahan, W. and Cáceres, J.F. (2008) Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. *PLoS One*, **3**.
126. Fu, Y., Huang, B., Shi, Z., Han, J., Wang, Y., Huangfu, J. and Wu, W. (2013) SRSF1 and SRSF9 RNA binding proteins promote Wnt signalling-mediated tumorigenesis by enhancing β -catenin biosynthesis. *EMBO Mol. Med.*, **5**, 737–750.
127. Han, S.P., Tang, Y.H. and Smith, R. (2010) Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.*, **430**, 379–92.
128. Matunis, M., Xing, J. and Dreyfuss, G. (1994) The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization. *Nucleic Acids Res.*, **22**, 1059–1067.
129. Fox, T.D. (2012) Mitochondrial protein synthesis, import, and assembly. *Genetics*, **192**, 1203–1234.
130. Caputi, M. and Zahler, a M. (2001) Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J. Biol. Chem.*, **276**, 43850–9.
131. Samatanga, B., Dominguez, C., Jelesarov, I. and Allain, F.H.-T. (2013) The high kinetic stability of a G-quadruplex limits hnRNP F qRRM3 binding to G-tract RNA. *Nucleic Acids Res.*, **41**, 2505–16.
132. Decorsière, A., Cayrel, A., Vagner, S. and Millevoi, S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev.*, **25**, 220–5.
133. Romig, H., Fackelmayer, F.O., Renz, a, Ramsperger, U. and Richter, a (1992) Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *EMBO J.*, **11**, 3431–3440.
134. Kiledjian, M. and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.*, **11**, 2655–2664.
135. Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y. and Fu, X.-D. (2012) Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. *Mol. Cell*, **45**, 656–68.

136. Van der Velden, A.W., van Nierop, K., Voorma, H.O. and Thomas, A.A.M. (2002) Ribosomal scanning on the highly structured insulin-like growth factor II-leader 1. *Int. J. Biochem. Cell Biol.*, **34**, 286–97.
137. Short, J.D. and Pfarr, C.M. (2002) Translational regulation of the JunD messenger RNA. *J. Biol. Chem.*, **277**, 32697–705.
138. Warner, J.R. and McIntosh, K.B. (2009) How common are extraribosomal functions of ribosomal proteins? *Mol. Cell*, **34**, 3–11.
139. Hellen, C.U.T. (2009) IRES-induced conformational changes in the ribosome and the mechanism of translation initiation by internal ribosomal entry. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **1789**, 558–570.
140. Spriggs, K. a, Bushell, M., Mitchell, S. a and Willis, a E. (2005) Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-trans-acting factors. *Cell Death Differ.*, **12**, 585–91.
141. Avramidou, a, Kroczeck, C., Lang, C., Schuh, W., Jäck, H.-M. and Mielenz, D. (2007) The novel adaptor protein Swiprosin-1 enhances BCR signals and contributes to BCR-induced apoptosis. *Cell Death Differ.*, **14**, 1936–47.
142. Kwon, M.S., Park, K.R., Kim, Y.D., Na, B.R., Kim, H.R., Choi, H.J., Piragyte, I., Jeon, H., Chung, K.H., Song, W.K., et al. (2013) Swiprosin-1 Is a Novel Actin Bundling Protein That Regulates Cell Spreading and Migration. *PLoS One*, **8**, e71626.
143. Jiang, P., Du, W., Mancuso, A., Wellen, K.E. and Yang, X. (2013) Reciprocal regulation of p53 and malic enzymes modulates metabolism and senescence. *Nature*, **493**, 689–93.
144. Chang, Y.-L., Gao, H.-W., Chiang, C.-P., Wang, W.-M., Huang, S.-M., Ku, C.-F., Liu, G.-Y. and Hung, H.-C. (2014) Human Mitochondrial NAD(P)⁺-Dependent Malic Enzyme Participates in Cutaneous Melanoma Progression and Invasion. *J. Invest. Dermatol.*, **135**, 807–815.
145. Chacinska, A., Koehler, C.M., Milenkovic, D., Lithgow, T. and Pfanner, N. (2009) Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell*, **138**, 628–644.
146. O. Tucker, W., T. Shum, K. and A. Tanner, J. (2012) G-quadruplex DNA Aptamers and their Ligands: Structure, Function and Application. *Curr. Pharm. Des.*, **18**, 2014–2026.
147. Sasahara, M., Amano, S., Sato, H., Yang, J.G., Hayase, Y., Kaneko, M., Sato, I., Suzuki, M. and Hazama, F. (1998) Normal developing rat brain expresses a platelet-derived growth factor B chain (c-sis) mRNA truncated at the 5' end. *Oncogene*, **16**, 1571–1578.
148. Gaildrat, P., Killian, A., Martins, A., Tournier, I., Frébourg, T. and Tosi, M. (2010) Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants. *Methods Mol. Biol.*, **653**, 249–57.

9. ABKÜRZUNGSVERZEICHNIS

APS	Ammoniumperoxodisulfat
ARP2/3 complex	<i>actin related protein 2/3 complex</i>
ARPC2	<i>actin related protein 2/3 complex subunit 2</i>
Bcl-2	B-cell lymphoma 2
bp	Basenpaar(e)
BSA	<i>bovine serum albumin</i>
CD	Circulardichroismus
cDNA	<i>complementary DNA</i>
CDS	<i>coding sequence</i>
CMV	Cytomegalievirus
C _t	<i>cycle threshold</i>
DNA	deoxyribonucleic acid, Desoxyribonucleinsäure
dNTP	Deoxynucleosidtriphosphat
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
EFHD2	<i>EF-hand domain family, member D2</i>
EGF	<i>epidermal growth factor</i>
GFP	<i>green fluorescent protein</i>
G-Quadruplex, GQ	Guanin-Quadruplex
hnRNP	<i>heterogeneous nuclear ribonucleoprotein</i>
IRES	<i>internal ribosome entry site</i>
k _a	Assoziationsgeschwindigkeitskonstante
Kb	Kilobasenpaare
k _d	Dissoziationsgeschwindigkeitskonstante
K _D	Dissoziationskonstante
LB	<i>lysogeny broth</i>
MMP16	<i>membrane-type matrix metalloproteinase 3</i>
mRNA	<i>messenger RNA</i>
MYC	Myelocytomatose
NHE	<i>nuclease hypersensitive element</i>
NRAS	neuroblastoma rat sarcoma
PBS	<i>Phosphate-buffered saline</i>

PCR	<i>polymerase chain reaction</i> , Polymerase-Kettenreaktion
qPCR	quantitative PCR
RNA	<i>ribonucleic acid</i> , Ribonucleinsäure
RNAse	Ribonuclease
rpm	<i>revolutions per minute</i> ; Umdrehungen pro Minute
RT	reverse Transkription
SDS	<i>sodium dodecyl sulfate</i>
SRSF1	<i>serine/arginine-rich splicing factor 1</i>
TAE	TRIS-Acetat-EDTA
TBE	TRIS-Borat-EDTA
TBS	TRIS- <i>buffered saline</i>
TEMED	Tetramethylethylendiamin
TERC	<i>telomerase RNA component</i>
TRIS	Tris(hydroxymethyl)-aminomethan
U2AF65	<i>U2 auxiliary factor 65</i>
UTR	untranslated region, untranslatierte Region
UV	Ultraviolettstrahlung
VEGF	<i>vascular endothelial growth factor</i>
Zic-1	<i>zinc finger of the cerebellum 1</i>

10. DANKSAGUNG

Vor allem möchte ich Prof. Dr. Jens Kurreck für die Möglichkeit danken meine Doktorarbeit in seiner Arbeitsgruppe auf dem spannenden Gebiet der G-Quadruplexe anzufertigen. Auch möchte ich mich für die Betreuung und die Zeit, die er sich immer genommen hat, bedanken.

Prof. Dr. Mario Mörl und Prof. Dr. Peter Neubauer danke ich für die Begutachtung meiner Arbeit. Des Weiteren Danke ich Prof. Dr. Neubauer für die Unterstützung bei der Proteinexpression.

Christoph Weise danke ich für die Durchführung der, sehr interessanten, Massenspektrometrie-Messungen. Tatjana Schütze, Zoltan Konthur und Markus Menger danke ich für die Hilfe bei den SPR-Messungen, wobei Markus Menger zusätzlich ganz besonderer Dank für die Unterstützung bei den Auswertungen gilt.

Meinen Kollegen aus der Angewandten Biochemie danke ich für die Unterstützung und für die nette Arbeitsatmosphäre. Besonders bedanken möchte ich mich bei Viola Röhrs, die sowohl bei den praktischen Arbeiten immer eine große Hilfe war als auch diese Arbeit Korrektur gelesen hat. Bernd Krostitz danke ich für die Unterstützung bei der Zellkultur sowie für den großen Keks-Vorrat. Petra Seifert möchte ich für die Hilfe bei allerlei administrativen Aufgaben danken. Und natürlich Tatsuo Serikawa für die tolle Verstärkung des GQ-Teams. Bei den ehemaligen Mitarbeiterinnen Diana Rothe und Anke Wagner möchte ich mich für die hilfreichen Ratschläge und Diskussionen bedanken. Des Weiteren danke ich meinen Studenten Robert Shahab und Delia Yang für ihre fleissige Arbeit.

11. ANHANG

11.1. PLASMIDKARTEN

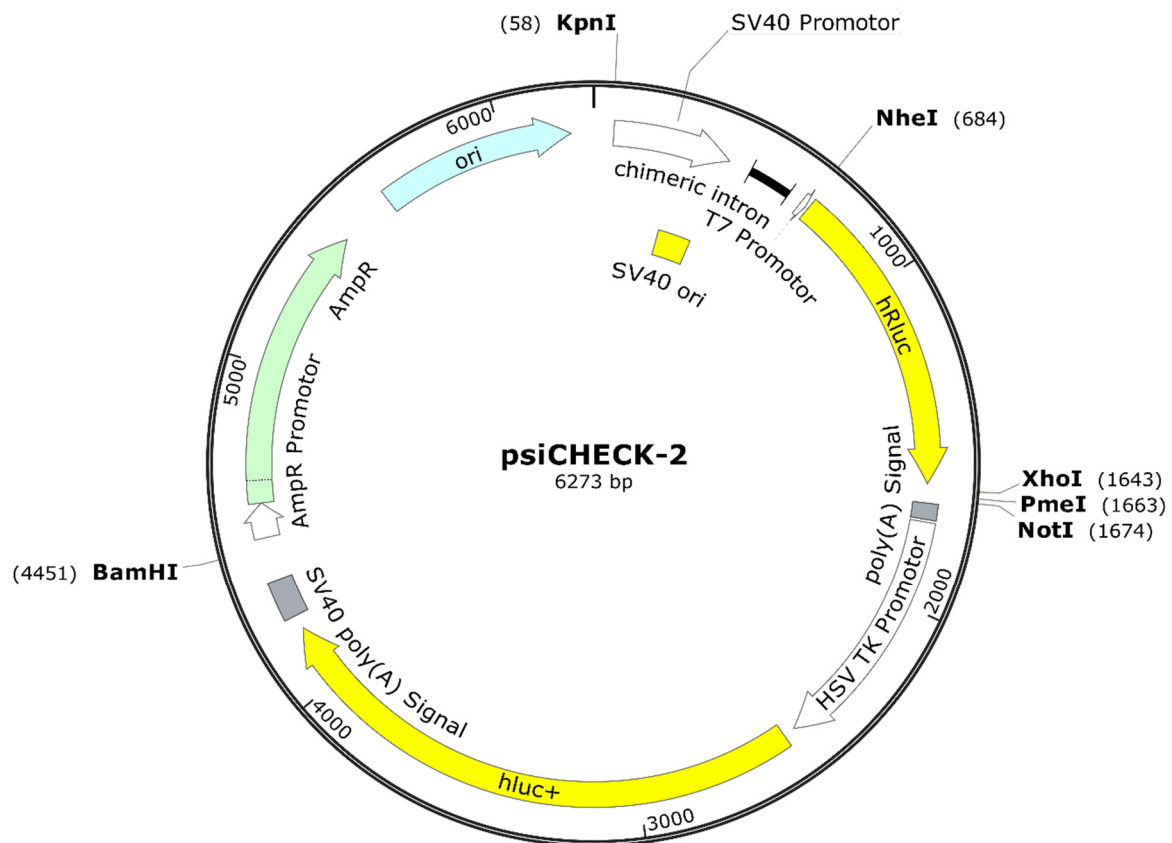


Abbildung 41: Vektorkarte des Reporterplasmids psiCHECK-2. Zur Klonierung in die 5'-UTR der Renilla-Luciferase wurde die NheI-Schnittstelle verwendet, zur Klonierung in die 3'-UTR wurden die XhoI und NotI Schnittstellen benutzt. Schnittstellen, die nur einmal vorkommen, sind eingezeichnet. Plasmidelemente: AmpR: Ampicillinresistenz; ori: *origin of replication*; SV40 ori: *simian virus 40 ori*; hRluc: synthetisches Renilla-Luciferase-Gen; HSV-TK Promotor: *herpes simplex virus thymidine kinase* Promotor; hLuc+: synthetisches Firefly-Luciferase-Gen; poly(A)-Signal: Polyadenylierungssignal

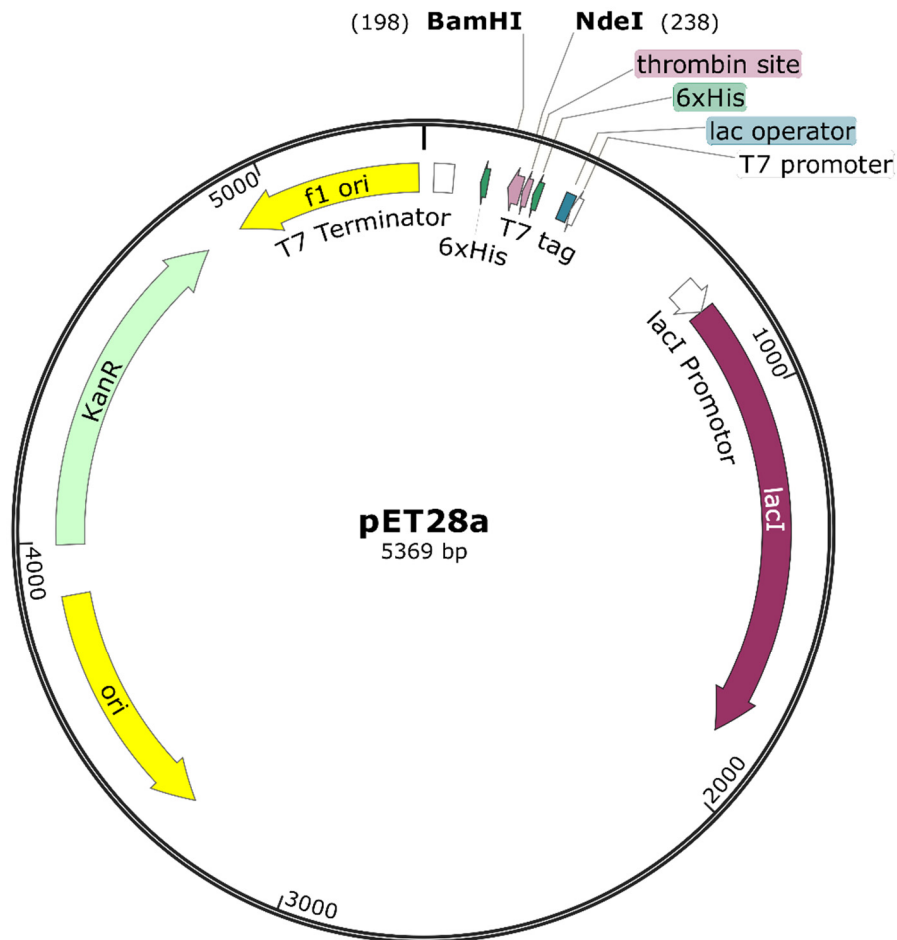


Abbildung 42: Vektorkarte des prokaryotischen Expressionsvektors pET28a. Die codierenden Sequenzen für die zu exprimierenden Proteine wurden in die BamHI und NdeI Schnittstellen eingefügt. Plasmidelemente: KanR: Kanamycin-Resistenzgen; ori: *origin of replication*; 6xHis: Polyhistidin-Tag; lacI: Lac-Repressor

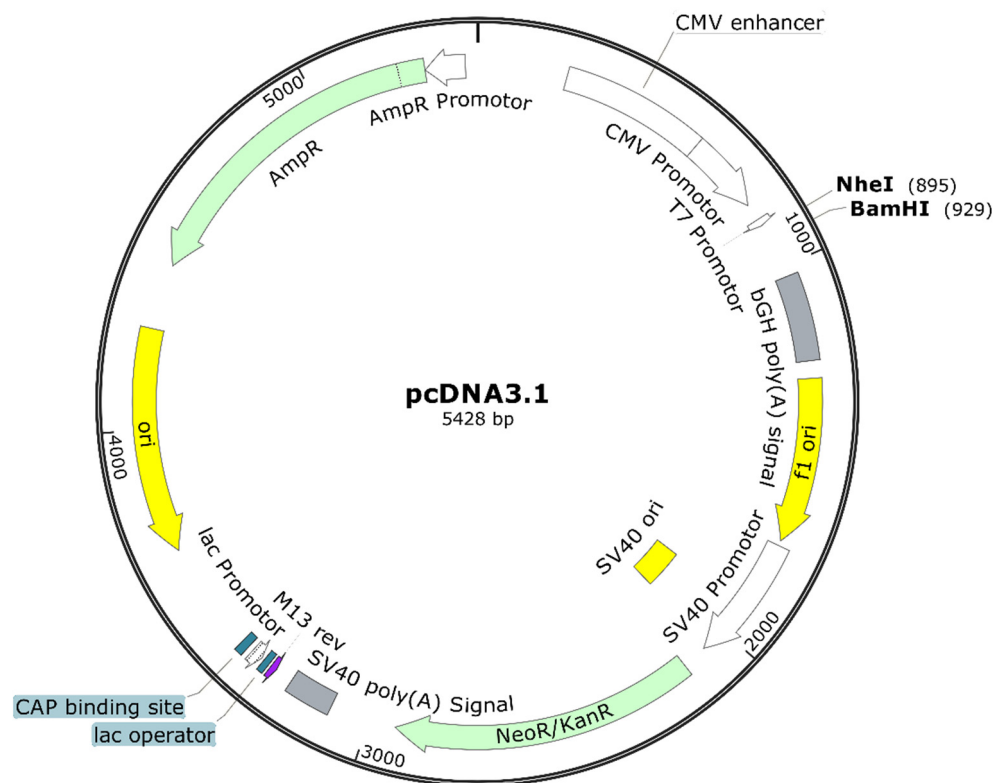


Abbildung 43: Vektorkarte des eukaryotischen Expressionsvektors pcDNA3.1. Die codierenden Sequenzen für die zu exprimierenden Proteine wurden in die BamHI und NheI Schnittstellen eingefügt. Plasmidelemente: AmpR: Ampicillinresistenz; CMV Promotor: Cytomegalievirus Promotor; ori: *origin of replication*; SV40 ori: *simian virus 40 ori*; poly(A)-Signal: Polyadenylierungssignal

11.2. SDS-PAGES DER PULL-DOWN-ASSAYS

11.2.1. MMP16

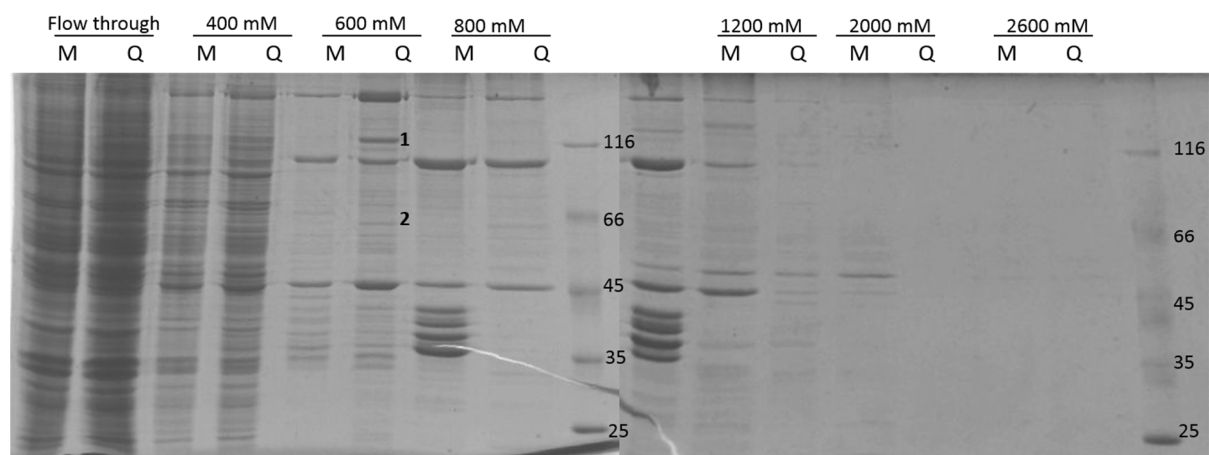


Abbildung 44: Pull-down-Assay mit Hek293 Zelllysats. Die markierten Banden wurden massenspektrometrisch untersucht.
1. hnRNP 2 MAOM, U2AF2

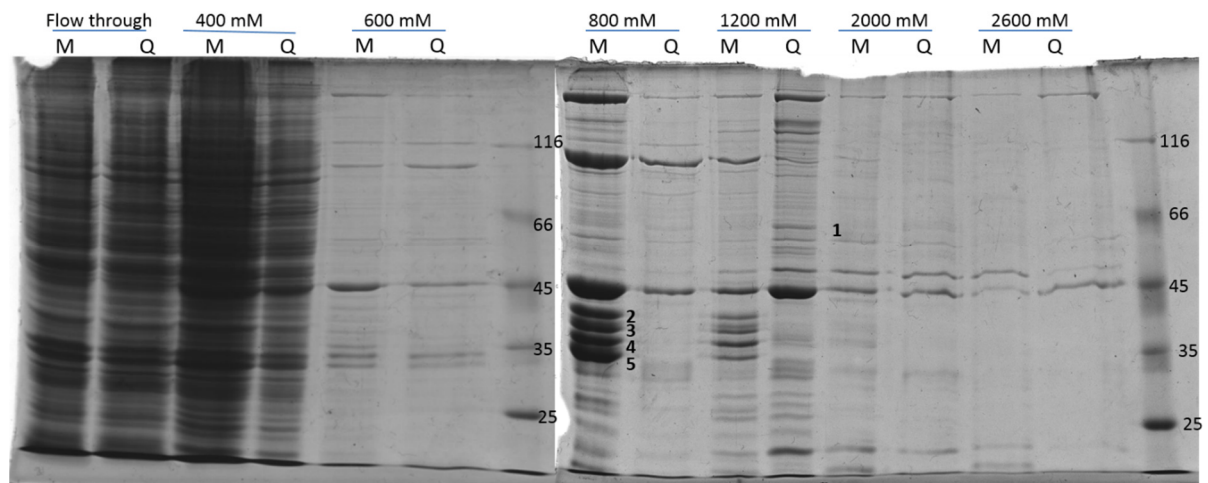


Abbildung 45: Pull-down-Assay mit Hek293 Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. MMP16 G-Quadruplex 1. MAOM und U2AF2 MMP16 mut: 2. hnRNPA3 3. hnRNPA3 und hnRNPA2/B1 4. hnRNPA2/B1, 5. hnRNPA1

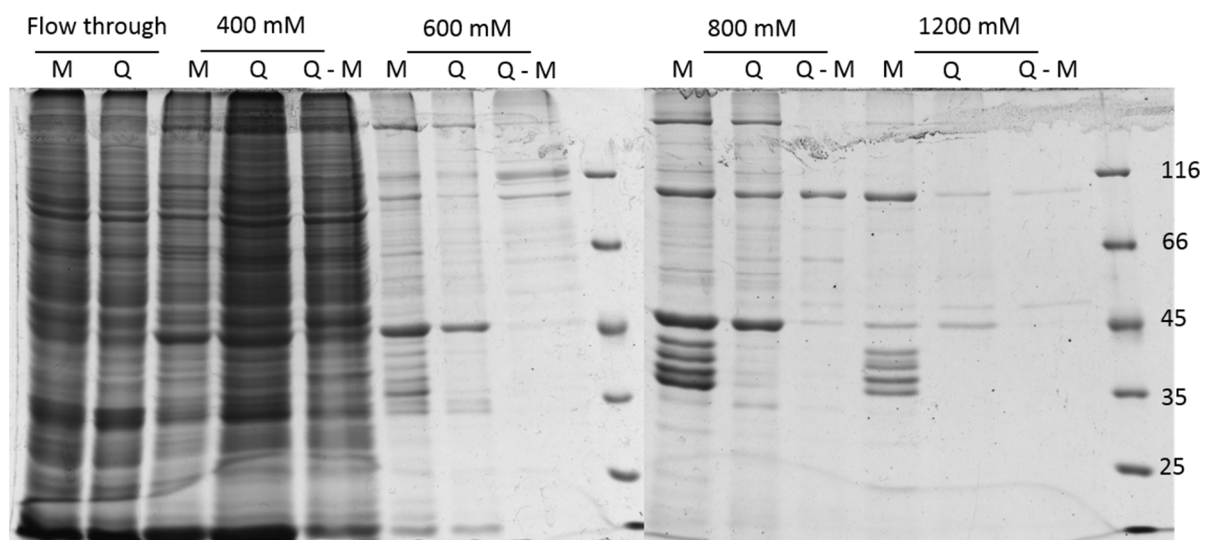


Abbildung 46: Pull-down-Assay mit Hek293 Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Hier wurde kein Protein gefunden, das nur an das G-Quadruplex Oligonucleotid gebunden hat.

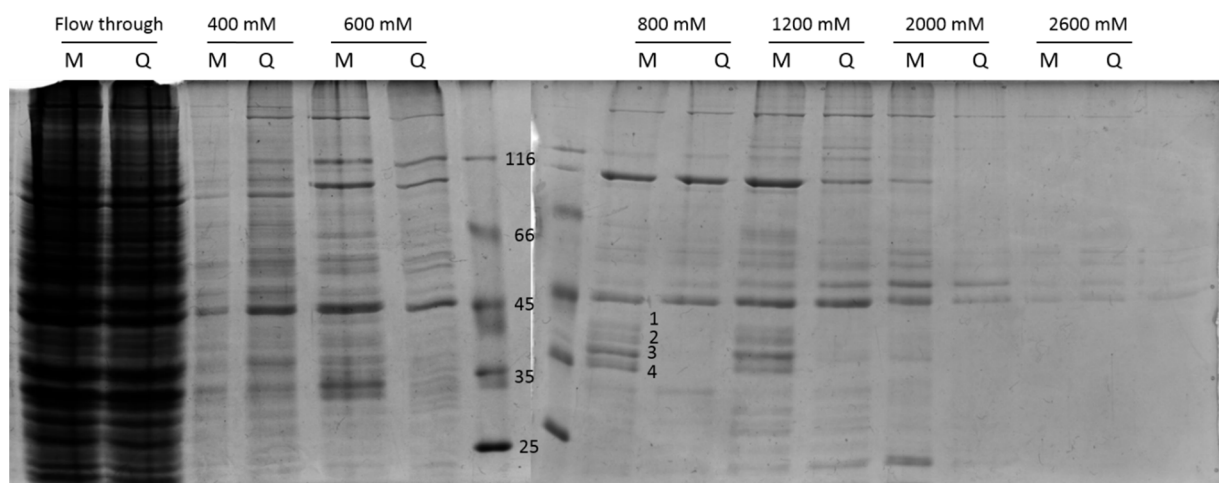


Abbildung 47: Pull-down-Assay mit HeLa Zelllysatsat. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. hnRNP A3 2. hnRNP A3 und hnRNP A2/B1 3. hnRNP A2/B1 4. hnRNP A1

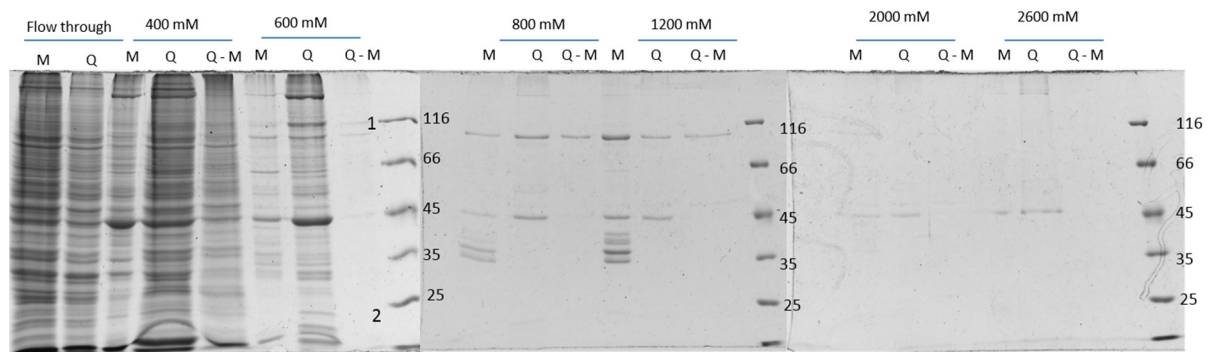


Abbildung 48: Pull-down-Assay mit HeLa Zelllysatsat. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. hnRNP U und Aktin 2. rps5

11.2.2. ARPC2

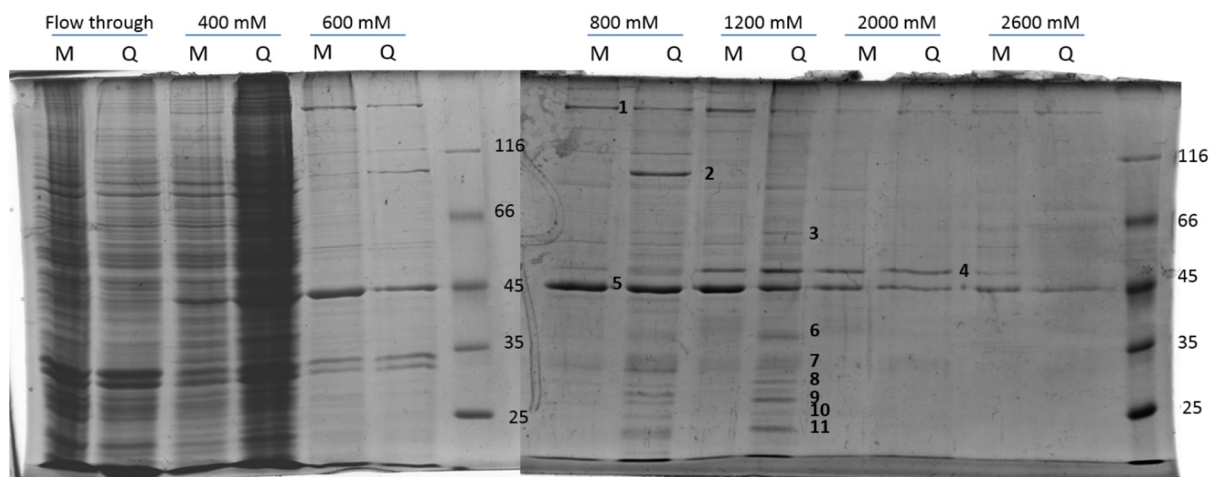


Abbildung 49: Pull-down-Assay mit Hek293 Zelllysatsat. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. Myosin 2. Nucleolin 3. DBPA, 4. YB-1 5. Aktin 6. RPL6 7. EFHD2, SRSF1 8. RPL7 9. SRSF9, RPL19 10. RPL10 11. RPS9

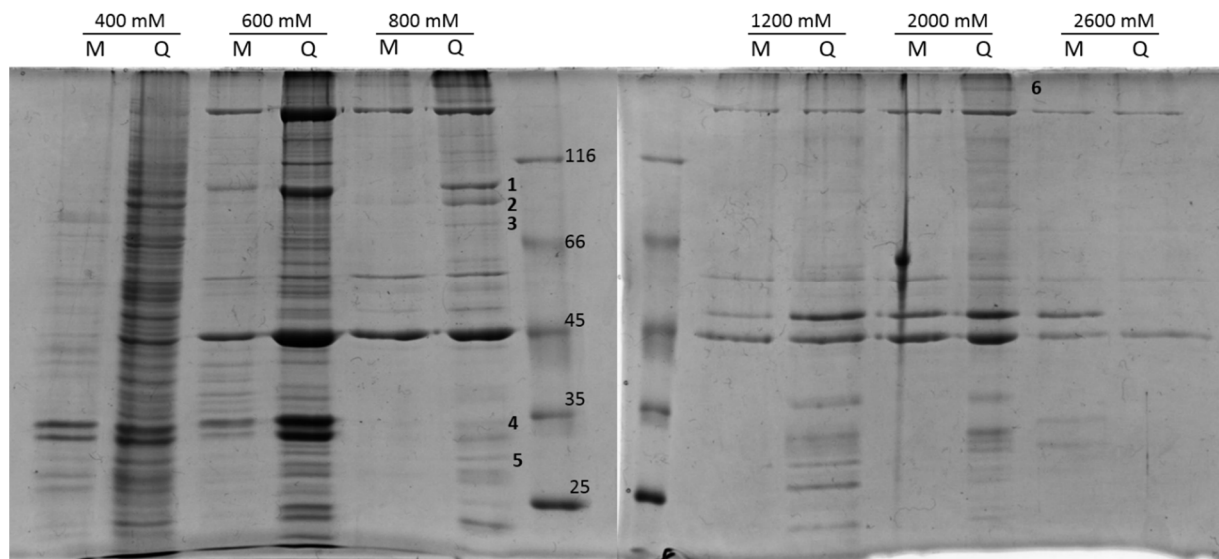


Abbildung 50: Pull-down-Assay mit Hek293 Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. Nucleolin 2. Actin 3. RNA binding protein 14 4. Tropomyosin alpha 4 5. EFHD2 SRSF1 6. Myosin9

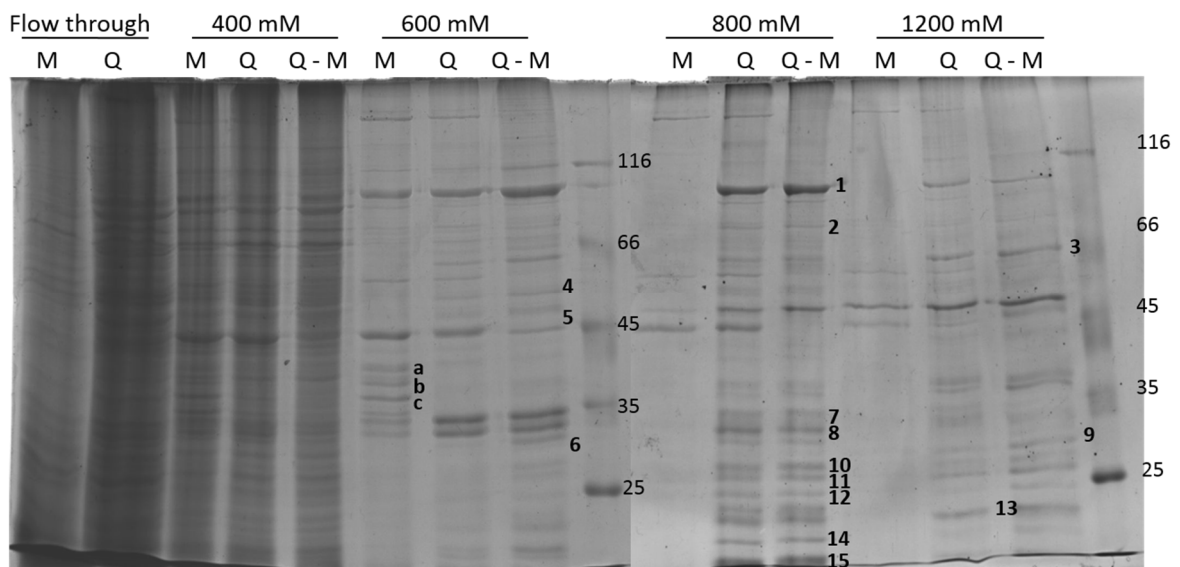


Abbildung 51: Pull-down-Assay mit Hek293 Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. Nucleolin 2. RNA-binding Protein 14 3. U2AF2 4. hnRNPH 5. YBOX1, hnRNPF 6. rpS2 7. SRSF1 8. rpS6, SRSF1 9. rpL7 10. rpL19, SRSF9 11. SRSF9, rpL14 12. rpS9, rpL10 13. rpS9 14. rpL26 15. rpL27a ARPC2 mut2: (a) hnRNPA3, (b) hnRNPA2/B1 (c) hnRNPA2/B1

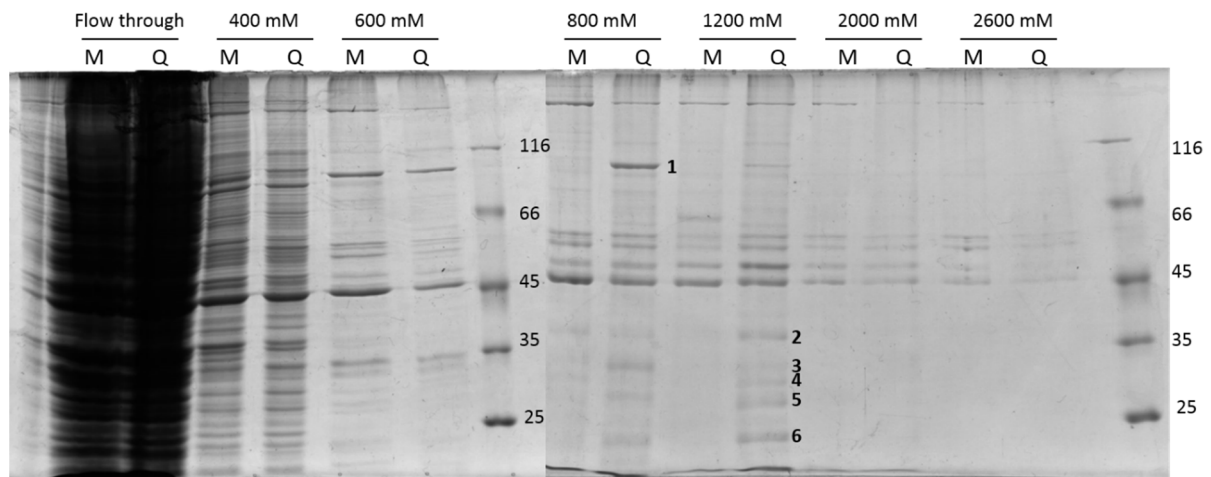


Abbildung 52: Pull-down-Assay mit HeLa Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. Nucleolin, 2. rpL6, 3. EFHD2, SRSF1, 4. rpL7, 5. rpL19, SRSF9, 6. rpS9

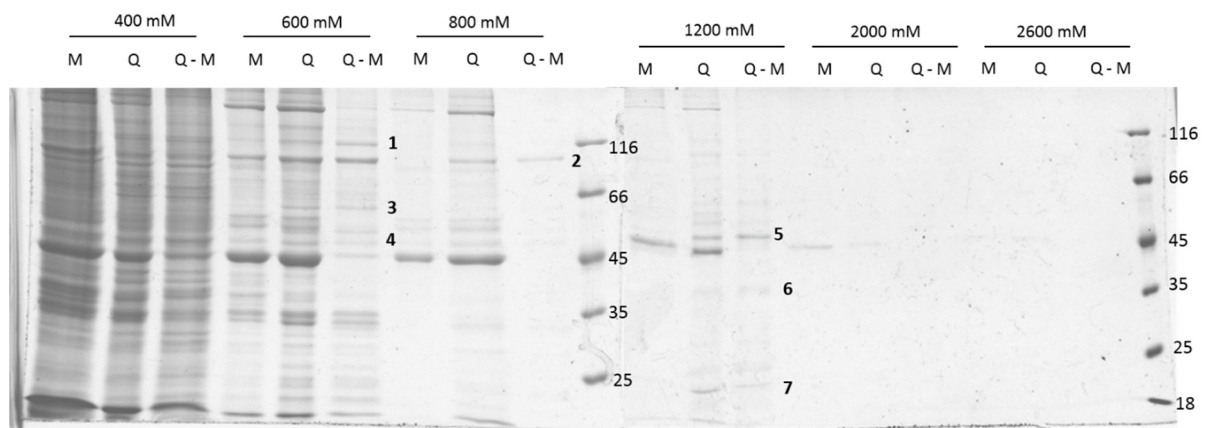


Abbildung 53: Pull-down-Assay mit HeLa Zelllysate. M: mutiertes Oligonucleotid; Q: G-Quadruplex. Die markierten Banden wurden massenspektrometrisch untersucht. 1. hnRNP, 2. Nucleolin, 3. G3BP2, U2AF2 4. hnRNPH, hnRNPH2, 5. YBOX1, 6. rpL6, 7. rpL12

11.3. MASSENSPEKTROMETRIE DATEN

Tabelle 19: Massenspektrometrie-Daten der Proteine die an die G-Quadruplex oder Kontrollsequenzen gebunden haben. Datenbankvergleiche wurden mit Mascot gegen die SwissProt Datenbank durchgeführt. Die Suche wurde auf humane Sequenzen beschränkt.

Identified Protein	SwissProt Accession code	Spot #	Apparent MW (kDa)	Theoretical MW (kDa)	Mass finger-print data Expect value	Peptide tolerance (PMF, +/- ppm)	Assigned Peptides	Sequence Coverage (%)	MS+MS/MS Combined (Protein summary deprecated) Expect value	Number of MS/MS	Selected masses for MS/MS
Ribosomal Proteins											
60S ribosomal protein L6	RL6_HUMAN	10	35	33	0.002	75	9	23	6.60E-07	1	1158.68
					0.001	150	13	29			
							1*		2.00E-10	1	2509.20
60S ribosomal protein L7	RL7_HUMAN	11	30	29	2.0E-06	75	11	35			
					8.3E-06	100	12	39			
					0.00011	75	9	32			
60S ribosomal protein L10	RL10_HUMAN	14	25	25	0.0036	75	8	27			
					0.003	60	10	30			
60S ribosomal protein L12	RL12_HUMAN		20	18			1*		0.1	2	714.31, 1685.89
60S ribosomal protein L14	RL14_HUMAN		25	24	0.032	75	9	32			
60S ribosomal protein L19	RL19_HUMAN		25	24	0.00019	125	12	40	0.00019	3	1065.41, 1454.54, 1942.75
					5.6E-06	75	15	32			
60S ribosomal protein L26	RL26_HUMAN	15	17	17		75			4.0E-10		955.44, 1078.62
60S ribosomal protein L27a	RL27A_HUMAN	16	16	17	0.0042	75	8	41			
40S ribosomal protein S2	RS2_HUMAN	6	30	32	2.4E-06	50	11	28			
40S ribosomal protein S5	RS5_HUMAN		23	23	0.00011	75	11	31	2.6E-07	1	1100.60
40S ribosomal protein S6	RS6_HUMAN		28	29	0.023	100	9	30			
40S ribosomal protein S9	RS9_HUMAN	17	23	23	5.1E-07	75	14	45			
					0.003	60	10	30			
					1.0E-12	75	17	48			
					4.0E-08	50	12	42			
Splicing Factors											
Heterogeneous ribonucleoprotein F	nuclear HNRPF_HUMAN	5	48	46	2.0E-08	75	15	39			

Identified Protein	SwissProt Accession code	Spot #	Apparent MW (kDa)	Theoretical MW (kDa)	Mass finger-print data Expect value	Peptide tolerance (PMF, +/- ppm)	Assigned Peptides	Sequence Coverage (%)	MS+MS/MS Combined (Protein summary deprecated) Expect value	Number of MS/MS	Selected masses for MS/MS
Heterogeneous nuclear ribonucleoprotein H	HNRH1_HUMAN	4	50	50	1.0E-08	125	17	42			
					5.1E-07	70	18	42			
Heterogeneous nuclear ribonucleoprotein U	HNRPU_HUMAN		116	91	1.6E-07	75	27	27			
					0.00072	75	17	20	n.s.	2	976.44, 1198.69
					0.00042	70	12	14			
Serine/arginine-rich factor 1	SRSF1_HUMAN		30	28	0.11	50	10	46			
		9			0.017	100	10	39			
		10			4.2E-06	100	11	40			
					6.9E-06	75	12	50			
Serine/arginine-rich factor 9	SRSF9_HUMAN		25	26	0.066	125	15	55	8.1E-12	3	1065.41, 1454.54, 1942.75
					6.7E-05	75	13	43			
		12			0.0002	75	11	36			
		13			0.0099	75	8	28			
Splicing factor U2AF	U2AF2_HUMAN		66	54	6.4E-13	75	18	41	4.1E-17	4	960.50, 1285.72, 1909.06, 1923.00
					3.2E-08	75	12	39			
					2.2E-05	70	12	31	2.0E-11	3	955.49, 1368.60, 1909.02
		3			0.0012	75	10	21			
Others											
EF-hand domain-containing protein D2	EFHD2_HUMAN		32	27	6.4E-10	75	15	52			
					0.0002	50	12	42			
					1.3E-05	100	12	46			
Ras GTPase-activating protein-binding protein 2	G3BP2_HUMAN				2.0E-07	54	16	32			

Identified Protein	SwissProt Accession code	Spot #	Apparent MW (kDa)	Theoretical MW (kDa)	Mass finger-print data Expect value	Peptide tolerance (PMF, +/- ppm)	Assigned Peptides	Sequence Coverage (%)	MS+MS/MS Combined (Protein summary deprecated) Expect value	Number of MS/MS	Selected masses for MS/MS
NAD-dependent malic enzyme, mitochondrial	MAOM_HUMAN			66	0.0022	75	11	23	1.0E-12	4	960.50, 1285.72, 1909.06, 1923.00
		1			0.00032	75	11	24			
					4.0E-08	75	16	27			
Nucleolin	NUCL_HUMAN		100	77	4.7E-06	75	21	27			
		7			1.3E-10	100	23	28			
					8.1E-11	75	20	25			
					2.6E-12	75	24	35			
					2.5E-09	70	20	27			
RNA binding protein 14	RBM14_HUMAN		70	70	1.0E-12	75	23	34			
		8			2.0E-07	75	12	22			
Tropomyosin alpha 4	TPM4_HUMAN		35	29	1.3E-06	75	16	45			
Y-box-binding protein 3	YBOX3_HUMAN		60	40	3.2E-07	75	13	31	8.3E-12	1	1559.80
Proteins binding to controls											
Actin, aortic smooth muscle	ACTA_HUMAN	g	45	42	0.067	120	6	18	0.00038	1	1198.70
Actin, cytoplasmic 1	ACTB_HUMAN	a	45	42	6.4E-11	75	20	60			
Actin, cytoplasmic 2	ACTG_HUMAN	a	45	42	6.4E-11	75	20	60			
Heterogeneous nuclear ribonucleoprotein A1	ROA1_HUMAN	e	34	39	5.1E-16	75	23	50			
Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	c	38	38	1.6E-12		22	63			
Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	d	36	38	1.0E-10	50	17	43			
Heterogeneous nuclear ribonucleoprotein A3	ROA3_HUMAN	b	40	40	1.0E-07	75	17	46			
		c	38	40	0.00066	75	12	36			
Nuclease-sensitive element-binding protein 1	YBOX1_HUMAN	2	48	36	2.5E-11	75	17	37			
		f			2.6E-08	75	12	23			
		g			4.0E-07	120	11	25	2.00E-07	1	1198.70

PMF, peptide mass fingerprint; * Proteins were identified by unique peptide sequences determined by MALDI-TOF-MS/MS experiments

Tabelle 20: Ergebnisse der Orbitrap-Analyse der SILAC-Experimente.

ARPC2	Ratio H/L	MMP16	Ratio H/L	NRAS	Ratio H/L
60S ribosomal protein L35	42,643	NAD-dependent malic enzyme, mitochondrial	13,953	Nucleolar RNA helicase 2	17,411
60S ribosomal protein L17	38,022	RRP15-like protein	10,203	60S ribosomal protein L23a	17,36
60S ribosomal protein L32	26,791	Putative methyltransferase NSUN5	9,2438	Enhancer of rudimentary homolog	16,608
60S ribosomal protein L37	23,502	Suppressor of SWI4 1 homolog	8,0555	Heterogeneous nuclear ribonucleoprotein U	15,553
60S ribosomal protein L13	17,553	60S ribosomal protein L26;60S ribosomal protein L26-like 1	6,9603	Transcriptional repressor CTCF	12,904
60S ribosomal protein L26	17,104	60S ribosomal protein L23a	6,5728	Bcl-2-associated transcription factor 1	10,538
40S ribosomal protein S6	16,801	Zinc finger protein 579	6,3344	ATP-dependent RNA helicase DDX50	9,7722
60S ribosomal protein L23a	16,738	Enhancer of rudimentary homolog	6,2982	CCAAT/enhancer-binding protein zeta	9,4426
40S ribosomal protein S24	16,289	40S ribosomal protein S24	6,031	Putative methyltransferase NSUN5	8,9039
Nucleolar RNA helicase 2	15,255	7SK snRNA methylphosphate capping enzyme	5,7578	RNA pseudouridylate synthase domain-containing protein 2	8,693
RNA-binding protein 4	14,896	Bcl-2-associated transcription factor 1	5,0324	RRP15-like protein	8,5087
Nucleolin	14,882	40S ribosomal protein S2	5,0309	General transcription factor 3C polypeptide 3	8,4514
40S ribosomal protein S2	14,756	Glioma tumor suppressor candidate region gene 2 protein	5,0237	Chromodomain-helicase-DNA-binding protein 6;Chromodomain-helicase-DNA-binding protein 7;Chromodomain-helicase-DNA-binding protein 9;Chromodomain-helicase-DNA-binding protein 8	8,4336
40S ribosomal protein S9	13,547	ATP-dependent RNA helicase DHX8	4,9815	General transcription factor 3C polypeptide 4	8,3987
60S ribosomal protein L27a	13,42	Methyl-CpG-binding domain protein 1	4,6487	A-kinase anchor protein 17A	8,2155
Peptidyl-prolyl cis-trans isomerase G	13	Heterogeneous nuclear ribonucleoprotein F;Heterogeneous nuclear ribonucleoprotein F, N-terminally processed	4,3896	G-rich sequence factor 1	7,8008
Heterogeneous nuclear ribonucleoprotein F;Heterogeneous nuclear ribonucleoprotein F, N-terminally processed	12,939	R3H and coiled-coil domain-containing protein 1	4,371	General transcription factor 3C polypeptide 5	7,7334
Ribosomal L1 domain-containing protein 1	12,896	ATP-dependent RNA helicase DDX50	4,1982	Uncharacterized protein C19orf29	7,5116
PUF60_HUMAN Isoform 6 of Poly(U)-binding-splicing factor PUF60	12,151	Elongation factor 1-alpha 1;Putative elongation factor 1-alpha-like 3;Elongation factor 1-alpha 2	4,1182	General transcription factor 3C polypeptide 1	7,4448
Serine/threonine-protein kinase PRP4 homolog	11,804	Thyroid hormone receptor-associated protein 3	4,084	Nucleolar complex protein 2 homolog	7,436
40S ribosomal protein S29	11,48	Translational activator GCN1	4,0769	Peptidyl-prolyl cis-trans isomerase G	7,2124
Bcl-2-associated transcription factor 1	11,261	A-kinase anchor protein 17A	4,0401	Suppressor of SWI4 1 homolog	7,1889
Apoptotic chromatin condensation inducer in the nucleus	10,799	Peptidyl-prolyl cis-trans isomerase G	4,0185	ATP-dependent RNA helicase DHX8	6,9345
Serine/arginine-rich splicing factor 1	10,661	Tubulin beta-8 chain	3,8775	Arginine--tRNA ligase, cytoplasmic	6,9023
Splicing factor, arginine/serine-rich 19	10,584	60S ribosomal protein L32	3,8258	U3 small nucleolar ribonucleoprotein MPP10	6,8439

Splicing factor U2AF 65 kDa subunit	10,488	60S ribosomal protein L39; Putative 60S ribosomal protein L39-like 5	3,7819	Coiled-coil domain-containing protein 86	6,7815
Putative RNA-binding protein Luc7-like 2	10,446	40S ribosomal protein S13	3,6988	DBIRD complex subunit KIAA1967	6,6953
60S ribosomal protein L10; 60S ribosomal protein L10-like	10,273	Nucleolar RNA helicase 2	3,6957	Heterogeneous nuclear ribonucleoprotein Q	6,5989
ATP-dependent RNA helicase DDX54	10,083	DnaJ homolog subfamily A member 1	3,6801	Thyroid hormone receptor-associated protein 3	6,5849
Enhancer of rudimentary homolog	9,9435	40S ribosomal protein S4, X isoform; 40S ribosomal protein S4, Y isoform 2	3,5239	Histone-binding protein RBBP4; Histone-binding protein RBBP7	6,5655
60S ribosomal protein L36	9,6851	Putative RNA-binding protein Luc7-like 2	3,501	Probable ATP-dependent RNA helicase DDX27	6,5423
Histone H2B type 1-K; Histone H2B type 1-D; Histone H2B type 1-C/E/F/G/I; Histone H2B type 2-F; Histone H2B type 1-H; Histone H2B type 1-N; Histone H2B type 1-M; Histone H2B type F-S; Histone H2B type 1-L	9,5349	60S ribosomal protein L17	3,4978	Tubulin beta chain	6,3555
Putative RNA-binding protein Luc7-like 1	9,3672	Putative RNA-binding protein Luc7-like 1	3,389	Death domain-associated protein 6	6,3509
Insulin-like growth factor 2 mRNA-binding protein 3	9,3367	Tubulin beta-4B chain	3,2942	Heterogeneous nuclear ribonucleoprotein F; Heterogeneous nuclear ribonucleoprotein F, N-terminally processed	6,2931
Histone H2A type 2-A; Histone H2A type 2-C; Histone H2A type 1; Histone H2A type 1-D; Histone H2A type 1-H; Histone H2A type 1-J; Histone H2A.J	9,3109	60S ribosomal protein L10; 60S ribosomal protein L10-like	3,2669	General transcription factor 3C polypeptide 2	6,2894
Heterogeneous nuclear ribonucleoprotein U-like protein 2	9,1623	ATP-dependent RNA helicase DDX54	3,2593	Chromodomain-helicase-DNA-binding protein 3	6,0023
Serine/arginine-rich splicing factor 5	9,1151	Poly(rC)-binding protein 2; Poly(rC)-binding protein 3	3,2592	Apoptotic chromatin condensation inducer in the nucleus	5,9449
RNA-binding protein 39	8,9798	General transcription factor 3C polypeptide 5	3,2396	Tubulin beta-8 chain	5,94
40S ribosomal protein S15	8,582	Splicing factor U2AF 35 kDa subunit	3,2364	Oxysterol-binding protein-related protein 3	5,6679
Serine/arginine-rich splicing factor 3	8,4996	Tubulin beta chain	3,2109	Methyl-CpG-binding domain protein 1	5,6258
Isoleucine--tRNA ligase, cytoplasmic	8,375	60S ribosomal protein L37	3,1193	Tubulin beta-4B chain; Tubulin beta-4A chain	5,6121
Serine/arginine-rich splicing factor 11	8,2172	40S ribosomal protein S18	3,1096	Aspartate--tRNA ligase, cytoplasmic	5,5746
40S ribosomal protein S17; 40S ribosomal protein S17-like	8,207	RNA-binding protein 39	3,088	DnaJ homolog subfamily A member 1	5,5317
40S ribosomal protein S8	8,117	Splicing factor U2AF 65 kDa subunit	3,0562	Leucine--tRNA ligase, cytoplasmic	5,5238
60S ribosomal protein L5	8,0712	40S ribosomal protein S23	3,0262	Protein SCAF11	5,4488
Histone H2B type 1-J; Histone H2B type 1-O; Histone H2B type 1-B; Histone H2B type 2-E; Histone H2B type 3-B	7,5652	Replication initiator 1	3,0078	60S ribosomal protein L32	5,4284
Histone H2A type 1-B/E; Histone H2A type 3; Histone H2A type 1-C	7,4237	Uncharacterized protein C19orf29	2,9304	40S ribosomal protein S29	5,3792
MKI67 FHA domain-interacting nucleolar phosphoprotein	7,2852	40S ribosomal protein S15	2,8757	Matrin-3	5,2452
Splicing factor U2AF 35 kDa subunit	7,1457	60S ribosomal protein L24	2,8744	Zinc finger protein 579	5,2003
G-rich sequence factor 1	6,9825	CCAAT/enhancer-binding protein zeta	2,8658	Protein polybromo-1	5,1748
40S ribosomal protein S23	6,9064	60S ribosomal protein L36	2,8377	Nucleolar GTP-binding protein 2	5,1395
Luc7-like protein 3	6,9002	Transcriptional repressor CTCF	2,7938	Something about silencing protein 10	4,9771
Probable ATP-dependent RNA helicase DHX36	6,8828	RNA-binding protein 26	2,7891	Protein KRI1 homolog	4,7202

Poly(rC)-binding protein 1	6,5925	Vimentin;Neurofilament medium polypeptide;Alpha-internexin;Desmin	2,7798	Fragile X mental retardation syndrome-related protein 2	4,5859
Suppressor of SWI4 1 homolog	6,5594	Tubulin alpha-1A chain;Tubulin alpha-3C/D chain;Tubulin alpha-3E chain	2,7334	DNA-directed RNA polymerase II subunit RPB2	4,5511
60S ribosomal protein L19	6,5129	60S ribosomal protein L21	2,7242	Bifunctional glutamate/proline--tRNA ligase;Glutamate--tRNA ligase;Proline--tRNA ligase	4,4748
GC-rich sequence DNA-binding factor 1	6,5059	DBIRD complex subunit KIAA1967	2,718	60S ribosomal protein L17	4,4486
CCAAT/enhancer-binding protein zeta	6,3614	Luc7-like protein 3	2,6915	60S ribosomal protein L10	4,4426
Uncharacterized protein C19orf29	6,3453	Chromodomain-helicase-DNA-binding protein 5;Chromodomain-helicase-DNA-binding protein 4;Chromodomain-helicase-DNA-binding protein 3	2,6735	Splicing factor U2AF 35 kDa subunit	4,3833
Heterogeneous nuclear ribonucleoprotein U	6,1913	Poly(rC)-binding protein 1	2,6502	Isoleucine--tRNA ligase, cytoplasmic	4,2397
Ribosome biogenesis protein NSA2 homolog	6,1376	40S ribosomal protein S11	2,5669	Probable ATP-dependent RNA helicase DDX46	4,1592
Ribosome biogenesis protein BRX1 homolog	6,1111	Serine/arginine-rich splicing factor 11	2,5337	40S ribosomal protein S2	4,1384
Serine/arginine-rich splicing factor 9	5,8706	Heterogeneous nuclear ribonucleoprotein H;Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	2,5185	Splicing factor U2AF 65 kDa subunit	4,1203
Probable ATP-dependent RNA helicase DDX46	5,7945	Nuclease-sensitive element-binding protein 1	2,506	Probable RNA-binding protein 23	4,1121
60S ribosomal protein L24	5,5978	60S ribosomal protein L28	2,479	ATP synthase subunit alpha, mitochondrial	3,9435
ATP-dependent RNA helicase DDX50	5,5751	Poly(U)-binding-splicing factor PUF60	2,4562	Leucine-rich PPR motif-containing protein, mitochondrial	3,8562
rRNA 2-O-methyltransferase fibrillarin	5,4894	DNA mismatch repair protein Msh6	2,4557	Heterochromatin protein 1-binding protein 3	3,8553
Thyroid hormone receptor-associated protein 3	5,4684	CAD protein;Glutamine-dependent carbamoyl-phosphate synthase;Aspartate carbamoyltransferase;Dihydroorotase	2,4483	CAD protein;Glutamine-dependent carbamoyl-phosphate synthase;Aspartate carbamoyltransferase;Dihydroorotase	3,847
Periodic tryptophan protein 1 homolog	5,4371	Ornithine aminotransferase, mitochondrial;Ornithine aminotransferase, hepatic form;Ornithine aminotransferase, renal form	2,4291	60S ribosomal protein L21	3,8316
H/ACA ribonucleoprotein complex subunit 4	5,4148	60S ribosomal protein L35	2,4211	General transcription factor II-I	3,8072
60S ribosomal protein L7a	5,3646	Heterogeneous nuclear ribonucleoprotein U	2,4168	KRR1 small subunit processome component homolog	3,8035
Zinc finger CCCH-type antiviral protein 1	5,3276	40S ribosomal protein S16	2,4122	6-phosphofructokinase, liver type	3,8022
H/ACA ribonucleoprotein complex subunit 1	5,2515	40S ribosomal protein S6	2,2964	DNA-directed RNA polymerase II subunit RPB1	3,7978
Probable rRNA-processing protein EBP2	5,2395	Ribosome biogenesis protein NSA2 homolog	2,248	DNA mismatch repair protein Msh6	3,7925
60S ribosomal protein L3	5,1627	6-phosphofructokinase, liver type	2,2374	SRA stem-loop-interacting RNA-binding protein, mitochondrial	3,788
Splicing regulatory glutamine/lysine-rich protein 1	5,1395	60S ribosomal protein L13	2,2201	Heterogeneous nuclear ribonucleoprotein H;Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	3,787
RNA-binding protein 26	5,1099	Heterogeneous nuclear ribonucleoprotein U-like protein 1	2,2015	Small subunit processome component 20 homolog	3,7468
Serine/arginine-rich splicing factor 2	5,0695	Protein FAM111B	2,1929	Protein AATF	3,6671

Poly(rC)-binding protein 2	5,0643	60S ribosomal protein L3	2,187	Transcription intermediary factor 1-beta	3,6289
Histone-binding protein RBBP4;Histone-binding protein RBBP7	5,0451	Serine/arginine-rich splicing factor 3	2,1779	Protein PRRC2C	3,6119
60S ribosomal protein L13a	5,0333	ATP synthase subunit alpha, mitochondrial	2,1366	60S ribosomal protein L23	3,4693
Heat shock cognate 71 kDa protein	4,9825	Matrin-3	2,1295	ATP-dependent RNA helicase DDX3X;ATP-dependent RNA helicase DDX3Y	3,4075
Arginine and glutamate-rich protein 1	4,9508	Proline-, glutamic acid- and leucine-rich protein 1	2,1169	RNA-binding protein 39	3,4048
A-kinase anchor protein 17A	4,8681	Heat shock cognate 71 kDa protein	2,0988	Putative RNA-binding protein Luc7-like 2	3,4013
40S ribosomal protein S18	4,7515	Heterogeneous nuclear ribonucleoprotein Q;Heterogeneous nuclear ribonucleoprotein R	2,0886	Elongation factor 1-alpha 1;Putative elongation factor 1-alpha-like 3;Elongation factor 1-alpha 2	3,3736
Surfeit locus protein 6	4,531	60S ribosomal protein L19	2,0846	U3 small nucleolar ribonucleoprotein protein IMP3	3,2878
H/ACA ribonucleoprotein complex subunit 3	4,5067	Heterogeneous nuclear ribonucleoprotein U-like protein 2	2,0208	Testis-expressed sequence 10 protein	3,2459
40S ribosomal protein S13	4,4949	Heat shock 70 kDa protein 1A/1B	2,0029	Proline-, glutamic acid- and leucine-rich protein 1	3,2162
Ribosome production factor 1	4,4869			Ornithine aminotransferase, mitochondrial;Ornithine aminotransferase, hepatic form;Ornithine aminotransferase, renal form	3,206
Heterogeneous nuclear ribonucleoprotein H;Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	4,4464			Luc7-like protein 3	3,1815
60S ribosomal protein L18a	4,3477			6-phosphofructokinase type C	3,154
Putative ribosomal RNA methyltransferase NOP2	4,3083			Tubulin beta-2B chain;Tubulin beta-2A chain	3,1499
Fragile X mental retardation syndrome-related protein 2	4,2094			Probable ATP-dependent RNA helicase DDX41	3,1498
H/ACA ribonucleoprotein complex subunit 2	4,1811			Putative RNA-binding protein Luc7-like 1	3,0725
Insulin-like growth factor 2 mRNA-binding protein 1	4,1527			60S ribosomal protein L38	3,0269
60S ribosomal protein L15	4,1412			Heat shock cognate 71 kDa protein	3,0082
Heat shock protein HSP 90-beta;Heat shock protein HSP 90-alpha;Putative heat shock protein HSP 90-beta 2;Putative heat shock protein HSP 90-alpha A2	4,1157			Chromodomain-helicase-DNA-binding protein 4;Chromodomain-helicase-DNA-binding protein 5	2,9834
Lupus La protein	4,0474			ATP-dependent RNA helicase DDX54	2,9656
Glioma tumor suppressor candidate region gene 2 protein	4,0349			40S ribosomal protein S11	2,9411
Nucleolar protein 56	4,0212			60S ribosomal protein L24	2,9294
Leucine--tRNA ligase, cytoplasmic	3,9963			Putative rRNA methyltransferase 3	2,9176
SURP and G-patch domain-containing protein 2	3,9943			RRP12-like protein	2,8756
60S ribosomal protein L21	3,9694			MKI67 FHA domain-interacting nucleolar phosphoprotein	2,8402
60S ribosomal protein L28	3,8687			Cyclin-dependent kinase 12	2,828
40S ribosomal protein S16	3,8654			RNA-binding protein 26	2,8276
ATP synthase subunit alpha, mitochondrial	3,75			60S ribosomal protein L39;Putative 60S ribosomal protein L39-like 5	2,7969
Ribosome biogenesis protein BOP1	3,6921			SURP and G-patch domain-containing protein 2	2,7705

Cold-inducible RNA-binding protein	3,6692
Pescadillo homolog	3,6616
Guanine nucleotide-binding protein-like 3	3,6442
40S ribosomal protein S26	3,5274
Ribosome production factor 2 homolog	3,4728
60S ribosomal protein L7-like 1	3,4675
Heat shock 70 kDa protein 1A/1B	3,4007
Serine/arginine-rich splicing factor 6	3,386
RNA-binding protein 14	3,3781
Putative rRNA methyltransferase 3	3,3627
Matrin-3	3,328
Serine/arginine repetitive matrix protein 2	3,3233
R3H and coiled-coil domain-containing protein 1	3,2632
Serine/arginine-rich splicing factor 4	3,229
Proteasome activator complex subunit 4	3,2196
Tubulin beta chain	3,1217
Tubulin beta-2B chain;Tubulin beta-2A chain	3,0863
Tubulin beta-4B chain;Tubulin beta-4A chain	3,0661
ATP-dependent RNA helicase DHX8	3,0427
Nucleolar GTP-binding protein 2	3,036
NHP2-like protein 1	2,9642
40S ribosomal protein S11	2,9257
Peptidyl-tRNA hydrolase ICT1, mitochondrial	2,918
60S ribosomal protein L27	2,914
Small nuclear ribonucleoprotein-associated proteins B and B;Small nuclear ribonucleoprotein-associated protein N	2,9074
39S ribosomal protein L3, mitochondrial	2,897
78 kDa glucose-regulated protein	2,8202
Tubulin beta-8 chain	2,7572
40S ribosomal protein S3a	2,7475
HEAT repeat-containing protein 1	2,7443
39S ribosomal protein L27, mitochondrial	2,7408
Heterogeneous nuclear ribonucleoprotein U-like protein 1	2,7355
Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	2,7325
39S ribosomal protein L1, mitochondrial	2,7136
Pumilio domain-containing protein KIAA0020	2,6874

DnaI homolog subfamily A member 2	2,7692
Heat shock protein HSP 90-beta	2,7446
Heat shock 70 kDa protein 1A/1B	2,7179
Tubulin beta-6 chain	2,6905
Poly(rC)-binding protein 1	2,6399
Replication initiator 1	2,5706
Tubulin alpha-1C chain	2,5336
40S ribosomal protein S27;40S ribosomal protein S27-like	2,5114
Large proline-rich protein BAG6	2,4483
Insulin-like growth factor 2 mRNA-binding protein 3	2,4272
60S ribosomal protein L26	2,4266
60S ribosomal protein L37	2,4159
Protein PRRC2A	2,3812
Melanoma-associated antigen D2	2,3659
Poly(rC)-binding protein 2	2,3549
40S ribosomal protein S26;Putative 40S ribosomal protein S26-like 1	2,3484
RNA-binding motif protein, X chromosome;RNA-binding motif protein, X chromosome, N-terminally processed	2,3323
Serine/threonine-protein kinase PRP4 homolog	2,2128
Periodic tryptophan protein 1 homolog	2,2096
Translational activator GCN1	2,174
Protein FAM111B	2,0815
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5;Probable global transcription activator SNF2L1	2,0765
Serine/arginine-rich splicing factor 11	2,0523
Heterogeneous nuclear ribonucleoprotein U-like protein 2	2,0443
Glioma tumor suppressor candidate region gene 2 protein	2,0169
116 kDa U5 small nuclear ribonucleoprotein component	2,0082

Ribosome biogenesis regulatory protein homolog	2,6673
40S ribosomal protein S14	2,6646
60S ribosomal protein L8	2,6598
Nucleolar protein 58	2,6479
60S ribosomal protein L37a	2,6179
39S ribosomal protein L19, mitochondrial	2,5994
60S ribosomal protein L7	2,5896
ATP-dependent RNA helicase DDX24	2,57
Tubulin alpha-1A chain;Tubulin alpha-3C/D chain;Tubulin alpha-3E chain	2,5662
60S ribosomal protein L6	2,5544
Histone H4	2,5443
Ornithine aminotransferase, mitochondrial;Ornithine aminotransferase, hepatic form;Ornithine aminotransferase, renal form	2,4989
Serine/arginine repetitive matrix protein 1	2,486
60S ribosomal protein L11	2,4766
60S ribosomal protein L14	2,4577
Small nuclear ribonucleoprotein Sm D1	2,4074
La-related protein 1	2,3718
39S ribosomal protein L35, mitochondrial	2,3538
Small nuclear ribonucleoprotein Sm D3	2,3391
60S ribosomal protein L35a	2,296
60S ribosomal protein L4	2,2439
60S acidic ribosomal protein P1	2,2421
39S ribosomal protein L39, mitochondrial	2,2396
39S ribosomal protein L38, mitochondrial	2,2383
60S acidic ribosomal protein P2	2,2065
60S ribosomal protein L12	2,2013
60S ribosomal protein L10a	2,1114
Voltage-dependent anion-selective channel protein 1	2,1062
60S ribosomal protein L30	2,0906
60S acidic ribosomal protein P0;60S acidic ribosomal protein P0-like	2,0735
Histone H3.1;Histone H3.3;Histone H3.1t;Histone H3.3C	2,037

11.4. PUBLIKATIONEN

von Hacht, A., Seifert, O., Menger, M., Schütze, T., Arora, A., Konthur, Z., Neubauer, P., Wagner, A., Weise, C., Kurreck, J. (2014). Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Research*, 42(10), 6630–6644. doi:10.1093/nar/gku290

Foit, L., Morgan, G. J., Kern, M. J., Steimer, L. R., von Hacht, A. A., Titchmarsh, J., Warriner, S. L., Radford, S. E., Bardwell, J. C. A. (2009). Optimizing protein stability in vivo. *Molecular Cell*, 36(5), 861–71. doi:10.1016/j.molcel.2009.11.022

Mac, T.-T., von Hacht, A., Hung, K.-C., Dutton, R. J., Boyd, D., Bardwell, J. C. A., & Ulmer, T. S. (2008). Insight into disulfide bond catalysis in Chlamydia from the structure and function of DsbH, a novel oxidoreductase. *The Journal of Biological Chemistry*, 283(2), 824–32. doi:10.1074/jbc.M707863200

Quan, S., Schneider, I., Pan, J., Von Hacht, A., & Bardwell, J. C. A. (2007). The CXXC motif is more than a redox rheostat. *The Journal of Biological Chemistry*, 282(39), 28823–33. doi:10.1074/jbc.M705291200

KONFERENZBEITRÄGE

Posterpräsentation beim “third international meeting on G-quadruplex and G-assembly”, Juni 2012, Sorrent, Italien; von Hacht, A., Arora, A., Weise, C., Kurreck, J. “Regulation of Gene Expression through RNA G-Quadruplexes”