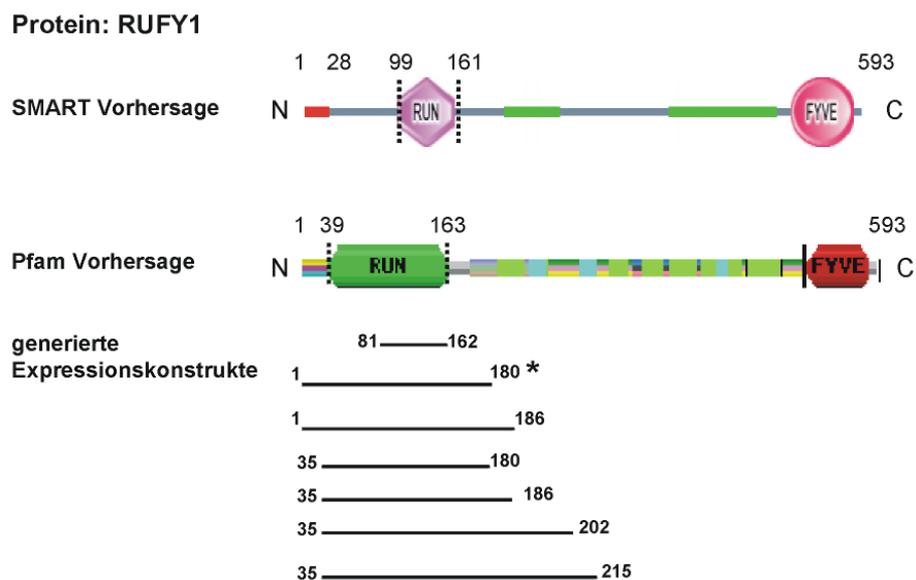


### 3 Generierung von Domänenkonstrukten für die NMR-Strukturbestimmung

#### 3.1 Herstellung von Expressionskonstrukten am Beispiel der RUN-Domäne

Aufgrund der unterschiedlichen Vorhersage der SMART- und Pfam-Datenbank für die Domänengrenze der RUN-Domäne wurden unterschiedlich lange Konstrukte hergestellt (Abbildung 3.1). Unter Verwendung des Gateway®-Systems wurden insgesamt 16 Konstrukte, die Teilsequenzen aus 7 verschiedenen humanen Proteinen enthalten, generiert. Abbildung 3.1 zeigt einen schematischen Überblick über die erzeugten RUN-Konstrukte am Beispiel des Proteins RUFY1, ausgehend von den vorhergesagten Domänengrenzen nach SMART und Pfam.



**Abbildung 3.1:** Schematischer Überblick über die generierten RUN-Konstrukte. Expressionskonstrukte (unten) ausgehend von den in der SMART- und Pfam-Datenbank vorhergesagten Domänengrenzen am Beispiel des humanen Proteins RUFY1 (SwissProt:Q96T51). Das Konstrukt in den Grenzen 1-180 (markiert mit \*) erwies sich als einzig lösliches unter den generierten Konstrukten und den verwendeten Expressionsbedingungen.

Insgesamt wurden 16 Konstrukte aus den folgenden 8 humanen Proteinen mit den Zugangsnummern bzw. -namen Q9NT69, Q96C34, Q9UF50, KIAA1091, KIAA0397, BC008078, NESCA (BC001045) und RUFY1 (Q96T51) hergestellt. Zusammenfassend gibt Tabelle. 3.1 einen Überblick über den Verlauf und die Erfolgsrate der einzelnen Schritte PCR, Klonierung, Expressions- und Löslichkeitstest. Dabei zeigte sich, dass die

PCR für 94 % aller Konstrukte ein geeignetes PCR-Produkt lieferte, was dann in den weiteren Schritten eingesetzt werden konnte. 74 % der Klonierungen waren erfolgreich und führten zu Expressionsplasmiden und den entsprechenden Expressionsklonen. Dabei wurden sowohl GST- als auch His-Tag Expressionsplasmide erzeugt. Diese wurden dann in einem Expressionstest in LB-Medium bezüglich ihrer Expressionsstärke, Größe und Löslichkeit des Proteins untersucht. Der Expressionstest wurde bei vier verschiedenen Temperaturen (16, 25, 30 und 35°C) durchgeführt. 65 % aller Expressionsplasmide zeigten die Produktion des gewünschten Proteins in *E.coli*. Nur eines von 16 generierten Konstrukten exprimiert lösliches Protein, während die anderen 10 Konstrukte *inclusion bodies* produzierten.

**Tabelle. 3.1:** Übersicht der generierten RUN-Konstrukte mit Hilfe des Gateway® Systems.

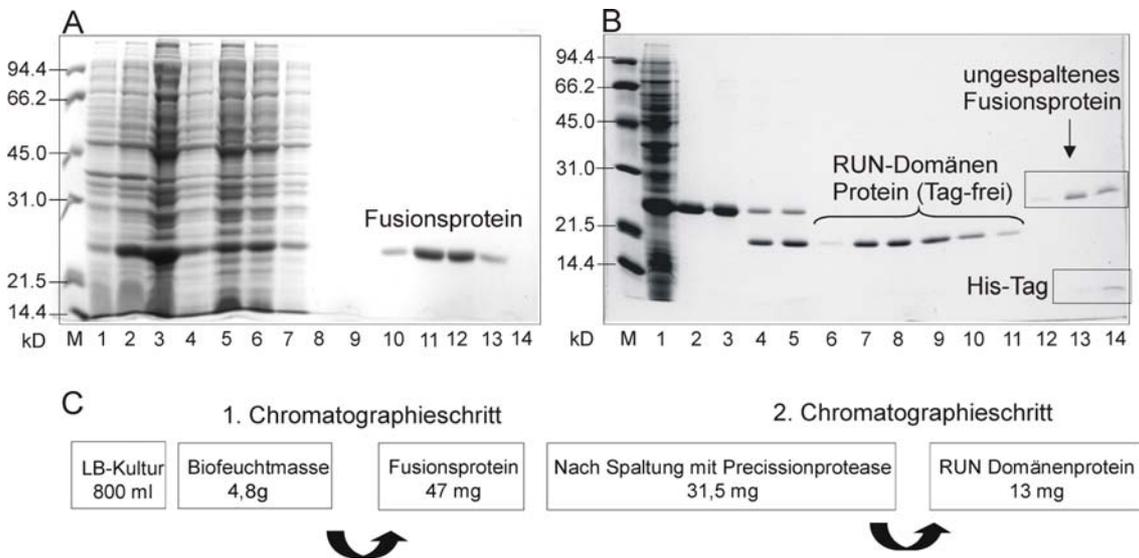
Proteinname	Start-Ende	PCR	Klonierung	Expressionstest	Löslichkeit
Q9NT69	1-135	+	+	+	-
KIAA1091	1282-1359	+	+	+	-
KIAA0397	132-204	-	-	-	-
Q96C34	528-614	+	-	-	-
NESCA	127-198	+	+	-	-
Q9UF50	52-199	+	+	+	-
Q9UF50	120-189	+	+	-	-
RUFY 1	81-162	+	+	+	-
RUFY 1	1-180	+	+	+	+
RUFY 1	1-186	+	+	+	-
RUFY 1	35-180	+	+	+	-
RUFY 1	35-186	+	+	+	-
RUFY 1	35-202	+	+	+	-
RUFY 1	35-215	+	+	+	-
BC008078	562-777	+	+	+	-
BC008078	614-777	+	-	-	-
<b>Erfolgsrate gesamt (%)</b>		<b>15/16 (94)</b>	<b>13/16 (74)</b>	<b>11/16 (65)</b>	<b>1/16 (6)</b>

Erfolgreiche Schritte wurden mit einem Pluszeichen (+), erfolglose mit einem Minuszeichen (-) gekennzeichnet.

Aus der Aufstellung der Ergebnisse wird ersichtlich, dass die Herstellung geeigneter Expressionkonstrukte relativ effizient verläuft und kaum Verluste auftreten. Der größte Engpaß liegt in der Bereitstellung von löslichem und gefaltetem Proteinprobenmaterial zur weiteren strukturellen Charakterisierung.

### 3.2 Proteinpräparation und Reinigung der RUN-Domäne (1-180) aus dem humanen Protein RUFY1

Die Anzucht der Biofeuchtmasse erfolgte bei einer Temperatur von 37°C welche vor Induktion mit IPTG (0,5 mM) auf 16°C reduziert wurde, um die Löslichkeit des exprimierten RUN-Konstruktes zu erhöhen. Expressionen bei höheren Temperaturen führten zur Bildung von unlöslichen *inclusion bodies*. Nach Zentrifugation und Aufschluss der Zellen erfolgten die weiteren chromatografischen Reinigungsschritte. In einem ersten Reinigungsschritt wird das His-Tag Fusionsprotein über Affinitätschromatographie an eine Ni<sup>2+</sup>-Chelat Säule gebunden, von löslichen Fremdbestandteilen des Zellüberstandes getrennt und anschließend eluiert. Der His-Tag wird übernacht mittels PreScissionProtease abgespalten. Das RUN-Konstrukt wird dann in einem zweiten analogen Affinitätschromatographie-Reinigungsschritt erhalten. Abbildung 3.2 zeigt die SDS-Page-Gele mit den Proben der einzelnen Reinigungsschritte.

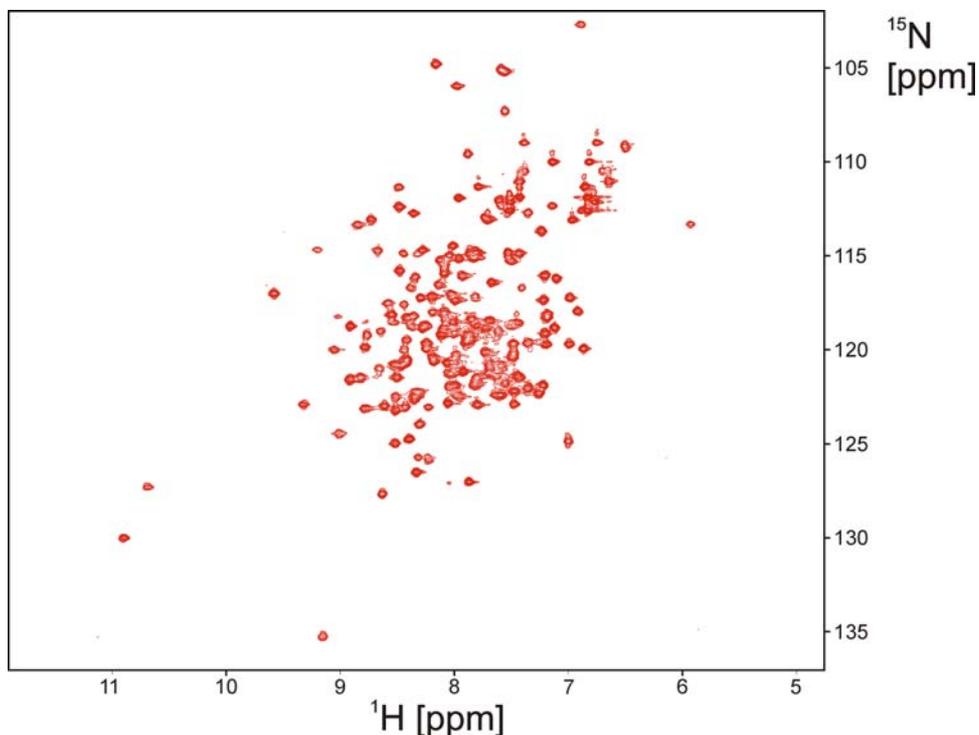


**Abbildung 3.2:** 12 % SDS-Page-Gele mit Proben der beiden Reinigungsschritte des RUN-Konstruktes (A) nach primärer Affinitätschromatographie. M: Proteinmarker (kD), 1: nicht induziert, 2: nach Induktion mit IPTG, 3: löslicher Überstand nach French Press, 4-9: Durchfluß des nicht gebundenen Zellextraktes, 10-13: His-Fusionsprotein, 14: Elutionsende (B) nach Abspaltung des His-Tags mit PreScissionProtease. M: Proteinmarker wie in (A), 1: A3, 2: A11, 3: A12, 4-5: partielle Abspaltung des His-Tags, nach dem zweiten Reinigungsschritt 6-11: gereinigtes Tag-freies RUN-Domänen Protein, 12-14: ungespaltene His-Fusionsprotein und abgespaltener His-Tag (zusätzlich in 13/14) (C) Überblick über die Erfolgsraten der einzelnen Präparationsschritte der Expression und Aufreinigung.

Die im linken Gel (A) aufgetragenen Proben zeigen die Situation bei Zellinduktion und nach dem ersten Affinitätschromatographieschritt. Mit dem Gel auf der rechten Seite (B) wurde die proteolytische Abspaltung des His-Tags und der Verlauf des sich direkt anschließenden zweiten Affinitätschromatographieschrittes verfolgt. In (C) abgebildet ist die Ausbeute der einzelnen Präparationsschritte ausgehend von einer 800 ml LB-Kultur. Aus 4,8 g Biofeuchtmasse konnten nach Durchlauf des ersten Chromatographieschrittes 47 mg Fusionsprotein gewonnen werden. Diese zunächst hohe Ausbeute reduziert sich auf 31,5 mg Gesamtproteinmenge nach Abspaltung des His-Tags und bei erneuter Aufreinigung auf 13 mg tagfreie RUN-Domäne. Die Verluste entstanden durch Aggregation des Proteins bei Dialyse- und Reinigungsschritten.

### 3.3 NMR-spektroskopische Untersuchungen

Von einer  $^{15}\text{N}$ -markierten Proteinprobe wurde ein  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum bei 289 K auf einem 750 MHz Spektrometer (Bruker) aufgenommen. Abbildung 3.3 zeigt das nicht zugeordnete  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum des RUN-Konstruktes (1-180) aus dem humanen Protein RUFY1. Die Resonanzen des Spektrums sind gut aufgelöst und zeigen chemische Verschiebungen, die für ein gefaltetes Protein typisch sind. Das  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektrum stellt einen *fingerprint* des Proteins dar.



**Abbildung 3.3:**  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum des RUN-Konstruktes (1-180) aus RUFY1 aufgenommen auf einem 750 MHz Spektrometer bei 289 K. Das Spektrum zeigt gut aufgelöste Signale, die für ein gefaltetes Protein typisch sind.

Weitere spektroskopische Untersuchungen hinsichtlich der Resonanzzuordnung befinden sich in Arbeit. Sie gestalten sich als sehr schwierig, da die Proteinproben einem starken Aggregationsverhalten unterliegen, was zum Ausflocken des Proteins nach wenigen Tagen führt. Auch durch Zuhilfenahme von reduzierenden Substanzen, wie beispielsweise DTT (Dithiothreitol), konnte dieses Verhalten nicht unterdrückt werden. Experimentelle Versuche, diesen Zustand in einem Aggregationsscreen durch Zugabe von Kosolventien (n-Octyl- $\beta$ -D-glucopyranosid, Tritin x-100, Glycerol, Betain, Harnstoff), chaotropen Salzen ( $(\text{NH}_4)_2\text{SO}_4$ ), reduzierenden Substanzen (DTT) und hypothetischen Kofaktoren (GTP,  $\text{Zn}^{2+}$ ) zu stabilisieren, waren nicht erfolgreich.

Alternativ wurden Kristallisationsansätze in der FG von Herrn Prof. Dr. Udo Heinemann (Max-Delbrück-Centrum, Berlin) durchgeführt, die ebenfalls zu keinen positiven Ergebnissen führten.

### **3.4 Welche Anforderungen an die Probengüte sind für die NMR-Strukturbestimmung zu stellen?**

Zu allererst ist es notwendig, das gefaltete Protein vorliegt. Dies erkennt man beispielsweise an gut aufgelösten Signalen eines  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrums und an schmalen Linienbreiten im allgemeinen. Je größer das zu untersuchende Protein ist, desto breiter werden die Signale. Daraus folgt eine schlechtere Dispersion der Signale durch Überlagerung dieser. Zweitens muss die Probe mindestens über einen Zeitraum von mindestens 7 Tagen stabil sein, um den derzeitigen technischen Möglichkeiten der Spektrenaufnahme Rechnung zu tragen. Drittens muss das Protein in ausreichender Menge vorhanden sein, da ein gutes Signal-Rausch Verhältnis für NMR-Untersuchungen unerlässlich ist. Man verwendet daher routinemäßig Proteinproben mit einer Probenkonzentration größer als 0,5 mM. Viertens ist eine erfolgreiche Strukturbestimmung davon abhängig, wie komplex das Protein aufgebaut ist. Flexible Bereiche tragen aufgrund der schnellen Relaxation zu einem schwachen oder nicht detektierbaren Signal bei. Weiter bedingt die Art der Reste wie das Vorhandensein von Aromaten komplizierter interpretierbare Spektren aufgrund von Überlagerungen der Resonanzsignale.

#### **3.4.1 Welche Bedingungen erfüllt das RUN-Konstrukt (1-180) ?**

Das RUN-Konstrukt (1-180) zeigt ein gut aufgelöstes  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum, welches dem eines gefalteten Proteins entspricht. Damit ist eine wichtige Voraussetzung für eine Strukturbestimmung erfüllt. Allerdings kann die mangelnde Stabilität der Probe ein Problem darstellen. Dies zeigte sich im starken Aggregationsverhalten, was auch zu großen Verlusten während des Reinigungsprozesses führte. Das RUN-Konstrukt konnte maximal bis zu einer Konzentration von 0,2-0,3 mM aufkonzentriert werden. Oberhalb dieser Grenze begann die Lösung zu aggregieren und auszuflocken. Proteinlösungen mit einer Konzentration von 0,3 mM zeigten schon nach wenigen Tagen die Tendenz zu aggregieren. Bei dem hier hergestellten RUN-Konstrukt (1-180) handelt es sich um kein optimales Target für eine Strukturbestimmung im Rahmen eines *structural genomics* Projektes, da nur wenige der obengenannten Anforderungen erfüllt sind.