



Doctoral Thesis

Characterization of RNA-binding proteins involved in post-transcriptional regulation

Author(s):

Daubner, Gerrit Martin

Publication Date:

2013

Permanent Link:

<https://doi.org/10.3929/ethz-a-009910037> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH No 21049

**Characterization of RNA-binding proteins involved in
post-transcriptional regulation**

A dissertation submitted to the

ETH ZÜRICH

for the degree of

DOCTOR OF SCIENCES

presented by

GERRIT MARTIN DAUBNER

M.Sc. (Biochemistry), Ludwig-Maximilians-Universität München

born 21. November 1981

citizen of Germany

accepted on the recommendation of

Prof. Dr. Frédéric Allain

Prof. Dr. Jonathan Hall

Prof. Dr. James Stevenin

Prof. Dr. Gerhard Wider

2013

Summary

The origin of eukaryotes is a milestone in the evolution of multicellular organism and was accompanied by the development of complex intracellular networks to ensure specific and proper regulation of gene expression. Post-transcriptional processing determines the reliability and diversity of gene expression and its tight regulation is guided by a large number of RNA-binding proteins. Already tiny defects like a single nucleotide substitution can severely affect protein-RNA interactions within the regulation machinery and subsequently lead to severe diseases. Therefore it is indispensable to study the protein-RNA interactions important for post-transcriptional regulation on a molecular basis in order to rationalize their impact on cellular homeostasis. Here, we have investigated the function of three different RNA-binding proteins at a molecular level by solving the solution structure of the respective protein-RNA complex and/ or by conducting functional assays.

The serine/ arginine rich splicing factor 2 (SRSF2) is a splicing regulator and was shown to bind highly degenerated RNA sequences. In solving the solution structure of the RNA-binding motif (RRM) of SRSF2 with two different RNA targets and through combination with extensive mutagenesis and two different functional assays, we were able to identify its precise specificity and consensus sequence. Furthermore, the RRM binds RNA in a novel sandwich-like structure with the involvement of both N- and C-terminal ends, emphasizing the versatility of RNA binding by RRMs.

In contrast to SRSF2, the polypyrimidine tract binding protein (PTB) is primarily a repressor of alternative splicing and was proposed to induce RNA looping through its two interacting C-terminal RRM domains. As part of a multidisciplinary approach, we conducted a functional assay to observe this mechanism *in vivo*. Together with NMR spectroscopy and FRET data from our collaborators, we were finally able to support the thesis of PTB-induced looping of exons to repress alternative splicing.

The third RNA-binding protein we investigated is the germline defective 1 (GLD-1), a *C. elegans* protein involved in the regulation of translation. By solving the solution structure of its KH-QUA2 domain in complex with its native RNA target, we were able to understand its specificity for RNA on a molecular level. Interestingly, its additional QUA2 domain seems to modulate RNA-binding affinity and specificity within the whole STAR protein family.

In summary, the structural and functional studies of three different RNA-binding proteins in this thesis describe novel features of RNA recognition and their important implications on cellular homeostasis.

Zusammenfassung

Die Entstehung der Eukaryoten war ein Meilenstein in der Evolution von mehrzelligen Organismen. Sie wurde begleitet von der Entwicklung komplexer intrazellulärer Netzwerke, um eine spezifische und exakte Regulation der Genexpression zu ermöglichen. Die post-transkriptionale Prozessierung bestimmt die Zuverlässigkeit und Vielfalt der Genexpression und ihre straffe Regulierung wird von einer großen Anzahl von RNA-bindenden Proteinen begleitet. Schon winzige Defekte wie eine einzige Nukleotid-Substitution können schwerwiegende Auswirkungen auf Protein-RNA Interaktionen innerhalb dieser Regulations-Maschinerie haben und letztendlich zu ernsthaften Krankheiten führen. Deshalb ist es unerlässlich die für die post-transkriptionale Regulation wichtigen Protein-RNA Interaktionen auf molekularer Ebene zu untersuchen und ihre Bedeutung für die Homeostase der Zelle zu ermitteln. In dieser Doktorarbeit haben wir die Funktion von drei RNA-bindenden Proteinen auf molekularer Ebene untersucht, indem wir die Struktur in Lösung der jeweiligen Protein-RNA Komplexverbindung bestimmt und/ oder funktionelle Experimente durchgeführt haben.

Der Serin-/ Arginin-reiche Spleiß-faktor 2 (SRSF2) ist ein Spleiß-regulator und es ist bekannt, dass dieser stark degenerierte RNA Sequenzen bindet. Um die Spezifität von SRSF2 herauszufinden, haben wir die Struktur des RNA-bindenden Motifs (RRM) mit zwei unterschiedlichen RNAs gelöst. Diese Strukturen zeigen, dass RNA von SRSF2 in einer neuartigen Sandwich-ähnlichen Struktur unter Beteiligung beider N- und C-terminalen Enden gebunden wird. Zusammen mit thermodynamischen Daten und zwei unterschiedlichen funktionellen Test konnten wir letztendlich die präzise RNA Spezifität und Konsensus-Sequenz von SRSF2 für RNA bestimmen.

Im Gegensatz zu SRSF2 ist das Polypyrimidin-Sequenzen bindende Protein (PTB) hauptsächlich ein Repressor von alternativen Spleißen und es wurde vorgeschlagen, dass es durch seine beiden C-terminalen RRM Domänen Schlingenbildung von RNA induziert. Als Teil eines multidisziplinären Ansatzes haben wir einen funktionellen Test ausgeführt, um diesen Mechanismus *in vivo* nachzuweisen. Zusammen mit Daten von NMR Spektroskopie und FRET Experimenten, durchgeführt von unseren Kollaborateuren, konnten wir letztendlich die These unterstützen, dass PTB induzierte Schlingenbildung von Exonen zu einer Unterdrückung von Alternativen Spleißen führen kann.

Das dritte zu untersuchende RNA-bindende Protein ist das Keimbahn defektiv 1 (GLD-1). Es handelt sich dabei um ein *C. elegans* Protein, welches in der translationalen Regulation involviert ist. Indem wir die Struktur von der KH-QUA2 Domäne in Verbindung mit ihrer natürlichen Ziel-RNA mittels NMR Spektrometrie gelöst haben, konnten wir deren Spezifität für RNA auf einer molekularen Ebene verstehen. Interessanterweise scheint die

zusätzliche QUA2 Domäne RNA Affinität und Spezifität der gesamten STAR Protein Familie zu modulieren.

In dieser Doktorarbeit werden anhand struktureller und funktioneller Studien von drei unterschiedlichen RNA-bindenden Proteinen neuartige Eigenschaften der RNA Erkennung und deren wichtige Folgen für die zelluläre Homeostase beschrieben.