

RNA Protein Interaction

Axel Wintsche

February 16, 2009

Table of Contents

- 1 Introduction
- 2 RNA binding in detail
- 3 structural analysis of local contexts of consensus sequences
- 4 Examples

Introduction

- RNA is more than an intermediate step in protein expression
- not every RNA is translated (ncRNAs)
- RNA can be linked with regulatory processes in cell
- UTRs of some mRNAs have conserved regions
- carry information important for fitness

Introduction

- one way to participate in regulatory mechanisms is to bind to a protein
 - can induce conformational changes and results in new functionality
 - RNA-protein binding leads to direct regulation of the RNA (translational repression of mRNA)
- ⇒ understanding of RNA-protein binding is important to understand many regulatory processes

Introduction

- one way to participate in regulatory mechanisms is to bind to a protein
 - can induce conformational changes and results in new functionality
 - RNA-protein binding leads to direct regulation of the RNA (translational repression of mRNA)
- ⇒ understanding of RNA-protein binding is important to understand many regulatory processes
- prediction of RNA-protein interaction is hard to compute

Introduction

- How recognise proteins their RNA targets?
- Is RNA sequence and/or RNA structure important for specific binding?

Two Examples of RNA binding protein domains

RRM: RNA Recognition Motif

The information of structure, RNA binding and pictures are taken from [4] and [1]

KH: K Homology Domain

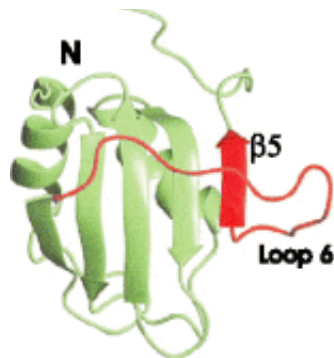
The information of structure, RNA binding and pictures are taken from [5] and [2]

RNA Recognition Motif

- RRM can be found in all eukaryotic organisms
- one of the most widely spread domains in eukaryots (H. sapiens has at least 497 proteins)
- often more than one RRM per protein
- identified in the 80s while investigating mRNA precursors
- RRM is part of many different functions in cell

RNA Recognition Motif

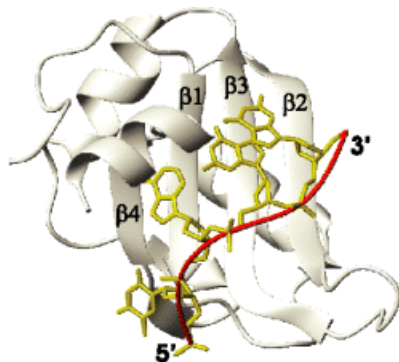
- RRM has a length of about 90 AA
- it folds into an $\alpha\beta$ sandwich ($\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$)
- spacially arranged in one four-stranded antiparallel β -sheet ($\beta_4\beta_1\beta_3\beta_2$)



PTB-RRM3

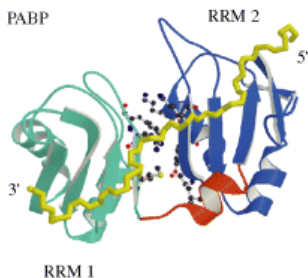
RNA Recognition Motif

- RNA binds mainly to the β -sheet
- 5' end is located on $\beta_4\beta_1$, 3' end on $\beta_3\beta_2$
- sequence specificity by sidechain interaction
- β -sheet detects only a dinucleotide



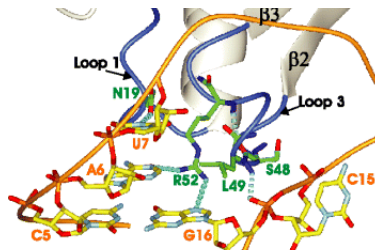
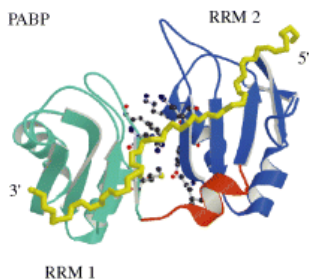
RNA Recognition Motif

- additional RRMs allow recognition of longer sequences



RNA Recognition Motif

- additional RRMs allow recognition of longer sequences
- RRM of UA1 needs secondary RNA structure to bind (stem loop or internal loop)

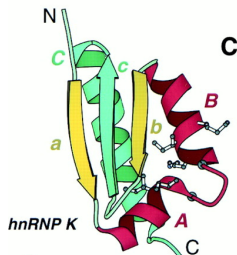


K homology domain

- first discovered in K-proteins (human hnRNP)
- KH is a common and widespread domain
- KH exists in all three kingdoms and hence arose early in evolution
- most KH-proteins have more than one KH domain

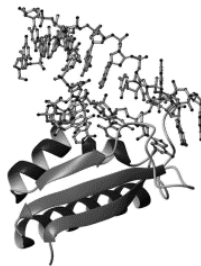
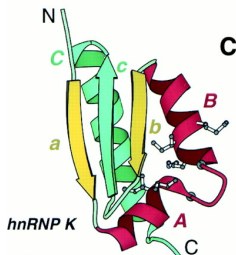
K homology domain

- KH has a length of about 70 AA
- fold into $\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3$
- spacially forming a $\beta_1\beta_3\beta_2$ -sheet



K homology domain

- KH has a length of about 70 AA
- fold into $\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3$
- spacially forming a $\beta_1\beta_3\beta_2$ -sheet
- binds RNA in a different manner to RRM, suggesting that the structural convergence of RNA binding proteins has little to do with RNA recognition



Materials and Software

Materials

- Proteins and related RNA consensus sequences [3]
- yeast genome from Genbank
- yeast mRNA annotations from Genbank

Materials and Software

Materials

- Proteins and related RNA consensus sequences [3]
- yeast genome from Genbank
- yeast mRNA annotations from Genbank

Software

- `pwmatch`: find sequence matches in a genome (threshold parameter 1 = exact match)
- `RNAfold`: predict RNA secondary structure

first simple approach

- 1 translate rna consensus sequence into PWM
- 2 search matches on genome with `pwmacth`
- 3 select matches laying in annotated mRNAs
- 4 cut out matches plus offset
- 5 fold sequences with `RNAfold`

first simple approach

- 1 translate rna consensus sequence into PWM
- 2 search matches on genome with `pwmatch`
- 3 select matches laying in annotated mRNAs
- 4 cut out matches plus offset
- 5 fold sequences with `RNAfold`

Example of consensus sequence to PWM conversion

HUCNBCAWH →

A:	1	0	0	1	0	0	1	1	1
C:	1	0	1	1	1	1	0	0	1
G:	0	0	0	1	1	0	0	0	0
U:	1	1	0	1	1	0	0	1	1

Example Proteins

SCP160

- function: involved in the control of mitotic chromosome transmission
- contains 7 KH domains
- RNA consensus sequence: UUUAAUGA

Example Proteins

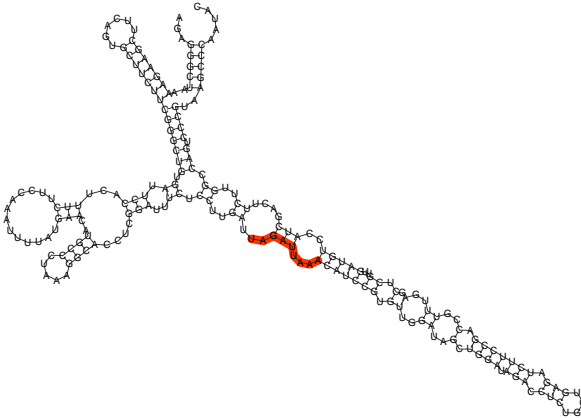
SCP160

- function: involved in the control of mitotic chromosome transmission
- contains 7 KH domains
- RNA consensus sequence: UUUAAUGA

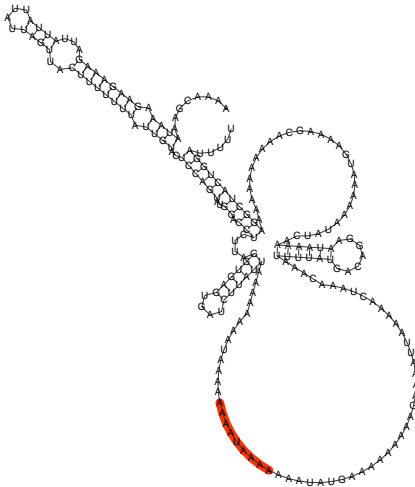
PIN4

- function: involved in normal G2/M phase transition of the mitotic cell cycle
- contains 1 RRM domain
- RNA consensus sequence: KUUWAYBUN

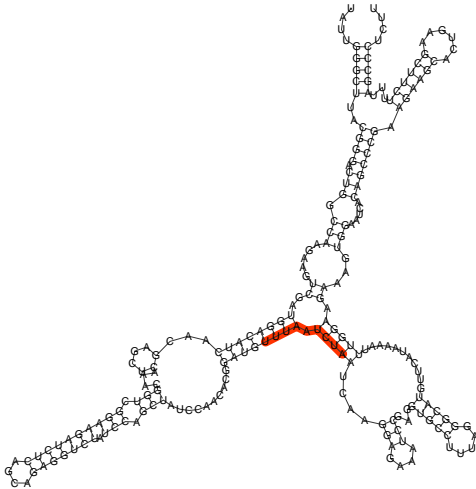
Examples SCP160



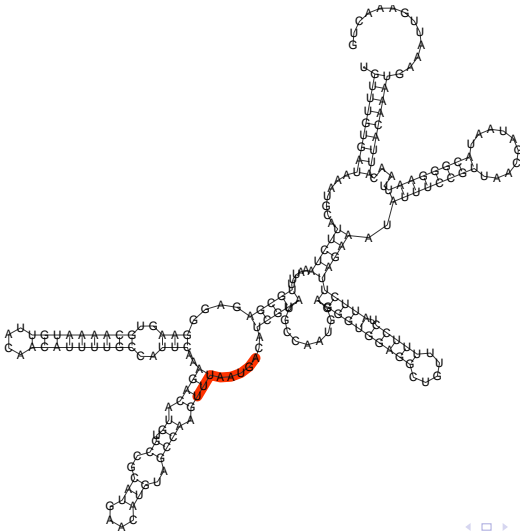
Examples SCP160



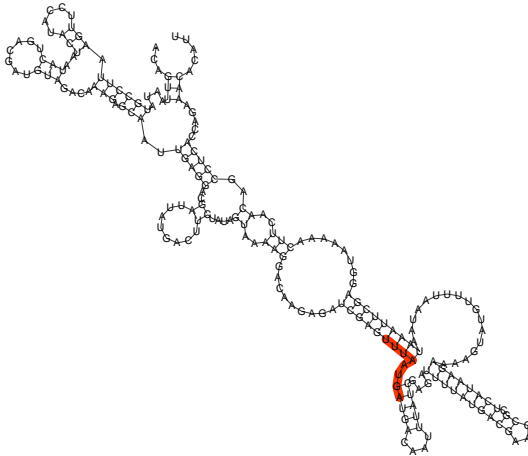
Examples SCP160



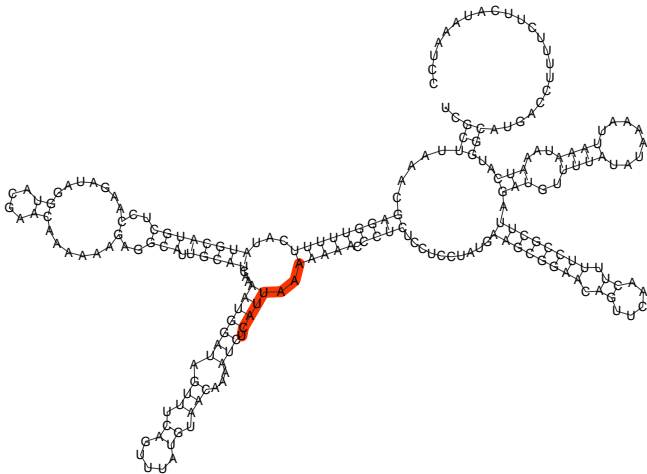
Examples PIN4



Examples PIN4



Examples PIN4



Literature



A. Cléry, M. Blatter, and F.H.T. Allain.

RNA recognition motifs: boring? Not quite.

Current Opinion in Structural Biology, 18(3):290–298, 2008.



N.V. Grishin.

KH domain: one motif, two folds.

Nucleic Acids Research, 29(3):638–643, 2001.



D.J. Hogan, D.P. Riordan, A.P. Gerber, D. Herschlag, and P.O. Brown.

Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System.

PLoS Biol, 6(10):e255, 2008.



C. Maris, C. Dominguez, and F.H.T. Allain.

The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression.

FEBS Journal, 272(9):2118–2131, 2005.



J.M. Pérez-Cañadillas and G. Varani.

Recent advances in RNA–protein recognition.

Current Opinion in Structural Biology, 11(1):53–58, 2001.

Thank you for listening!