

3'Processing of messenger RNA in the yeast *Saccharomyces cerevisiae*

Doctoral Thesis

Author(s):

Egli, Christoph Mathias

Publication date:

1995

Permanent link:

<https://doi.org/10.3929/ethz-a-001459754>

Rights / license:

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


18. Juli 1995

3' Processing of Messenger RNA in the Yeast

Saccharomyces cerevisiae

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH
for the degree of
DOCTOR OF NATURAL SCIENCES

presented by
CHRISTOPH MATHIAS EGLI
Dipl. Natw. ETH
born on October 6, 1964
citizen of Kirchberg SG



1. Jul. 1995

accepted on the recommendation of
Prof. Dr. R. Hütter, examiner
Prof. Dr. G. Braus, co-examiner
Prof. Dr. H. Hennecke, co-examiner

Zürich 1995

Summary

The process of 3' end formation of mRNA transcripts was investigated in the yeast *Saccharomyces cerevisiae*. 3' Processing consists of two steps i. e. cleavage of the primary transcript and addition of a poly(A) tail to the 5' cleavage product (polyadenylation). The sequences required for recognition of the position where cleavage takes place are unique in higher eukaryotes but seem to be degenerate in yeast. Earlier investigations revealed that polyadenylation elements in *Saccharomyces cerevisiae* can be grouped into two classes. (I) Elements of one class direct 3' processing in a unidirectional manner in an *in vivo* test system and are efficient even in combination with a strong promoter. (II) Representatives of the other class act bidirectionally and less efficiently.

The GCN4 polyadenylation element was analyzed in more detail. 3' Processing experiments confirmed that the GCN4 polyadenylation element belongs to the class of unidirectional and efficient elements. A large minimal polyadenylation element of 116 nucleotides was found to be necessary for full 3' processing activity. This element was subdivided by deletion analysis into five distinct regions which were named EL1 to EL5. Two of them (EL1 and EL4) contained the heptanucleotide motif TTTTAT, which was shown to be present in all elements of class I polyadenylation elements. These two motifs are essential for full 3' processing efficiency. Either deletion or point mutations within these motifs strongly impaired 3' processing efficiency. The major poly(A) addition site which is used in at least 90% of the wildtype mRNAs could be deleted without a decrease in 3' processing activity when the correct upstream sequences were present. However, the poly(A) addition site was shifted to cryptic downstream cleavage sites if 15 nucleotides (EL3) between the two TTTTAT sequence motifs were deleted indicating that well-defined sequences are responsible for exact positioning of the poly(A) addition site. The tripartite sequence TAG...TATGT...TTT which is required in many other polyadenylation elements was without 3' processing activity in this polyadenylation element. This confirmed the hypothesis of different classes of polyadenylation elements.

The TRP4 polyadenylation element was taken as representative of the class of bidirectional and weak polyadenylation elements (class II) for a detailed investigation. Three sequence motifs TA(T)GT(T/A), on the sense DNA strand TATGTT and on the nonsense DNA strand TAGTA and TAGTT,

respectively were shown by deletion and mutation analysis to be involved in *TRP4* 3' processing. The motif TAGTTT on the nonsense strand had the strongest influence in this respect. Activity was completely abolished when a single point mutation was introduced within this sequence motif. Most modifications, either deletions or point mutations within these motifs had similar effects on 3' processing whether the mutated polyadenylation element was inserted in the positive or in the negative orientation into the test gene. An RNA hairpin secondary structure with the essential TAGTTT motif located in the hairpin loop was found by computer analysis and could be involved in poly(A) site recognition. The symmetrical core element was only fully active if additional upstream sequences were present. For the positive orientation the T nucleotide in the *TRP4* open reading frame at position -17 relative to the translational stop codon was essential whereas sequences between the positions -161 and +70 were required as upstream sequences in case of the negative orientation. Since so far no *in vitro* processing was detected for the *TRP4* polyadenylation element, it cannot be excluded that *TRP4* mRNA 3' end formation is performed in a process of termination and polyadenylation as it is described for prokaryotic mRNA 3' ends.

The processes of cleavage and polyadenylation are generally conserved in the eukaryotic kingdom. However, there are differences in the sequence requirements or the involved protein factors in this reaction. The reaction steps, cleavage and polyadenylation can be uncoupled *in vivo* in the yeast *Saccharomyces cerevisiae*. This was shown by introduction of a heterologous ribozyme derived from tobacco ringspot virus into an artificial yeast test gene. Very efficient cleavage was observed using this sequence. The transcript was polyadenylated, although the transcript was cleaved independently of the cellular apparatus. The poly(A) tail length was similar to the poly(A) tail length of an endogenous yeast mRNA. The fact that cleavage and polyadenylation can be separated *in vivo* is in contrast to the situation in higher eukaryotes. A possible reason for this is that the poly(A)osome in mammals is a more stable and tight complex than in yeast.

Kurzfassung

In dieser Arbeit wurde in der Hefe *Saccharomyces cerevisiae* die mRNA 3'Prozessierung untersucht. Diese Reaktion kann funktionell in zwei Teilschritte unterteilt werden: Spaltung des primären Transkriptes und Polyadenylierung des 5'Spaltproduktes. Für die Erkennung der Stelle, an der das Transkript gespalten wird, sind Sequenzen notwendig, die in höheren Eukaryonten stark konserviert, in der Hefe jedoch degeneriert sind. Frühere Untersuchungen haben ergeben, dass die Polyadenylierungselemente in *Saccharomyces cerevisiae* in zwei Klassen eingeteilt werden können: (I) Polyadenylierungselemente, die in einem *in vivo* Testsystem unidirektionell und effizient funktionieren, und (II) Elemente, die bidirektionell und weniger effizient wirken.

Das GCN4 Polyadenylierungselement gehört zur Klasse der unidirektionellen und effizienten Elemente, was mittels *in vitro* 3'Prozessierungsexperimente bestätigt werden konnte. Ein Minimalelement von 116 Nukleotiden wird für eine vollständige 3'Prozessierungseffizienz benötigt. Dieses Minimalelement wurde durch Deletionsanalyse in fünf Regionen (EL1 bis EL5) unterteilt, von denen die Regionen EL1 und EL4 das Heptanukleotid-Motiv TTTTAT enthalten. Diese zwei Motive sind essentiell für eine vollständige 3'Prozessierungsaktivität. Sowohl Deletion als auch Mutation dieser Motive verursachte einen starken Aktivitätsabfall. Im Wildtyp-Polyadenylierungselement wird bei der Spaltung zu über 90% eine Polyadenylierungsstelle benützt. Diese konnte ohne Aktivitätsverlust deletiert werden, wenn Sequenzen 'upstream' des Minimalelementes vorhanden waren. Deletion von 15 Nukleotiden (EL3) zwischen den beiden TTTTAT Motiven führte zu einer Verschiebung der dominanten Wildtyp-Polyadenylierungsstelle zu kryptischen 'downstream' gelegenen Spaltstellen. Folglich sind definierte Sequenzen für die exakte Positionierung der Polyadenylierungsstelle nötig. Die dreiteilige Sequenz TAG...TATGT...TTT, die in zahlreichen Polyadenylierungsstellen benötigt wird, ist hier ohne offensichtliche Funktion.

Aus der Klasse II der bidirektionellen und schwachen Polyadenylierungselemente wurde diejenige vom TRP4 Gen ausgewählt. Mittels Deletionsanalyse und Punktmutagenese konnte gezeigt werden, dass alle drei Sequenzmotive TA(T)GT(T/A), auf dem kodierenden DNA Strang TATGTT und auf dem nichtkodierenden DNA Strang TAGTA und TAGTTT,

für die *TRP4* 3'Prozessierung benötigt werden. Das Sequenzmotiv TAGTTT auf dem nichtkodierenden DNA Strang hat den stärksten Einfluss auf die 3'Prozessierung. Einzelne Punktmutationen bewirkten einen vollständigen Aktivitätsverlust. Alle Modifikationen in diesen Elementen beeinträchtigten die Effizienz der 3'Prozessierung qualitativ gleich, unabhängig davon, ob sie in positiver oder negativer Orientierung getestet wurden. Eine RNA-'hairpin'-Sekundärstruktur wurde mit Hilfe einer Computeranalyse gefunden und könnte an der Erkennung der Polyadenylierungsstelle beteiligt sein. Diese Annahme wird insofern erhärtet, als das essentielle Sequenzmotiv TAGTTT bei dieser Struktur genau im 'hairpin loop' liegt. Das symmetrische Element war nur in Kombination mit zusätzlichen 'upstream' gelegenen Sequenzen effizient. Im Fall der positiven Orientierung ist hierfür das T-Nukleotid im 'open reading frame' an der Position -17 relativ zum Translationsstopcodon essentiell. In der negativen Orientierung waren dazu Sequenzen in der Region zwischen den Positionen + 161 und + 70 erforderlich. Bislang konnte für das *TRP4* Polyadenylierungselement mittels *in vitro* Experimenten keine 3'Prozessierung nachgewiesen werden, und es ist möglich, dass das *TRP4* 3'Ende wie bei Prokaryonten durch Termination und Polyadenylierung gebildet wird.

Der Vorgang der 3'Prozessierung ist in Eukaryonten im Prinzip konserviert. Es treten in den Erkennungssequenzen und beteiligten Proteinfaktoren jedoch Unterschiede auf. Die beiden Teilprozesse Spaltung und Polyadenylierung lassen sich in der Hefe *Saccharomyces cerevisiae in vivo* trennen. Ein aus einem Tabakvirus abgeleitetes heterologes Ribozym wurde in ein künstliches Hefe-Testgen gebracht. Dabei war eine vollständige Transkriptspaltung zu beobachten. Das entstandene Transkript wurde polyadenyliert, obwohl die primäre mRNA durch einen zellfremden Mechanismus gespalten wurde. Die Länge des Poly(A) Schwanzes war mit derjenigen eines endogenen Transkriptes vergleichbar. Die Tatsache, dass die beiden Reaktionsschritte Spaltung des primären Transkriptes und Polyadenylierung *in vivo* voneinander getrennt werden können, steht im Gegensatz zur Situation in höheren Eukaryonten. Es lässt sich vermuten, dass die Komponenten des Poly(A)osoms in der Hefe lockerer aneinander gebunden sind als in höheren Eukaryonten.