Transcription of the HIV genome

Matthews, L., Peterlin, BM., Rice, AP.

European Bioinformatics Institute, New York University Langone Medical Center, Ontario Institute for Cancer Research, Oregon Health and Science University.

The contents of this document may be freely copied and distributed in any media, provided the authors, plus the institutions, are credited, as stated under the terms of Creative Commons Attribution 4.0 International (CC BY 4.0) License. For more information see our license.

This is just an excerpt of a full-length report for this pathway. To access the complete report, please download it at the Reactome Textbook.

27/12/2022
Introduction

Reactome is open-source, open access, manually curated and peer-reviewed pathway database. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff and cross-referenced to many bioinformatics databases. A system of evidence tracking ensures that all assertions are backed up by the primary literature. Reactome is used by clinicians, geneticists, genomics researchers, and molecular biologists to interpret the results of high-throughput experimental studies, by bioinformaticians seeking to develop novel algorithms for mining knowledge from genomic studies, and by systems biologists building predictive models of normal and disease variant pathways.

The development of Reactome is supported by grants from the US National Institutes of Health (P41 HG003751), University of Toronto (CFREF Medicine by Design), European Union (EU STRP, EMI-CD), and the European Molecular Biology Laboratory (EBI Industry program).

Literature references


Reactome database release: 83

This document contains 10 pathways and 2 reactions (see Table of Contents)

https://reactome.org
Expression of the integrated HIV-1 provirus is dependent on the host cell Pol II transcription machinery, but is regulated in critical ways by HIV-1 Tat and Rev proteins. The long terminal repeats (LTR) located at either end of the proviral DNA contain regulatory sequences that recruit cellular transcription factors. The U3 region of the 5' LTR contains numerous cis-acting elements that regulate Pol II-mediated transcription initiation. The full-length transcript, which encodes nine genes, functions as an mRNA and is packaged as genomic RNA. Smaller (subgenomic) viral mRNAs are generated by alternative splicing. The activities of Tat and Rev create two phases of gene expression (see Karn 1999; Cullen 1991). The Tat protein is an RNA specific trans-activator of LTR-mediated transcription. Association of Tat with TAR, a RNA stem-loop within the RNA leader sequence, is required for efficient elongation of the HIV-1 transcript. In the early phase of viral transcription, a multiply-spliced set of mRNAs is generated, producing the transcripts of the regulatory proteins, Tat, Rev, and Nef. In the late phase, Rev regulates nuclear export of HIV-1 mRNAs, repressing expression of the early regulatory mRNAs and promoting expression of viral structural proteins. Nuclear export of the unspliced and partially spliced late HIV-1 transcripts that encode the structural proteins requires the association of Rev with a cis-acting RNA sequence in the transcripts (Rev Response Element, RRE).

**Literature references**


<table>
<thead>
<tr>
<th>Editions</th>
<th>Date</th>
<th>Action</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2005-01-05</td>
<td>Reviewed</td>
<td>Peterlin, BM.</td>
</tr>
<tr>
<td></td>
<td>2005-07-27</td>
<td>Authored</td>
<td>Matthews, L., Rice, AP.</td>
</tr>
<tr>
<td></td>
<td>2005-07-27</td>
<td>Edited</td>
<td>Matthews, L.</td>
</tr>
</tbody>
</table>
HIV Promoter Opening: First Transition

Location: Transcription of the HIV genome

Stable identifier: R-HSA-167097

Type: transition

Compartments: nucleoplasm

Diseases: Human immunodeficiency virus infectious disease

Inferred from: RNA Polymerase II Promoter Opening: First Transition (Homo sapiens)

After assembly of the complete RNA polymerase II-preinitiation complex, the next step is separation of the two DNA strands. This isomerization step is known as the closed-to-open complex transition and occurs prior to the initiation of mRNA synthesis. In the RNA polymerase II system this step requires the hydrolysis of ATP or dATP into Pi and ADP or dADP (in contrast to the other RNA polymerase systems) and is catalyzed by the XPB subunit of TFIIH. The region of the promoter, which becomes single-stranded, spans from –10 to +2 relative to the transcription start site.

Negative supercoiling in the promoter region probably induces transient opening events and can alleviate requirement of TFIIE, TFIIH and ATP-hydrolysis for open complex formation. ATP is also used in this step by the cdk7-subunit of TFIIH to phosphorylate the heptad repeats of the C-terminal domain of the largest subunit of RNA polymerase II (RPB1) on serine-2

Followed by: Fall Back to Closed Pre-initiation Complex, HIV Transcription Initiation

Literature references


**Fall Back to Closed Pre-initiation Complex**

**Location:** Transcription of the HIV genome

**Stable identifier:** R-HSA-167484

**Type:** transition

**Compartments:** nucleoplasm

**Diseases:** Human immunodeficiency virus infectious disease

**Inferred from:** Fall Back to Closed Pre-initiation Complex (Homo sapiens)

At the beginning of this reaction, 1 molecule of 'HIV-1 open pre-initiation complex' is present. At the end of this reaction, 1 molecule of 'HIV-1 closed pre-initiation complex' is present.

This reaction takes place in the 'nucleus'.

**Preceded by:** HIV Promoter Opening: First Transition

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-07-27</td>
<td>Authored</td>
<td>Matthews, L., Rice, AP.</td>
</tr>
<tr>
<td>2005-10-16</td>
<td>Edited</td>
<td>Matthews, L.</td>
</tr>
</tbody>
</table>
Formation of the open complex exposes the template strand to the catalytic center of the RNA polymerase II enzyme. This facilitates formation of the first phosphodiester bond, which marks transcription initiation. As a result of this, the TFIIB basal transcription factor dissociates from the initiation complex.

The open transcription initiation complex is unstable and can revert to the closed state. Initiation at this stage requires continued (d)ATP-hydrolysis by TFIIH. Dinucleotide transcripts are not stably associated with the transcription complex. Upon dissociation they form abortive products. The transcription complex is also sensitive to inhibition by small oligo-nucleotides.

Dinucleotides complementary to position -1 and +1 in the template can also direct first phosphodiester bond formation. This reaction is independent on the basal transcription factors TFIIE and TFIIH and does not involve open complex formation. This reaction is sensitive to inhibition by single-stranded oligonucleotides.
RNA Polymerase II HIV Promoter Escape

Location: Transcription of the HIV genome

Stable identifier: R-HSA-167162

Compartments: nucleoplasm

Diseases: Human immunodeficiency virus infectious disease

Inferred from: RNA Polymerase II Promoter Escape (Homo sapiens)

RNA Polymerase II promoter escape occurs after the first phosphodiester bond has been created.

Editions

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-07-27</td>
<td>Authored</td>
<td>Matthews, L., Rice, AP.</td>
</tr>
<tr>
<td>2005-07-27</td>
<td>Edited</td>
<td>Matthews, L.</td>
</tr>
</tbody>
</table>
To facilitate co-transcriptional capping, and thereby restrict the cap structure to RNAs made by RNA polymerase II, the capping enzymes bind directly to the RNA polymerase II. The C-terminal domain of the largest Pol II subunit contains several phosphorylation sites on its heptapeptide repeats. The capping enzyme guanylyltransferase and the methyltransferase bind specifically to CTD phosphorylated at Serine 5 within the CTD. Kinase subunit of TFIH, Cdk7, catalyzes this phosphorylation event that occurs near the promoter. In addition, it has been shown that binding of capping enzyme to the Serine-5 phosphorylated CTD stimulates guanylyltransferase activity in vitro.
In the absence of the HIV-1 protein Tat, transcription of the proviral DNA is inefficient and results in the production of truncated transcripts (Kao et al., 1987). While initiation of transcription from the HIV-1 LTR and formation of the early elongation complex occurs normally, transcription elongation is incomplete with non-processive polymerases disengaging from the proviral DNA template prematurely (reviewed in Karn 1999). The mechanism of Tat-mediated elongation is described below.

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-07-27</td>
<td>Authored</td>
<td>Matthews, L., Rice, AP.</td>
</tr>
<tr>
<td>2005-07-27</td>
<td>Edited</td>
<td>Matthews, L.</td>
</tr>
</tbody>
</table>
RNA Pol II arrest is believed to be a result of irreversible backsliding of the enzyme by ~7-14 nucleotides. TFIIS reactivates arrested RNA Pol II by promoting the excision of nascent transcript ~7-14 nucleotides upstream of the 3' end.
During the formation of the HIV elongation complex in the absence of HIV Tat, elongation factors are recruited to form the HIV-1 elongation complex (Hill and Sundquist 2013) and P-TEFb complex hyperphosphorylates RNA Pol II CTD (Hermann and Rice, 2005, Zhou et al., 2000).
Pausing and recovery of HIV elongation

Location: Transcription of the HIV genome

Stable identifier: R-HSA-167290

Compartments: nucleoplasm

Diseases: Human immunodeficiency virus infectious disease

After Pol II pauses by back tracking 2-4 nucleotides on the HIV-1 template, elongation of the HIV-1 transcript resumes.
RNA Pol II arrest is believed to be a result of irreversible backsliding of the enzyme by ~7-14 nucleotides. TFIIS reactivates arrested RNA Pol II by promoting the excision of nascent transcript ~7-14 nucleotides upstream of the 3' end.
Pausing and recovery of Tat-mediated HIV elongation

**Location:** Transcription of the HIV genome

**Stable identifier:** R-HSA-167238

**Compartments:** nucleoplasm

**Diseases:** Human immunodeficiency virus infectious disease

After Pol II pauses by back tracking 2 -4 nucleotides on the HIV-1 template, elongation of the HIV-1 transcript resumes.
Table of Contents

Introduction 1

Transcription of the HIV genome 2

HIV Promoter Opening: First Transition 4

Fall Back to Closed Pre-initiation Complex 6

HIV Transcription Initiation 7

RNA Polymerase II HIV Promoter Escape 8

RNA Pol II CTD phosphorylation and interaction with CE during HIV infection 9

HIV Transcription Elongation 10

HIV elongation arrest and recovery 11

Formation of HIV elongation complex in the absence of HIV Tat 12

Pausing and recovery of HIV elongation 13

Tat-mediated HIV elongation arrest and recovery 14

Pausing and recovery of Tat-mediated HIV elongation 15

Table of Contents 16