Molecular insights into the mRNA poly(A) tail machinery

Lori Passmore
My lab at LMB

1. mRNA 3’-end processing

- Splicing
- Polyadenylation
- Nuclear export
- Degradation
- Ccr4-Not
- Pan2/3
- 80S
- Translation
My lab at LMB

1. mRNA 3’-end processing
2. poly(A) tail removal
My lab at LMB

1. mRNA 3’-end processing
2. poly(A) tail removal
3. DNA crosslink repair
The length and position of poly(A) tails are important for gene regulation.

Globin mRNA polyA site mutations (LOF) - thalassemia
Prothrombin mRNA polyA site mutations (GOF) - thrombophilia
Alternative polyadenylation - cancer
Disruption of host polyadenylation - influenza A
mRNA 3' end processing

Transcription termination

Poly(A) polymerase
Phosphatase
Nuclease
Identification of enzymatic modules in CPF
Identification of enzymatic modules in CPF

Ashley Easter, Yuliya Gordiyenko, Carol Robinson
Identification of enzymatic modules in CPF

38 sub complexes, computational network analysis
DNA
Pol II
polyadenylation
CPF
cleavage
dephosphorylation
ORF
CPF
CPF
Nuclease
Phosphatase
Poly(A) polymerase

Casañal, Kumar et al, Science 2017 (with Carol Robinson)
The polymerase module

- Bacmid

- Pfs2
- Yth1
- Cft1
- Fip1
- Pap1

Size (kDa)

- Cft1
- Pap1
- Pfs2-SII
- Fip1
- 8His-Yth1

CPSF160
PAP
WDR33
FIP
CPSF30
The polymerase module

~ 200 kDa
No polymerase
Cryo-EM of the polymerase module

Relion 2
79,950 particles
The polymerase module

- 3.5 Å resolution overall
- 3D classification
- Focussed classification for individual domains with heterogeneity/flexibility

Ananth Kumar
Ana Casañal
The polymerase module

~200 kDa

PDB 6EOJ, EMDB 3908, EMPIAR 10299
Casañal, Kumar et al Science 2017
The polymerase module
Interaction of Yth1 with the polymerase module

PDB 6EOJ, EMDB 3908, EMPIAR 10299
Casañal, Kumar et al Science 2017
Interaction of Pfs2 with polymerase module

PDB 6EOJ, EMDB 3908, EMPIAR 10299
Casañal, Kumar et al. Science 2017
Similarity to other nucleic acid binding complexes

mRNA 3' end processing
CPF polymerase module


DNA repair
DDB1–DDB2

Yan et al (2016)
Cretu et al (2016)
Plaschka et al (2017)

pre-mRNA splicing
SF3b

Rse1

Hsh155

Rds3

Cus1

DDB2

DNA

RNA

Pfs2

Cft1

Yth1

DDB1

DNA
Similarity to other nucleic acid binding complexes

mRNA 3' end processing
CPF polymerase module

Yth1
Pfs2
Cft1

putative RNA binding surface

DNA repair
DDB1–DDB2

DDB1
DDB2
Cus1

Rds3
Hsh155

Pre-mRNA splicing
SF3b

Rse1


Cretu et al (2016)
Plaschka et al (2017)

Yan et al (2016)
Pap1 is flexibly tethered
Pap1 is flexibly tethered

Kumar, Yu et al, 2021 Genes Dev
Fip1 is largely disordered

Kumar, Yu et al, 2021 Genes Dev
Mapping binding sites on Fip1

Relative peak intensities

+Yth1

Kumar et al, unpublished
Mapping binding sites on Fip1

Kumar et al, unpublished
Mapping binding sites on Fip1

Kumar et al, unpublished
Yth\textsubscript{1} binding (180–220)
Hamilton & Tong 2020

Low complexity region (110–180)
Ezeokonkwo et al., 2011

Pap\textsubscript{1} binding (60–110)
Meinke et al., 2008
Understanding the dynamics of CPF

Recombinant CPF

Kumar, Yu et al, 2021 Genes Dev
Understanding the dynamics of CPF

Kumar, Yu et al, 2021 Genes Dev
Understanding the dynamics of CPF

Recombinant CPF

Ananth Kumar, Conny Yu, Stefan Freund
Kumar; Yu et al, 2021 Genes Dev
Understanding the dynamics of CPF

Recombinant CPF

Ananth Kumar, Conny Yu, Stefan Freund
Kumar; Yu et al, 2021 Genes Dev
Understanding the dynamics of CPF

<table>
<thead>
<tr>
<th>Size (kDa)</th>
<th>Cft1</th>
<th>Cft2</th>
<th>Ysh1</th>
<th>Pta1</th>
<th>Ref2-SII</th>
<th>Pap1</th>
<th>Pfs2, Mpe1</th>
<th>Pti1, Fip1</th>
<th>Swd2, Glc7</th>
<th>Yth1, Ssu72</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Recombinant CPF

Ananth Kumar, Conny Yu, Stefan Freund
Kumar; Yu et al, 2021 Genes Dev
How is the nuclease activated?

Nuclease

Phosphatase

Poly(A) polymerase
Nuclease module

Poly(A) polymerase module

Phosphatase/APT module
X-ray crystal and cryoEM structure of Ysh1-Mpe1 complex

Ysh1

- β-CASP domain
- metallo-β-lactamase domain (MβL)

Mpe1

- ubiquitin-like domain (UBL)

PDB 6I1D, EMD-0325

Chris Hill, Vytaute Borekaite

Hill et al., Mol Cell 2019
The active site is located at the end of a long tunnel.
The active site is located at the end of a long tunnel.
Cryo-EM of Ysh1-Mpe1-Ipa1 complex

Disordered Mpe1 C-terminal ZnK and RING domains

Disordered Ysh1 C-terminal domain, bound to Yjr141w

cryoEM map

Hill et al., Mol Cell 2019
Reconstitution of cleavage activity by the Ysh1 nuclease

Accessory factors CF IA and CF IB also proposed to play a role in 3’ end processing
Reconstitution of cleavage activity by the Ysh1 nuclease

3'UTR

Substrate

Cleavage products

L (nt) 500 400 300 200 100

(-) 30 60 90 (+) 30 60 90

time (min)

Mw (kDa) 250 150 100 75 50 37 25 20 15

3'UTR

Substrate

Cleavage products

3'UTR

Polyadenylated product

CPF core

Ysh1-Mpe1

Ysh1-Yjr141w

Ysh1-Mpe1-Yjr141w

Cft1

Cft2

Ysh1

Pap1

Pfs2

Mpe1

Fip1

Yjr141w

Yth1

Rna14

Pcf11

Clp1

Rna15

Hrp1

CF IA

CF IB

250 150 100 75 50 37 25 20 15

Mw (kDa)
Reconstitution of cleavage activity

Hill et al., Mol Cell 2019
Reconstitution of cleavage activity

This may explain why cleavage had never been reconstituted with recombinant proteins
Negative Stain EM of Core CPF

Hill et al., Mol Cell 2019
Architecture of core CPF

Hill et al., Mol Cell 2019
Architecture of core CPF

Hill et al., Mol Cell 2019
Juan Rodriguez-Molina

Nuclease module

Mpe1

Cft2

Ysh1

Poly(A) polymerase module

Cft1

Yth1

Pap1

Fip1

Pfs2

Phosphatase/APT module

Glc7

Ref2

Pta1

Pti1

Syc1

Ssu72

Swd2

Vo et al 2001
Di Giammartino et al 2014
Lee and Moore 2014
Baejen et al 2014

UBL
ZnK
RING

Mpe1, 441 aa
Mpe1 binds the polymerase module

Rodriguez-Mollina et al, Mol Cell 2022

2.7 Å resolution
pre-mRNA sensing region (PSR)
Mpe1 is important for cleavage

Rodriguez-Mollina et al
Mol Cell 2022
## Mpe1 is important for cleavage and polyadenylation

<table>
<thead>
<tr>
<th>CPF</th>
<th>CPFΔMpe1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 2 4 6 8 10 12</td>
<td>0 2 4 6 8 10 12</td>
</tr>
</tbody>
</table>

- \( A_{n>500} \)
- \( A_{n~100} \)

Rodriguez-Mollina et al. Mol Cell 2022
Mpe1 is important for cleavage and polyadenylation

Rodriguez-Mollina et al
Mol Cell 2022
Mpe1 is required for transcription termination

Rodriguez-Mollina et al
Mol Cell 2022
Mpe1 is required for transcription termination

Mpe1 is required for transcription termination and polyadenylation. Baejen et al 2017

Rodriguez-Mollina et al Mol Cell 2022
Dynamics likely regulate CPF activation
Dynamics likely regulate CPF activation
Dynamics likely regulate CPF activation
Dynamics likely regulate CPF activation
What about the human complex?

Yeast CPF

Polymerase module

Nuclease module

Phosphatase module

Ysh1
Cft1
Fip1
Pfs2
Pap1

Phosphatases

PNUTS
WDR82
PP1
Tox4?
Ssu72
RBBP6
PAP

Human CPSF

mammalian cleavage factor (mCF)

mammalianPolyadenylation Specificity Factor (mPSF)

CPSF73
CPSF100
CPSF160
hFip1
CPSF30
WDR33

Yth1
Glc7
Swd2
Pti1
Ssu72

Cft2
Mpe1
Ref2

CPSF30
Pta1

Phosphatases

WDR82
PP1
Tox4?
Ssu72
RBBP6
PAP

Yeast CPF

Human CPSF
RBBP6/Mpe1 is not a core subunit of human CPSF

Vytaute Boreikaite

RBBP6/Mpe1 is not a core subunit of human CPSF

Vytaute Boreikaite
RBBP6/Mpe1 is not a core subunit of human CPSF

Vytaute Boreikaite

Boreikaite et al Genes Dev 2022
RBBP6/Mpe1 is not a core subunit of human CPSF

Vytaute Boreikaite

Boreikaite et al Genes Dev 2022
RBBP6 activates human pre-mRNA cleavage
RBBP6 activates human pre-mRNA cleavage

<table>
<thead>
<tr>
<th></th>
<th>CPSF</th>
<th>CStF</th>
<th>CFIIm</th>
<th>RBBP6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Boreikaite et al Genes Dev 2022
RBBP6 activates human pre-mRNA cleavage

<table>
<thead>
<tr>
<th></th>
<th>CPSF</th>
<th>CStF</th>
<th>CFIIm</th>
<th>RBBP6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Boreikaite et al Genes Dev 2022
RBBP6 activates human pre-mRNA cleavage

<table>
<thead>
<tr>
<th></th>
<th>length, nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPSF</td>
<td>+</td>
</tr>
<tr>
<td>CStF</td>
<td>-</td>
</tr>
<tr>
<td>CFIIm</td>
<td>-</td>
</tr>
<tr>
<td>RBBP6</td>
<td>-</td>
</tr>
</tbody>
</table>

Boreikaite et al. Genes Dev 2022
RBBP6 activates human pre-mRNA cleavage
RBBP6 activates human pre-mRNA cleavage
RBBP6 activates human pre-mRNA cleavage

Boreikaite et al. Genes Dev 2022
Reconstitution of 3'-end processing

Yeast

- Cft1
- Cft2
- Pfa1
- Ref2
- Pap1
- Pfs2, Mpe1
- Fip1/Pit1
- Swd2, Glc7
- Yth1
- Ssu72

Human

- CPSF160/ Symplekin
- CPSF100
- CPSF73
- WDR33
- hFip1
- CPSF30
- Pcf11
- Clp1
- RBBP6

Yeast proteins:
- CPF
- CF IA
- CF IB

Human proteins:
- CPSF
- CSTF
- CFIm
- RBBP6

Hill et al, Mol Cell 2019
Boreikaite et al, G&D 2022
Reconstitution of 3'-end processing

Yeast

Cleavage assay

Human

Cleavage assay

Hill et al, Mol Cell 2019

Boreikaite et al, G&D 2022
Juan Rodriguez-Molina
Ana Casañal
Ananthanarayanan Kumar
Chris Hill, Ashley Easter
Vytaute Boreikaite
Manuel Carminati
Eleanor Sheekey, Peter Kubik
Matthias Girbig, Chris Russo
Jana Wolf, Katrin Wiederhold
Holly Fagarasan, Terence Tang
Eva Absmeier, James Stowell
Shabih Shakeel, Pablo Alcon
Tamara Sijack, Cemre Manav

Chris Johnson, Stephen McLaughlin
MRC LMB Biophysics

Jianguo Shi, MRC LMB Baculovirus

Conny Yu, Stefan Freund
MRC LMB NMR facility

Jake Grimmett, Toby Darling
MRC LMB computing

Native Mass Spectrometry: Yuliya Gordiyenko, Carol Robinson
HDX-MS: Sarah Maslen, Mark Skehel
XL-MS: Francis O'Reilly, Juri Rappsilber
EM: Christos Savva, Shaoxia Chen, Giuseppe Cannone, Grigory Sharov
eBIC: Alistair Siebert, Dan Clare, Corey Hecksel
How does CPF interact with Pol II?
Pol II CTD phosphorylation

Initiation
- pre-mRNA
- TSS
- Mediator
- GTFs
- Pol II

Elongation
- CPF
- Polymerase
- PPase
- Nuclease
- Pol II

Termination
- Poly(A) signal
- pre-sn(o)RNA
- APT
- Pol II

5' - 3'
CPF and APT dephosphorylate Pol II CTD
CPF and APT interact directly with Pol II

Phosphatase module is necessary and sufficient for the interaction

Carminati et al, bioRxiv
CPF and APT interact directly with Pol II

<table>
<thead>
<tr>
<th>pull-down</th>
<th>CPF</th>
<th>APT</th>
<th>Pol II</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>input</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

Absorbance at 280 nm (mAU)

Elution volume (ml)

Ref2
Pta1
Rpb2
Rpb1

∆CTD

kDa

Carminati et al, bioRxiv
Pol II CTD is not required

Carminati et al, bioRxiv
Pol II CTD is not required
In vitro transcription is compatible with APT/CPF

Carminati et al, bioRxiv
In vitro transcription is compatible with APT/CPF

Carminati et al, bioRxiv
CPF is not stimulated by Pol II

Carminati et al, bioRxiv
CPF is not stimulated by Pol II

Carminati et al, bioRxiv
APT interacts directly with Pol II
Ref2-Glc7-Swd2 sub complex also interacts with Pol II

Carminati et al, bioRxiv
CPF promotes dimerisation of Pol II

Carminati et al, 2022 bioRxiv
The Pol II dimer is compatible with APT binding

Carminati et al, bioRxiv
The Pol II stalk is required for dimer formation

Carminati et al, bioRxiv
The Pol II stalk is required for dimer formation

Carminati et al, bioRxiv
CTD dephosphorylation promotes Pol II dimerization

Carminati et al, bioRxiv
CTD dephosphorylation promotes Pol II dimerization

Carminati et al, bioRxiv
The Pol II dimer clashes with transcription factors

Carminati et al, bioRxiv
Model for transcription termination

Carminati et al, bioRxiv
Model for transcription termination

Termination competent

Carminati et al, bioRxiv
Juan Rodriguez-Molina
Ana Casañal
Ananthanarayanan Kumar
Chris Hill, Ashley Easter
Vytaute Borekaitė
Manuel Carminati
Eleanor Sheekey, Peter Kubik
Matthias Girbig, Chris Russo
Jana Wolf, Katrin Wiederhold
Holly Fagarasan, Terence Tang
Eva Absmeier, James Stowell
Shabih Shakeel, Pablo Alcon
Tamara Sijack, Cemre Manav
Chris Johnson, Stephen McLaughlin
MRC LMB Biophysics

Jianguo Shi, MRC LMB Baculovirus

Conny Yu, Stefan Freund
MRC LMB NMR facility

Jake Grimmett, Toby Darling
MRC LMB computing

Native Mass Spectrometry: Yuliya Gordyenko, Carol Robinson
HDX-MS: Sarah Maslen, Mark Skehel
XL-MS: Francis O’Reilly, Juri Rappsilber
EM: Christos Savva, Shaoxia Chen, Giuseppe Cannone, Grigory Sharov
eBIC: Alistair Siebert, Dan Clare, Corey Hecksel