Two paths of evolution of RNA polymerases

**Multisubunit RNA polymerases**

- animal viruses (mostly pox-)
- archaea
  - eukaryote nuclei (pol I, II, III) and
  - plant nuclei (pol IV)

**Single subunit nucleic acid polymerases**

- bacteria
  - chloroplasts
- DNAP
- RNAP
- Reverse transcriptase

- mitochondria (and nuclei!)
- some chloroplasts
- linear plasmids of unicellular eukaryotes phage (e.g. T7)

- Eukaryotic cells have 4, 5, 6... RNAPs — of all types
Structural relationships of single subunit polymerases

Polymerase

Template

DNA

RNA

Product

DNA

RNA

Poly I

T7

α

β

T7

HIV-1

MMLV

Polio

Hep
A RNAP largest subunits
B RNAP second-largest subunits

*Diagram showing evolutionary relationships among different species' RNAP second-largest subunits.*
It’s not clear whether Pol IV is a DNA-dependent RNA polymerase (but I bet it is).
Well, I was (most likely) wrong

The common reaction sequence of transcription

- promoter finding
- promoter opening
- abortive initiation
- initial elongation
- (promoter clearance)
- productive elongation
  - pausing
  - hydrolytic retraction
  - clearing the path (chromatin)
- termination
- initiation factors
- elongation factors
- termination factors
The existence of the distant common evolutionary root of the multi-subunit bacterial, eukaryotic (nuclear) and archaeal RNA polymerases has been appreciated for some time, and is made brilliantly clear by structure determinations.

refs: Zhang et al. (Darst), 1999
      Cramer et al. (R. Kornberg), 2001
Fitting the high resolution (X-ray) structure of *T. aquaticus* RNA polymerase into the ~15Å (em-crystallography) structure of *E. Coli* RNA polymerase: mobile domains

Subunits of *Saccharomyces cerevisiae* and *Homo sapiens* RNA polymerase III

<table>
<thead>
<tr>
<th><em>S. cerevisiae</em> RNA Pol III subunits</th>
<th>MW (kD)</th>
<th>Guide for nomenclature</th>
<th>Corresponding <em>S. cerevisiae</em> RNA Pol II subunits</th>
<th><em>H. sapiens</em> RNA Pol III subunits</th>
<th>% amino acid identities between <em>H. sapiens</em> and <em>S. cerevisiae</em> Pol III subunits</th>
</tr>
</thead>
<tbody>
<tr>
<td>C160 (β'-like)</td>
<td>162.1</td>
<td>ScRPC1</td>
<td>RPB1</td>
<td>HsRPC1/RPC155</td>
<td>50% [1356/1391]a</td>
</tr>
<tr>
<td>C128 (β-like)</td>
<td>129.3</td>
<td>ScRPC2</td>
<td>RPB2</td>
<td>HsRPC2</td>
<td>63% [1115/1133]a</td>
</tr>
<tr>
<td>C82</td>
<td>73.6</td>
<td>ScRPC3</td>
<td><strong>HsRPC3/RPC62</strong></td>
<td></td>
<td>22% [163/534]a</td>
</tr>
<tr>
<td>C53</td>
<td>46.6</td>
<td>ScRPC4</td>
<td><strong>HsRPC4/RPC53</strong></td>
<td></td>
<td>28% [134/398]a</td>
</tr>
<tr>
<td>C37</td>
<td>32.1</td>
<td>ScRPC5</td>
<td><strong>HsRPC5</strong></td>
<td></td>
<td>26% [160/708]a</td>
</tr>
<tr>
<td>C34</td>
<td>36.1</td>
<td>ScRPC6</td>
<td><strong>HsRPC6/RPC39</strong></td>
<td></td>
<td>26% [216/316]a</td>
</tr>
<tr>
<td>C31</td>
<td>27.7</td>
<td>ScRPC7</td>
<td><strong>HsRPC7/RPC32</strong></td>
<td></td>
<td>35% [44/223]a</td>
</tr>
<tr>
<td>C25</td>
<td>24.3</td>
<td>ScRPC8</td>
<td>RPB7</td>
<td>HsRPC8</td>
<td>42% [201/204]a</td>
</tr>
<tr>
<td>C17</td>
<td>18.6</td>
<td>ScRPC9</td>
<td>RPB4</td>
<td>HsRPC9/CGRP-RC</td>
<td>30% [122/148]a</td>
</tr>
<tr>
<td>C11</td>
<td>12.5</td>
<td>ScRPC10</td>
<td>RPB9</td>
<td>HsRPC10/RPC11</td>
<td>52% [108/108]a</td>
</tr>
<tr>
<td>AC40 (α-like)</td>
<td>37.6</td>
<td>ScRPAC1</td>
<td>RPB3</td>
<td>HsRPAC1/RPA5,RPA39</td>
<td>47% [287/342]a</td>
</tr>
<tr>
<td>AC19 (α-like)</td>
<td>16.1</td>
<td>ScRPAC2</td>
<td>RPB11</td>
<td>HsRPAC2/RPA9,RPA16</td>
<td>45% [119/133]a</td>
</tr>
<tr>
<td>ABC27</td>
<td>25.1</td>
<td>ScRPABC1</td>
<td>RPB5</td>
<td>HsRPABC1/RPB5,RPB25</td>
<td>42% [207/210]a</td>
</tr>
<tr>
<td>ABC23 (ω-like)</td>
<td>17.9</td>
<td>ScRPABC2</td>
<td>RPB6</td>
<td>HsRPABC2/RPB6,RPB14.4</td>
<td>72% [83/127]a</td>
</tr>
<tr>
<td>ABC14.5</td>
<td>16.5</td>
<td>ScRPABC3</td>
<td>RPB8</td>
<td>HsRPABC3/RPB8,RPB17</td>
<td>35% [147/150]a</td>
</tr>
<tr>
<td>ABC10α</td>
<td>7.7</td>
<td>ScRPABC4</td>
<td>RPB12</td>
<td>HsRPABC4/RPB7.0</td>
<td>52% [42/58]a</td>
</tr>
<tr>
<td>ABC10β</td>
<td>8.2</td>
<td>ScRPABC5</td>
<td>RPB10</td>
<td>HsRPABC5/RPB10,RPB7.6</td>
<td>73% [67/67]a</td>
</tr>
</tbody>
</table>
Using homology and structure modelling to derive a detailed structure of the PolⅢ 11-subunit core
The C-terminal domain of the pol II largest subunit

Consensus:  Y (S) P T (S) P S

Number of repeats:

F. falciparum  20
S. cerevisiae  26
S. pombe      29
C. elegans    33
A. thaliana   41
D. melanogaster  44
H. sapiens    52
Unified two Mg2+-ion catalyzed nucleotide addition mechanism

1. The catalytic mechanism/structure of the active site mechanism
The same mechanism in DNA synthesis

The common reaction sequence of transcription

- promoter finding
- promoter opening
- abortive initiation
- initial elongation
- (promoter clearance)
- productive elongation
- pausing
- hydrolytic retraction
- clearing the path (chromatin)
- termination

initiation factors

elongation factors

termination factors
Core Promoters

**Bacteria**

```
| UP | -35 | -10 |
```

**Eukarya Pol II**

```
| BRE | TATA | INT | PTE | DPE |
```

**Archaea**

```
| BRE | TATA | INT |
```
E. Coli Promotor architecture (2006)
Alternative $\sigma$ are also used for modulating transcription.

Backtracking by RNAP and GreA-GreB transcript induced cleavage
Fig. 4. Evolutional conservation of the catalytic residues in transcript-cleavage factors. (A) Sequence alignment of Gre factors. D41 and E44 residues are marked (19). (B) Structure of TFIIS factor. Cyan, β-strands; gray, loops. Functionally important residues (Asp-261 and Glu-262) are shown in ball-and-stick representation and highlighted. The atomic coordinates were taken from Protein Data Bank (PDB ID code 1TFl). Sequence alignment for the
Figure 2. MccJ25 Binding Could Completely Block the Secondary Channel
(Top) Crystal structure of *Thermus aquaticus* RNAP at 3.3 Å resolution (Zhang et al., 1999). The bridge helix is shown in orange, G loop is shown in pink, and active site Mg^{2+} is depicted as a red sphere. The green contour encompasses RNAP elements that together form the secondary channel.
(Bottom left) The part of RNAP highlighted by the green contour at the top of the figure is shown in spacefill next to a spacefill model of the MccJ25 structure (Wilson et al., 2003) shown in blue.
(Bottom right) The MccJ25 structure docked within the RNAP secondary channel. A similar model could be built using the *T. thermophilus* RNAP holoenzyme structure (Vassylyev et al., 2002).

Figure 4. MccJ25 Obstructs the RNAP Secondary Channel

(A) Structure of MccJ25 (Bayro et al., 2003; PDB accession 1PP5; see also Rosengren et al., 2003; Wilson et al., 2003).

(B and C) Model for the structure of the MccJ25-RNAP complex (view orientation as in Figures 2 and 3). Green van der Waals surface, MccJ25 (oriented as in [A]); red and pink van der Waals surfaces, sites of substitutions in β’ and β that confer MccJ25 resistance; yellow labels, sites of substitutions in β’ that confer highest-level MccJ25 resistance (MBC = 1 mg/ml).

Mechanisms of promoter co-dependent on two activator proteins