Supplementary information

Structures of mammalian RNA polymerase II pre-initiation complexes

In the format provided by the authors and unedited
### Data collection and pre-processing

<table>
<thead>
<tr>
<th>Magnification</th>
<th>81,000x</th>
<th>81,000x</th>
<th>81,000x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage (kV)</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Electron exposure (e⁻/Å²)</td>
<td>41.1</td>
<td>44.8</td>
<td>41.5</td>
</tr>
<tr>
<td>Defocus range (µm)</td>
<td>0.2-3.0</td>
<td>0.2-3.0</td>
<td>0.2-3.0</td>
</tr>
<tr>
<td>Pixel size (Å)</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>Initial particle images (no.)</td>
<td>1,535,552</td>
<td>6,106,191</td>
<td>5,277,512</td>
</tr>
</tbody>
</table>

### Model/Map Name

- **Proximal cPIC**
  - **EMDB Code**: 12611 (Proximal cPIC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 63,190
  - **Map resolution (Å)**: 2.8 (Proximal cPIC)
  - **Mean B factors (Å²)**: 24.69

- **Distal cPIC**
  - **EMDB Code**: 12612 (Distal cPIC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 50,177
  - **Map resolution (Å)**: 2.4 (Consensus cPIC)
  - **Mean B factors (Å²)**: 56.24

- **TFIIH (No ADP·BeF₃)**
  - **EMDB Code**: 12615 (Overall)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 255,110
  - **Map resolution (Å)**: 4.3 (Overall)
  - **Mean B factors (Å²)**: 24.69

- **OC cPIC**
  - **EMDB Code**: 12616 (Overall)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 399,247
  - **Map resolution (Å)**: 2.5 (Overall)
  - **Mean B factors (Å²)**: 56.24

- **TFIIH (+ADP·BeF₃)**
  - **EMDB Code**: 12617 (Overall)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 15,226
  - **Map resolution (Å)**: 2.5 (Overall)
  - **Mean B factors (Å²)**: 56.24

- **Proximal CC**
  - **EMDB Code**: 12620 (Consensus cPIC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 271,685
  - **Map resolution (Å)**: 2.4 (Consensus cPIC)
  - **Mean B factors (Å²)**: 56.24

- **Distal CC**
  - **EMDB Code**: 12621 (Consensus cPIC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 255,110
  - **Map resolution (Å)**: 3.9 (Overall)
  - **Mean B factors (Å²)**: 56.24

- **OC**
  - **EMDB Code**: 12622 (Consensus UC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 714,395
  - **Map resolution (Å)**: 2.8 (Overall)
  - **Mean B factors (Å²)**: 56.24

- **IC**
  - **EMDB Code**: 12623 (Consensus UC)
  - **Symmetry imposed**: C1
  - **Final particle images (no.)**: 71,215
  - **Map resolution (Å)**: 4.1 (Overall)
  - **Mean B factors (Å²)**: 56.24

### Refinement

- **Initial model used (PDB code)**: 5FLM
- **Model composition**:
  - **Non-hydrogen atoms**: 44,073
  - **Protein residues**: 5,249
  - **Nucleotides**: 98
  - **Ligands**: ZN: 10, MG: 1
- **Mean B factors (Å²)**:
  - **Protein**: 24.69
  - **Nucleotides**: 62.46
  - **Ligand**: 40.35

### Statistics

- **R.m.s. deviations**:
  - **Bond lengths (Å)**: 0.003
  - **Bond angles (°)**: 0.453
  - **Validation**:
    - **MolProbity score**: 1.31
    - **Ramachandran plot**:
      - **Favored (%)**: 97.53
      - **Allowed (%)**: 2.47
      - **Disallowed (%)**: 0.00

**Supplementary Table 1 | Cryo-EM data collection and processing information**
Supplementary Figure 1 | Raw, uncropped images of gels
Uncropped images of all gels from SDS-PAGE analysis presented in this study. Molecular weight markers are indicated and the boxed area represent where the gel was cropped for presentation in the manuscript.
Supplementary Figure 2 | Further analysis of free-TFIIB compared to the PIC incorporated form

Superposition of the free-TFIIB over the PIC shows that our XPB domain is placed further away than in the conformation observed in the apo-TFIIB. For the XPB to make the conformational change required to engage the DNA in its current position, the MAT1 helical region must let go of the DRD domain.
Supplementary Figure 3 | Local resolution analysis of the distal CC and IC DNA binding cleft

Local resolution analysis of the DNA binding cleft of Pol II show a comparable distribution of local resolution at high and low thresholds, indicating that the differences in density intensity around the DNA are not due to resolution loss.
Supplementary Figure 4 | Multibody analysis and further 3D classification of cPIC-TFIIH

a. Multibody analysis shows that motion present between the cPIC and TFIH are Gaussian distributions and no notable alternative conformational minima are present.
b. Further 3D classification of IC (top) and OC (bottom) classes show a roughly equal distribution of classes with varying positions of TFIH with respect to the cPIC.