doi:10.1038/nature09785

Supplementary Tables

Table S1 Data collection and refinement statistics

	Arrested Pol II	Reactivation intermediate containing TFIIS
Data collection[1,2]		
Space group	C 2 2 2_1	C 2 2 2_1
Cell dimensions		
a, *b*, *c* (Å)	224.6 394.6 282.1	220.0, 395.1, 280.4
Resolution (Å)	50 – 3.3 (3.48 – 3.30)*	50 – 3.3 (3.48 – 3.30)
R_{sym} or R_{merge}	0.123 (0.910)	0.105 (0.785)
$I/\sigma I$	9.3 (2.2)	9.1 (2.0)
Completeness (%)	99.9 (99.9)	99.8 (99.9)
Redundancy	6.8 (6.4)	5.7 (5.5)
Refinement[3-5]		
Resolution (Å)	50 – 3.3	50 – 3.3
No. unique reflections	186691	181822
R_{work}/ R_{free}[5]	0.153/0.178[+]	0.163/0.190[+]
No. atoms	32287	32998
Protein	31275	32445
Ligand/ion	1012	553
B-factors (Å^2)		
Protein	121.7	126.5
Ligand/ion	207.5	213.8
R.m.s. deviations		
Bond lengths (Å)	0.010	0.010
Bond angles (°)	1.34	1.34

* Highest resolution shell is shown in parenthesis.
[+] Refinement with CNS 1.21[6] produced R_{work}/R_{free} values of 0.209/0.235 and 0.218/0.244 for arrested and reactivation complexes, respectively.

Table S2 Contacts of backtracked RNA nucleotides (nt) with Pol II amino acid residues forming the backtrack site, and their conservation in eukaryotes.

RNA nt	Pol II Subunit	Sc. residue	Interaction type	Conserved*
+2 ribose (O2)	Rpb1	T827 (OG1)	H-bond (3.6Å)	Sp, At, Ce, Hs, Mm, Dm
+2 base (N4)	Rpb2	E529 (OE1)	H-bond (3.5Å)	Sp, At, Ce, Hs, Mm, Dm
+2 base	Rpb2	Y769	Stacking	Sp, At, Ce, Hs, Mm, Dm
+3 base (O2)	Rpb1	Q1078 (OE1)	H-bond (3.0Å)	Sp, At, Ce, Hs, Mm, Dm
+3 base (N3)	Rpb1	Q1078 (O)	H-bond (3.3Å)	Sp, At, Ce, Hs, Mm, Dm
+3 base (N4)	Rpb1	Q1078 (O)	H-bond (3.1Å)	Sp, At, Ce, Hs, Mm, Dm
+3 base (N4)	Rpb1	T1080 (OG1)	H-bond (2.7Å)	Sp, At, Ce, Hs, Mm, Dm
+3 backbone (O1P)	Rpb2	R766 (NH2)	Salt bridge (5.1Å)	Sp, At, Ce, Hs, Mm, Dm
+4 backbone (O1P)	Rpb1	K752 (NZ)	H-bond (3.0Å)	Sp, At, Ce, Hs, Mm, Dm
+5 backbone (O2P)	Rpb1	S754 (OG)	H-bond (3.5Å)	Sp, At, Ce, Hs, Mm, Dm
+5/+6 base	Rpb1	I756	Intercalation	Sp, At, Ce, Hs, Mm, Dm
+6 backbone (O2P)	Rpb1	S754 (OG)	H-bond (3.0Å)	Sp, At, Ce, Hs, Mm, Dm
+6 base	Rpb1	R726	Stacking	Sp, At, Ce, Hs, Mm, Dm
+7 ribose (O2)	Rpb1	R731 (NH1)	H-bond (2.7Å)	Sp
+7 base	Rpb1	F755	Stacking	Sp, At

*Organism abbreviations are Sc: *Saccharomyces cerevisiae*, Sp: *Schizosaccharomyces pombe*, At: *Arabadopsis thaliana*, Ce: *Caenorhabditis elegans*, Hs: *Homo sapiens*, Mm: *Mus musculus*, Dm: *Drosophila melanogaster*.

Supplementary Figures

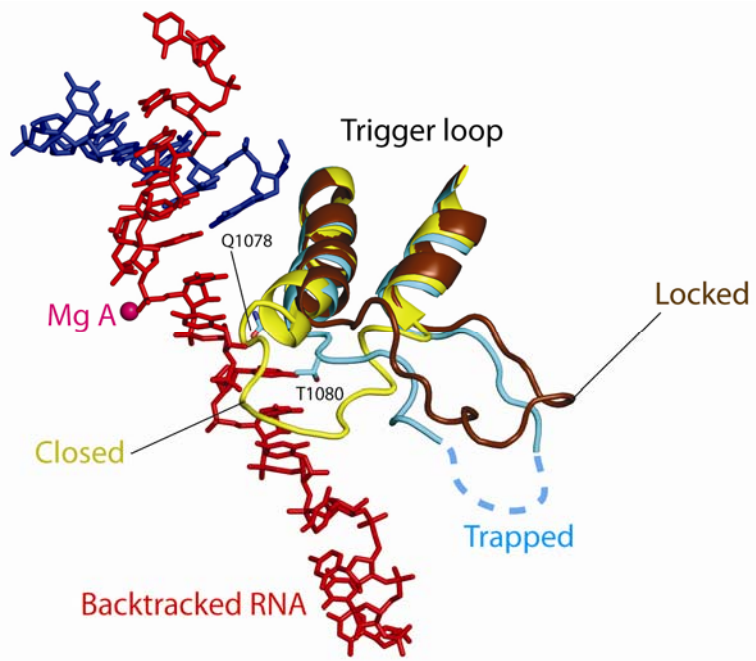
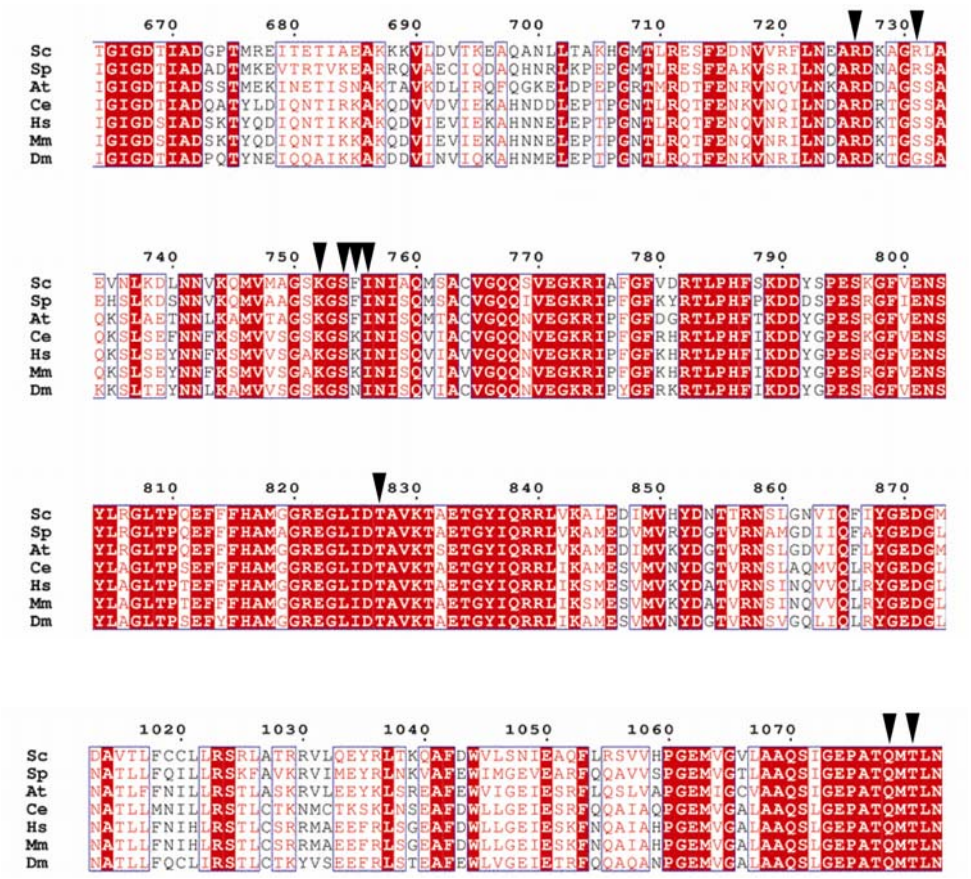


Figure S1 : Comparison of closed (pdb code 2E2H[7]), trapped (this study) and locked (pdb code 1Y1V[8] and this study) trigger loop conformations. Nucleic acids and the active site Mg A ion from the arrested Pol II complex are also shown. Residues Q1078 and T1080 are shown as stick models on the trapped trigger loop.

(a) Rpb1



(b) Rpb2

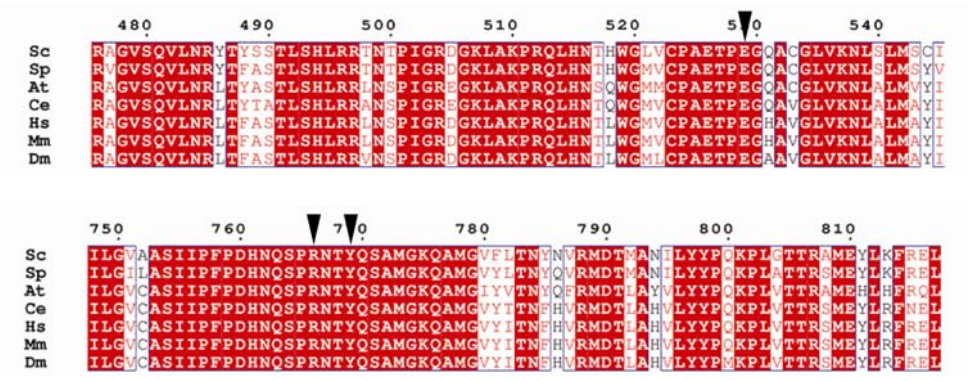


Figure S2 : Sequence conservation of Pol II residues forming the backtrack site within (a) Rpb1 and (b) Rpb2. Residues that contact the backtracked RNA are marked with an arrowhead. Organism abbreviations are: Sc, *Saccharomyces cerevisiae;* Sp, *Schizosaccharomyces pombe*; At, *Arabadopsis thaliana*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Dm, *Drosophila melanogaster.*

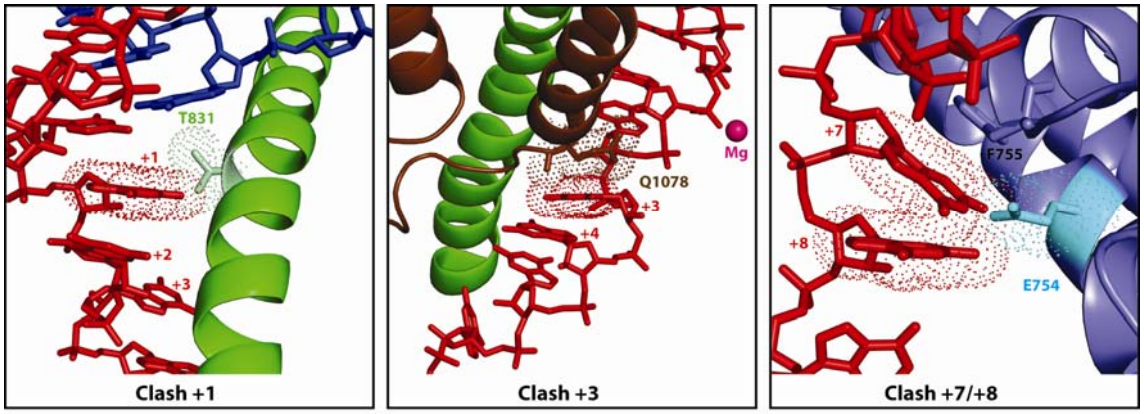


Figure S3 : Selected clashes of modelled guanine residues in backtracked RNA with Pol II residues in the backtrack site. The cytosine bases of the backtracked RNA in the arrested Pol II complex structure were mutated to guanine at all positions in COOT[9], maintaining the same dihedral angle about the glycosidic bond. The resulting structure was inspected visually for clashes and also submitted to the MOLPROBITY server[10] for hydrogen atom generation and clash analysis. Significant clashes for guanine bases were indicated at all backtracked positions except +5 and +9. The greatest clashes were at the backtracked positions +3, +7 and +8 and position +1 and are shown with van der Waals radii of clashing groups depicted as dots.

Supplementary References

1 The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763 (1994).

2 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

3 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).

4 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

5 autoBUSTER (Global Phasing Limited).

6 Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921 (1998).

7 Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941-954 (2006).

8 Kettenberger, H., Armache, K.-J. & Cramer, P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**, 347-357 (2003).

9 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

10 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).