

Dissertation zur Erlangung des Doktorgrades der Fakultät für Chemie
und Pharmazie der Ludwig-Maximilians-Universität München

**Molecular basis of translocation,
 α -amanitin inhibition,
and CPD damage recognition
by RNA polymerase II**



Florian Brückner
aus München

2008

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Patrick Cramer betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig und ohne unerlaubte Hilfe erarbeitet.

München, am 22. April 2008



Florian Brückner

Dissertation eingereicht am 22. April 2008

1. Gutachter: Prof. Dr. Patrick Cramer
2. Gutachter: Prof. Dr. Karl-Peter Hopfner

Mündliche Prüfung am 21. Mai 2008

Acknowledgements

This work was accomplished in the laboratory and under supervision of Patrick Cramer at the Gene Center in Munich. It was a very exciting and enjoyable experience for me to be involved in highly interesting basic research at this excellent place. I am deeply grateful to Patrick for this opportunity and for his continuous personal support. His excitement about science, cheerful character and his encouraging and respectful attitude towards his coworkers are infectious and create an extremely pleasant and motivating atmosphere in the lab. All this was crucial for keeping my motivation high and turning my efforts into success.

I am thankful to all present and former members of the Cramer lab for their highly collaborative attitude and all their help, for inspiring scientific discussions and for the great times together.

I particularly want to thank Karim Armache and Hubert Kettenberger for their pioneering work on Pol II and Pol II nucleic acid complexes in this lab, which formed the basis of my projects. Special thanks to Hubert for introducing me to many experimental techniques.

Moreover, I want to thank Alan Cheung, Gerke Damsma, Elisabeth Lehmann and Jasmin Sydow for joining to work on Pol II-nucleic acid complexes and letting me take part in their projects.

Special thanks to Stefan Benkert for his technical support in producing huge amounts of yeast, Claudia Buchen for helping with many of the everyday problems in the lab and Dirk Kostrewa for helping me on various crystallographic questions.

In addition, I am very grateful to all collaborators outside of the lab for sharing ideas and research projects, in particular Ulrich Hennecke, Aaron Alt and Thomas Carell from the organic chemistry department, Joanna Andrecka, Adam Muschielok and Jens Michaelis from the physical chemistry department, Karl-Peter Hopfner, Sheng Cui and Axel Kirchhofer from the Gene Center, Dmitry Temiakov from Stratford, New Jersey in the US and Julio Ortiz from the Max Planck Institute for Biochemistry. Many thanks to Clemens Schulze-Briese and his team at the beamline X06SA at the SLS in Villigen, Switzerland for their support for collecting x-ray diffraction data.

Thanks to Karl-Peter Hopfner, Thomas Carell, Dietmar Martin, Roland Beckmann and Elena Conti for being my PhD examiners.

Many thanks to my wife Viktoriya for all her support during the last three years of my PhD and for always cheering me up, when I was stressed or in a bad mood.

I am most deeply grateful to my parents, who made this education possible for me.

Summary

RNA polymerase II (Pol II) is the eukaryotic enzyme responsible for transcribing all protein-coding genes into messenger RNA (mRNA). This thesis describes studies on the molecular mechanisms of Pol II translocation, α -amanitin inhibition and DNA lesion recognition by Pol II. To study how Pol II translocates after nucleotide incorporation, we prepared elongation complex (EC) crystals in which pre- and post-translocation states interconvert. Crystal soaking with the inhibitor α -amanitin locked the EC in a new state that we identified as a translocation intermediate at 3.4 Å resolution. The DNA base entering the active site occupies a “pre-templating” position above the central bridge helix, which is shifted and occludes the standard templating position. A leucine residue in the trigger loop forms a wedge next to the shifted bridge helix, but moves by 13 Å to close the active site for nucleotide incorporation. Our results support a Brownian ratchet mechanism of elongation that involves swinging of the trigger loop between open, wedged, and closed positions, and suggest that α -amanitin impairs nucleotide incorporation and translocation by trapping the trigger loop and bridge helix in a translocation intermediate. Cells use transcription-coupled repair (TCR) to efficiently eliminate DNA lesions such as UV-induced cyclobutane pyrimidine dimers (CPDs). Here we present the structure-based mechanism for the first step in eukaryotic TCR, CPD-induced stalling of Pol II. A CPD in the transcribed strand slowly passes a translocation barrier, and enters the polymerase active site. The CPD 5'-thymine then directs uridine monophosphate (UMP) misincorporation into mRNA, which blocks translocation. Artificial replacement of the UMP by adenosine monophosphate (AMP) enables CPD bypass, thus Pol II stalling requires CPD-directed misincorporation. In the stalled complex, the lesion is inaccessible, and the polymerase conformation is unchanged. This is consistent with non-allosteric recruitment of repair factors and excision of a lesion-containing DNA fragment in the presence of Pol II. CPD recognition is compared with the recognition of a cisplatin-induced guanine-guanine intrastrand crosslink. Similarities and differences in the detailed mechanism of transcriptional stalling at the two different dinucleotide lesions are discussed.

Publications

Parts of this work have been published or are in the process of publication:

Kettenberger, H., Eisenfuhr, A., Brueckner, F., Theis, M., Famulok, M., and Cramer, P. (2006). Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by noncoding RNAs. *Nat Struct Mol Biol* **13**, 44-48.

Kashkina, E., Anikin, M., Brueckner, F., Pomerantz, R.T., McAllister, W.T., Cramer, P., Temiakov, D. (2006). Template misalignment in multisubunit RNA polymerases and transcription fidelity. *Mol Cell* **24**, 257-266.

Brueckner, F., Hennecke, U., Carell, T., Cramer, P. (2007). CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859-862.

Brueckner, F., Cramer, P. (2007). DNA photodamage recognition by RNA polymerase II. *FEBS Lett* **581**, 2757-2760.

Kashkina, E., Anikin, M., Brueckner, F., Lehmann, E., Kochetkov, S.N., McAllister, W.T., Cramer, P., Temiakov, D. (2007). Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site. *J Biol Chem* **282**, 21578-21582.

Lehmann, E., Brueckner, F., Cramer, P. (2007). Molecular basis of RNA-dependent RNA polymerase II activity. *Nature*, **450**, 445-449.

Damsma, G., Alt, A., Brueckner, F., Carell, T., Cramer, P. (2007). Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* **14**, 1127-33.

Andrecka, J., Lewis, R., Brueckner, F., Lehmann, E., Cramer, P., Michaelis, J. (2008). Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc Natl Acad Sci U S A*. **105**, 135-140.

Cramer, P., Armache, K.-J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C.-D., Lehmann, E., Leike, K., Sydow, J. and Vannini, A. (2008). Structure of Eukaryotic RNA Polymerases. *Annu. Rev. Biophys.* **37**, 337-352.

Brueckner, F., Cramer, P. (2008). Structural basis of transcription inhibition by α -amanitin and implications for RNA polymerase II translocation. *Revised*.

Table of contents

Erklärung	II
Ehrenwörtliche Versicherung.....	II
Acknowledgements	III
Summary.....	IV
Publications.....	V
1 Introduction	1
1.1 The flow of genetic information.....	1
1.2 DNA-dependent RNA polymerases	1
1.3 The mRNA transcription cycle	2
1.4 Composition and structure of RNA polymerases	4
1.5 Structure of RNA polymerase II	6
1.6 The RNA polymerase II elongation complex and nucleotide incorporation.....	7
1.7 Obstacles during elongation	10
1.8 Open questions and scope of this work	11
1.9 Additional contributions	11
2 Molecular basis of RNA polymerase II translocation and α-amanitin inhibition.....	13
2.1 Introduction.....	13
2.2 Results	14
2.2.1 Co-existence of pre- and post-translocation states in Pol II EC crystals	14
2.2.2 Structure of the α -amanitin-inhibited complete Pol II EC	16
2.2.3 The trigger loop forms a wedge next to the central bridge helix.....	19
2.2.4 The new state of the EC is a translocation intermediate.....	21
2.3 Discussion.....	23
2.3.1 Structure of a translocation intermediate.....	23
2.3.2 Extended NAC model	23
2.3.3 Stepwise translocation and helicase activity of Pol II.....	25
2.3.4 Comparison with bacterial RNA polymerase	26
2.3.5 A general translocation mechanism	26
2.3.6 The Brownian ratchet and transcription fidelity	27
2.3.7 Single-subunit RNA polymerase and a possible power stroke.....	27
2.3.8 Mechanism of α -amanitin inhibition.....	28
2.4 Specific experimental procedures.....	29

2.4.1	Preparation of Pol II elongation complexes	29
2.4.2	Crystallization and crystal treatment.....	29
2.4.3	X-ray structure analysis.....	30
2.4.4	Transcript extension assays.....	30
3	CPD damage recognition by transcribing RNA polymerase II	32
3.1	Introduction.....	32
3.2	Results and discussion.....	32
3.2.1	The molecular mechanism of CPD damage recognition by RNA polymerase II	32
3.2.2	A topological model for transcription-coupled DNA repair	43
3.2.3	Comparison of CPD recognition by different polymerases	45
3.2.4	Comparison with the mechanism of cisplatin damage recognition.....	45
3.2.5	Additional detailed information	48
3.3	Specific experimental procedures.....	50
3.3.1	Preparation of Pol II-nucleic acid complexes.....	50
3.3.2	Crystallization and crystal treatment.....	51
3.3.3	X-ray structural analysis.....	51
3.3.4	Transcript extension assays.....	52
4	Common experimental procedures	53
4.1	Measurement of protein concentration.....	53
4.2	Isolation of 10-subunit core Pol II from yeast	53
4.2.1	Yeast fermentation.....	53
4.2.2	Purification of 10-subunit core RNA Pol II	54
4.3	Purification of His-tagged RNA polymerase II.....	59
4.4	Purification of the subcomplex Rpb4/7.....	59
4.5	Purification of elongation factor TFIIIS.....	60
4.6	Buffers for Pol II and Pol II ECs	61
4.7	Crystallization solutions for Pol II and Pol II ECs	62
5	Appendix.....	63
5.1	Unpublished RNA polymerase II elongation complex structures	63
5.1.1	Nucleic acid scaffolds	63
5.1.2	Crystallographic results.....	64
5.2	Attempts to improve the resolution of diffraction data from complete 12-subunit RNA polymerase II crystals.....	78
5.2.1	Capillary mounting and measuring at room temperature	78
5.2.2	Crystallization of 12-subunit Pol II without CTD.....	79
5.2.3	Further attempts in table form	79

5.2.4 Ideas for the future	80
6 Conclusions and Outlook	82
7 Abbreviations	83
8 References.....	85
9 Curriculum vitae – Florian Brückner	97
9.1 Personal Details	97
9.2 Curriculum vitae	98
9.3 Presentations at conferences	99

1 Introduction

1.1 The flow of genetic information

The genetic information in a cell is stored in the form of deoxyribonucleic acid (DNA). In order to make use of the information, DNA is transcribed into ribonucleic acid (RNA) by DNA-dependent RNA polymerases. RNA molecules, in turn, are translated into protein by the ribosome. The directional flow of genetic information from DNA via RNA to protein was put forward as "The Central Dogma of molecular biology" by Francis Crick in 1958 (Crick, 1970; Thierry & Sarkar, 1998) (Figure 1).



Figure 1: The central dogma of molecular biology

1.2 DNA-dependent RNA polymerases

The existence of cellular DNA-dependent RNA polymerases (RNAPs) was first demonstrated independently by Jerard Hurwitz (Hurwitz et al, 1960) and Samuel Weiss (Weiss & Nakamoto, 1961). Whereas in prokaryotes there is only one RNAP, in eukaryotes different kinds of RNA molecules are synthesized by at least three distinct RNAPs, discovered by the end of the 1960s (Roeder & Rutter, 1969). In plants a fourth type of RNAP was recently discovered (Herr et al, 2005; Kanno et al, 2005). Eukaryotic RNA polymerase (Pol) I is located in the nucleoli and produces 5.8S, 18S and 28S ribosomal RNAs (rRNA). Pol II is located in the nucleoplasm and produces messenger RNAs (mRNA) from all protein-coding genes, small nucleolar RNAs (snoRNAs) and some small nuclear RNAs (snRNAs). Pol III is also located in the nucleoplasm and produces transfer RNAs (tRNA), 5S rRNA, some snRNAs and other small RNAs (Alberts, 2002; Berg, 2001). Pol IV is thought to play a role in the silencing of repetitive DNA sequences in plants by RNA-directed DNA methylation (Till & Ladurner, 2007).

The eukaryotic RNA polymerases also differ in their sensitivity to the inhibitor α -amanitin, a deadly mushroom toxin from *Amanita phalloides* (death cap) (Schwartz et al, 1974). Pol I is insensitive to the toxin, Pol III is only inhibited by higher concentration of α -amanitin ($K_i = 10\text{-}40\ \mu\text{M}$), and Pol II is most sensitive ($K_i = 3\ \text{nM}$) (Wieland & Faulstich, 1991).

1.3 The mRNA transcription cycle

The regulation of gene transcription is one of the most important steps for the control of cell growth, cell differentiation and the proper response to environmental signals. Defects in transcription regulation can be the cause for severe diseases like cancer.

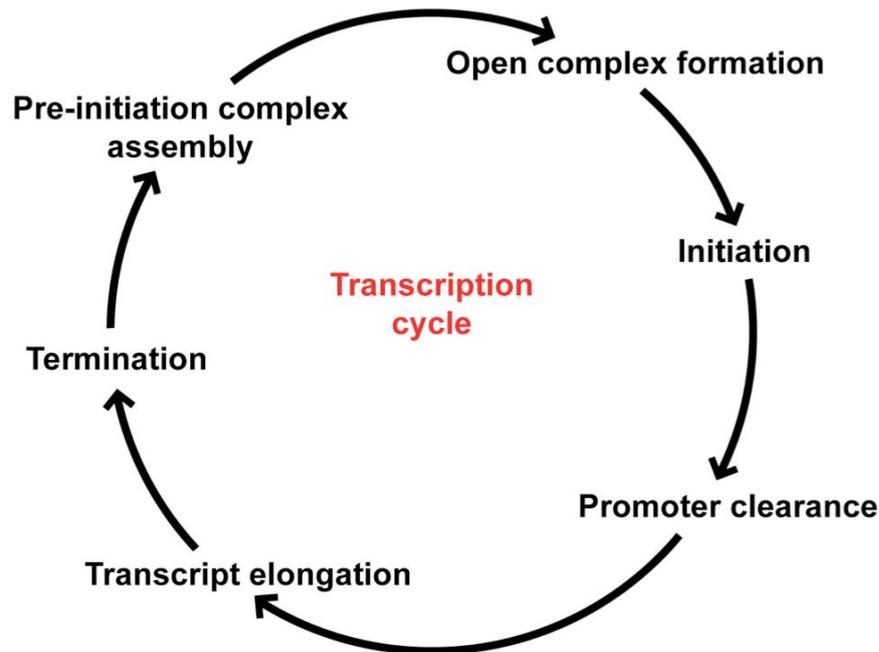


Figure 2: The transcription cycle. Adapted from (Svejstrup, 2004).

The transcription of a gene is characterized by a cycle of events, named "the transcription cycle" (Figure 2). The mRNA transcription cycle is most complex, and it involves a plethora of factors in addition to Pol II, many of which are important for gene regulation (Figure 3, Table 1) (Hahn, 2004; Svejstrup, 2004). The mRNA transcription cycle can be roughly divided into three stages: initiation, elongation and termination. Most of the regulation takes place during the initiation phase. The chromatin at the promoter region must be modified appropriately to allow initiation (Li et al, 2007). Pol II then assembles at the core promoter together with several general transcription factors (GTFs) and coactivators (see Table 1) to form the pre-initiation complex (PIC) (Bushnell et al, 2004; Chen et al, 2007; Hahn, 2004; Miller & Hahn, 2006). Additional regulatory factors, namely activators and repressors bind to enhancer and silencer elements on the DNA respectively, and transmit regulatory signals via the coactivators. In the PIC the double-stranded promoter DNA must be melted to make the template DNA accessible for the polymerase. This results in the formation of a transcription bubble, i.e. the open complex. Pol II can then initiate template DNA-dependent mRNA synthesis. Pol II leaves most of the initiation factors behind at the promoter (promoter clearance) to enter into elongation phase. These factors (scaffold complex) greatly support re-initiation by another Pol II molecule (Hahn, 2004). During the elongation phase the mRNA is produced in

a processive manner. The pre-mRNA is synthesized by Pol II and co-transcriptionally processed by capping and splicing. Besides the factors required for these processes, elongation factors are involved, which modulate the catalytic activity of Pol II, pausing of Pol II, and transcriptional arrest (TFIIS) (Sims et al, 2004). Additional factors are required to allow transcription through chromatin (Armstrong, 2007; Kulaeva et al, 2007; Li et al, 2007). Termination occurs at the end of a gene (Gilmour & Fan, 2008). The RNA is cleaved and polyadenylated upon a signal on the transcript. Pol II is then removed from the DNA and recycled to enter another round of transcription. The C-terminal domain of the largest subunit of Pol II (CTD) couples mRNA transcription to the subsequent events of mRNA biogenesis, namely co-transcriptional mRNA processing, chromatin remodeling and modification, DNA repair, mRNA packaging, RNA editing and nuclear mRNA export (for reviews see (Buratowski, 2003; Hirose & Manley, 2000; Meinhart et al, 2005)). It does so by recruiting factors, which recognize the phosphorylation pattern of the CTD, which changes during the transcription cycle in a specific manner.

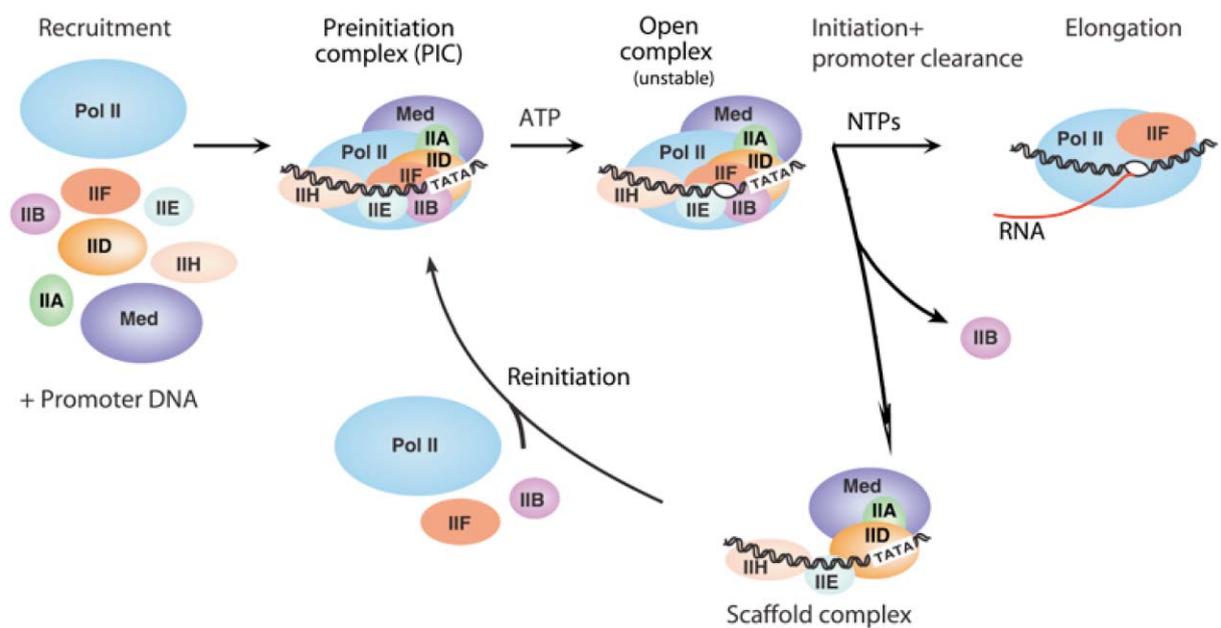


Figure 3: The pathway of transcription initiation and reinitiation for Pol II.

See Table 1 for a description of each transcription factor and Mediator (Med). Adapted from (Hahn, 2004).

Table 1: *Saccharomyces cerevisiae* Pol II general transcription factors and coactivators

Factor		number of subunits	Function (Hahn, 2004)
TFIIA		2	Stabilizes TBP and TFIID-DNA binding. Blocks transcription inhibitors. Positive and negative gene regulation.
TFIIB		1	Binds TBP, Pol II and promoter DNA. Helps fix transcription start site.
TFIID	TBP	1	Binds TATA element and deforms promoter DNA. Platform for assembly of TFIIB, TFIIA and TAFs.
	TAFs	14	Binds INR and DPE promoter elements. Target of regulatory factors.
Mediator		24	Binds cooperatively with Pol II. Kinase and acetyltransferase activity. Stimulates basal and activated transcription. Target of regulatory factors.
TFIIF		3	Binds Pol II and is involved in Pol II recruitment to PIC and in open complex formation.
TFIIE		2	Binds promoter near transcription start. May help open or stabilize the transcription bubble in the open complex.
TFIIH		10	Functions in transcription and DNA repair. Kinase and two helicase activities. Essential for open complex formation. Mutations in IIH can cause human disease.

1.4 Composition and structure of RNA polymerases

DNA-dependent RNA polymerases can be classified into two families, comprising single subunit and multisubunit RNA polymerases, respectively (Cramer, 2002a).

The single subunit family encompasses RNA polymerases of bacteriophages (e.g. T7, T3, SP6) and mitochondrial RNA polymerase. The respective enzymes all consist of a single protein subunit, characterized by a right hand-like architecture with palm, thumb and finger domains (canonical domains), similar to DNA polymerases. The best studied single subunit RNA polymerase is from the bacteriophage T7. It consists of 883 residues (99 kDa) and the crystal structure of the free enzyme was reported already in 1993 (Sousa et al, 1993).

The multisubunit RNA polymerase family includes the RNA polymerases from all three domains of life, bacteria, archaea and eukaryotes (Cramer, 2002a). Whereas bacteria and archaea possess only one type of RNA polymerase, all eukaryotes have the nuclear Pol I, II, and III. Plants possess additional multisubunit RNAPs, Pol IV and the RNAP of chloroplasts. There is no sequence or structural homology between single and multisubunit RNAPs detectable. Therefore the existence of common topological features and mechanistic principles could be the result of convergent evolution (Cramer, 2002a). Multisubunit RNAPs contain between five and 17 protein subunits and have total molecular weights between around 400 to around 700 kDa (Table 2). Five core subunits of multisubunit RNAPs are conserved between all three kingdoms of life and result in a similar general architecture of the enzymes (Cramer, 2002b). The overall structure is reminiscent of a crab claw with a large central cleft. The active center is located at the floor of the cleft, and it's high structural conservation suggests a general catalytic mechanism for all multisubunit RNAPs. The additional subunits lie

on the outside of the five-subunit core enzyme and thereby change the interaction surface of the RNAPs. Thus the additional subunits probably fulfill the domain-specific needs for transcription regulation by factors interacting with the RNAPs.

Crystal structures have been determined of bacterial RNA polymerases (*Thermus aquaticus*, *Thermus thermophilus*) (Vassilyev et al, 2002; Zhang et al, 1999), archaeal RNA polymerase (*Sulfolobus solfataricus*) (Hirata et al, 2008) and eukaryotic Pol II (*Saccharomyces cerevisiae*) in its 10-subunit form (Cramer et al, 2000; Cramer et al, 2001) and 12-subunit form (Armache et al, 2003; Armache et al, 2005; Bushnell & Kornberg, 2003). For the Pol I and Pol III enzymes crystal structures are still lacking, however electron microscopic structures have been published at 12 and 17 Å, respectively (Fernandez-Tornero et al, 2007; Kuhn et al, 2007).

Table 2: Subunit composition of multisubunit RNA polymerases

RNA polymerase	Pol I	Pol II	Pol III	Archaea	Bacteria
Ten-subunit core	A190	Rpb1	C160	A' + A''	β'
	A135	Rpb2	C128	B (B' + B'')	β
	AC40	Rpb3	AC40	D	α
	AC19	Rpb11	AC19	L	α
	Rpb6 (ABC23)	Rpb6	Rpb6	K	ω
	Rpb5 (ABC27)	Rpb5	Rpb5	H	-
	Rpb8 (ABC14.5)	Rpb8	Rpb8	-	-
	Rpb10 (ABC10α)	Rpb10	Rpb10	N	-
	Rpb12 (ABC10β)	Rpb12	Rpb12	P	-
	A12.2	Rpb9	C11	-	-
Stalk	A14	Rpb4	C17	F	-
	A43	Rpb7	C25	E'	-
TFIIF-like subcomplex ^a	A49	(Tfg1/Rap74)	C37	-	-
	A34.5	(Tfg2/Rap30)	C53	-	-
Pol III-specific subcomplex	-	-	C82	-	-
	-	-	C34	-	-
	-	-	C31	-	-
Number of subunits	14	12	17	11 (12)	5
Molecular weight in kDa (<i>species</i>)	589 (<i>S. cerevisiae</i>)	514 (<i>S. cerevisiae</i>)	693 (<i>S. cerevisiae</i>)	380 (<i>P. furiosus</i>)	375 (<i>T. aquaticus</i>)

^aThe two subunits in Pol I and Pol III are predicted to form heterodimers that resemble part of the Pol II initiation/elongation factor TFIIF, which is composed of subunits Tfg1, Tfg2, and Tfg3 in *Saccharomyces cerevisiae*, and of subunits Rap74 and Rap30 in human.

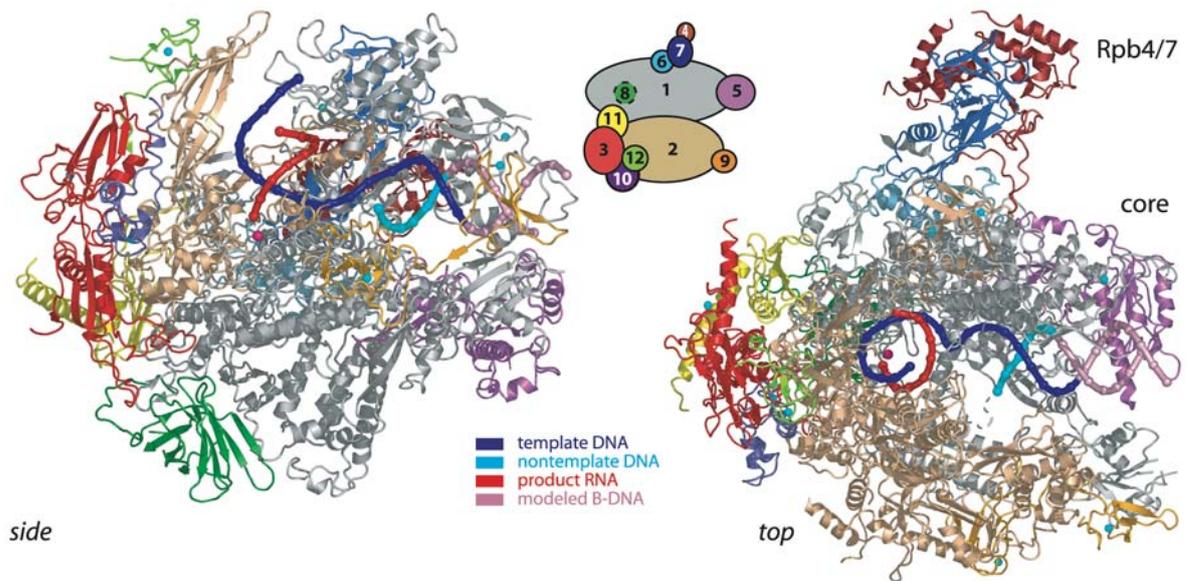


Figure 4: Structure of complete 12-subunit RNA polymerase II elongation complex (Kettenberger et al, 2004).

Two views of a ribbon model of the protein subunits and nucleic acids, side view (left) and top view (right). The polymerase subunits Rpb1–Rpb12 are colored according to the key between the views. Template DNA, nontemplate DNA, and product RNA are shown in blue, cyan, and red, respectively. Phosphorous atoms are indicated as spheres and extrapolated B-form downstream DNA is colored in light pink. Eight zinc ions and the active site magnesium ion are depicted as cyan spheres and a magenta sphere, respectively. This color code is used throughout. Secondary structure assignments for pol II are according to (Cramer et al, 2001) and (Armache et al, 2005). This figure is adapted from (Kettenberger et al, 2004).

1.5 Structure of RNA polymerase II

Pol II consists of a 10-subunit core enzyme and a peripheral heterodimer of subunits Rpb4 and Rpb7 (stalk, Table 2, Figure 4). The core enzyme comprises subunits Rpb1, Rpb2, Rpb3, and Rpb11, which contain regions of sequence and structural similarity in Pol I, Pol III, bacterial RNA polymerases (Vassilyev et al, 2002; Zhang et al, 1999), and the archaeal RNA polymerase (Kusser et al, 2008; Murakami, 2007). The Pol II core also comprises subunits Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12, which are shared between Pol I, II, and III (common subunits, Table 2). Counterparts of these common subunits except Rpb8 exist in the archaeal polymerase, but only a counterpart of Rpb6 exists in the bacterial enzyme (Minakhin et al, 2001). Finally, homologues of the core subunit Rpb9 exist in Pol I and Pol III, but not in the archaeal or bacterial enzyme. Initial electron microscopic studies of Pol II revealed the overall shape of the enzyme (Darst et al, 1991). The core Pol II could subsequently be crystallized, leading to an electron density map at 6 Å resolution (Fu et al, 1999). Crystal improvement by controlled shrinkage and phasing at 3 Å resolution resulted in a backbone model of the Pol II core (Cramer et al, 2000). This revealed that Rpb1 and Rpb2 form opposite sides of a positively charged active center cleft, whereas the smaller subunits

are arrayed around the periphery. Refined atomic structures of the core Pol II at 3.1 and 2.8 Å resolution were obtained in two different conformations and revealed domain-like regions within the subunits, as well as surface elements predicted to have functional roles (Cramer et al, 2001) (Figure 4 Figure 5). The active site and the bridge helix, which spans the cleft, line a pore in the floor of the cleft. The Rpb1 side of the cleft forms a mobile clamp, which was trapped in two different open states in the free core structures (Cramer et al, 2001) but was closed in the structure of a core complex that included DNA and RNA (Gnatt et al, 2001). The mobile clamp is connected to the body of the polymerase by five switch regions that show conformational variability. The Rpb2 side of the cleft consists of the lobe and protrusion domains. Rpb2 also forms a protein wall that blocks the end of the cleft. The Pol II core structures lacked subunits Rpb4 and Rpb7, which can dissociate from the yeast enzyme (Edwards et al, 1991). Crystallographic backbone models of the complete 12-subunit Pol II at ≈ 4 Å resolution revealed the exact position and orientation of Rpb4/7 and showed that it formed a wedge between the clamp and the linker to the unique tail-like C-terminal repeat domain (CTD) of the polymerase (Armache et al, 2003; Bushnell & Kornberg, 2003). In all crystal structures of the complete Pol II the clamp was observed in a closed conformation. The crystal structure of free Rpb4/7 together with an improved resolution of the complete Pol II crystals finally enabled refinement of a complete atomic model of Pol II (Armache et al, 2005). The CTD of Pol II (see chapter 1.3) is flexibly linked to the core enzyme and consists of heptapeptide repeats of the consensus sequence YSPTSPS. It is disordered in crystal structures of Pol II.

1.6 The RNA polymerase II elongation complex and nucleotide incorporation

During the elongation phase of transcription, the polymerases move along a DNA template and synthesize a complementary chain of ribonucleotides. The events required for the addition of each nucleotide to the product RNA form a cyclic process, referred to as the "Nucleotide Addition Cycle" (NAC). RNA extension begins with binding of a nucleoside triphosphate (NTP) substrate to the transcription elongation complex (EC) that is formed by the polymerase, DNA, and RNA. Catalytic addition of the nucleotide to the growing RNA 3' end then releases a pyrophosphate ion. Finally, translocation of DNA and RNA frees the substrate site for binding of the next NTP. The EC is characterized by an unwound DNA region, the transcription bubble. The bubble contains a short hybrid duplex formed between the DNA template strand and the RNA product emerging from the active site. The mechanism of RNA elongation was elucidated by structural studies of Pol II-nucleic acid complexes (elongation complexes, Figure 4, Figure 5). EM first revealed the point of DNA entry to the Pol II cleft (Poglitich et al, 1999). The first crystal structure of a Pol II-nucleic acid complex was that of the core Pol II transcribing a tailed template DNA (Gnatt et al, 2001), which allows for pro-

moter-independent transcription initiation. This structure revealed downstream DNA entering the cleft and a 8 to 9 base pair DNA-RNA hybrid in the active center. Comparison with the high-resolution core Pol II structure (Cramer et al, 2001) revealed protein surface elements predicted to play functional roles. Later, polymerase EC structures utilized different kinds of synthetic DNA-RNA scaffolds. A mismatch bubble scaffold contained upstream and downstream DNA duplex with RNA annealed to a central mismatched bubble region (Kettenberger et al, 2004) (Figure 4). The upstream DNA duplex and the nontemplate strand in the bubble region were however disordered in the crystal structure. Reduced scaffolds, lacking those disordered parts ("minimal nucleic acid scaffolds", first used in (Westover et al, 2004a)) or lacking the entire nontemplate strand (Westover et al, 2004b) were applied. The synthetic scaffold EC structures revealed the exact location of the downstream DNA and several nucleotides upstream of the hybrid. Mechanisms were suggested for how Pol II unwinds downstream DNA and how it separates the RNA product from the DNA template at the end of the hybrid. Additional structures of Pol II ECs included the NTP substrate (Kettenberger et al, 2004; Wang et al, 2006; Westover et al, 2004a). These studies suggested how the polymerase selects the correct NTP and how it incorporates a nucleotide into RNA. The NTP was crystallographically trapped in the insertion site (Wang et al, 2006; Westover et al, 2004a), which is apparently occupied during catalysis, but also in an overlapping, slightly different location, suggesting an inactive NTP-bound preinsertion state of the enzyme (Kettenberger et al, 2004). The NTPs in both states form Watson-Crick interactions with a base in the DNA template strand. Binding of the NTP to the insertion site involves folding of the trigger loop (Wang et al, 2006), a mobile part of the active center first observed in free bacterial RNA polymerase (Vassylyev et al, 2002), and in the Pol II-TFIIS complex (Kettenberger et al, 2003). Folding of the trigger loop closes the active site and may be involved in selection of the correct NTP. The NTP complex structures revealed contacts of the nucleotide with the polymerase, which explain discrimination of ribonucleotides against deoxyribonucleotides, and provided insights into the selection of the nucleotide complementary to the templating DNA base. Catalytic nucleotide incorporation apparently follows a two-metal ion mechanism suggested for all polymerases (Steitz, 1998). The Pol II active site contains a persistently bound metal ion (metal A) and a second, mobile metal ion (metal B) (Cramer et al, 2001). Metal A is held by three invariant aspartate side chains and binds the RNA 3' end (Cramer et al, 2001), whereas metal B binds the NTP triphosphate moiety (Westover et al, 2004a). Recent studies of functional complexes of the bacterial RNA polymerase revealed the close conservation of the EC structure (Vassylyev et al, 2007a) and provided additional insights into nucleotide incorporation (Vassylyev et al, 2007b). As for Pol II, NTP binding to the insertion site can induce folding of the trigger loop. However, in the presence of the antibiotic streptolydigin the NTP binds in the inactive, preinsertion state, in which the triphosphate and metal B are too far from metal A to permit catalysis. This finding supported a two-step mechanism of nucleotide incorporation (Kettenberger et al, 2004; Vassylyev et al, 2007b).

The NTP would first bind in the inactive state to an open active center conformation. Complete folding of the trigger loop then leads to closure of the active center, delivery of the NTP to the insertion site, and catalysis. An alternative model for nucleotide addition involves binding of the NTP to a putative entry site in the pore, in which the nucleotide base is oriented away from the DNA template, and rotation of the NTP around metal ion B directly into the insertion site (Westover et al, 2004a).

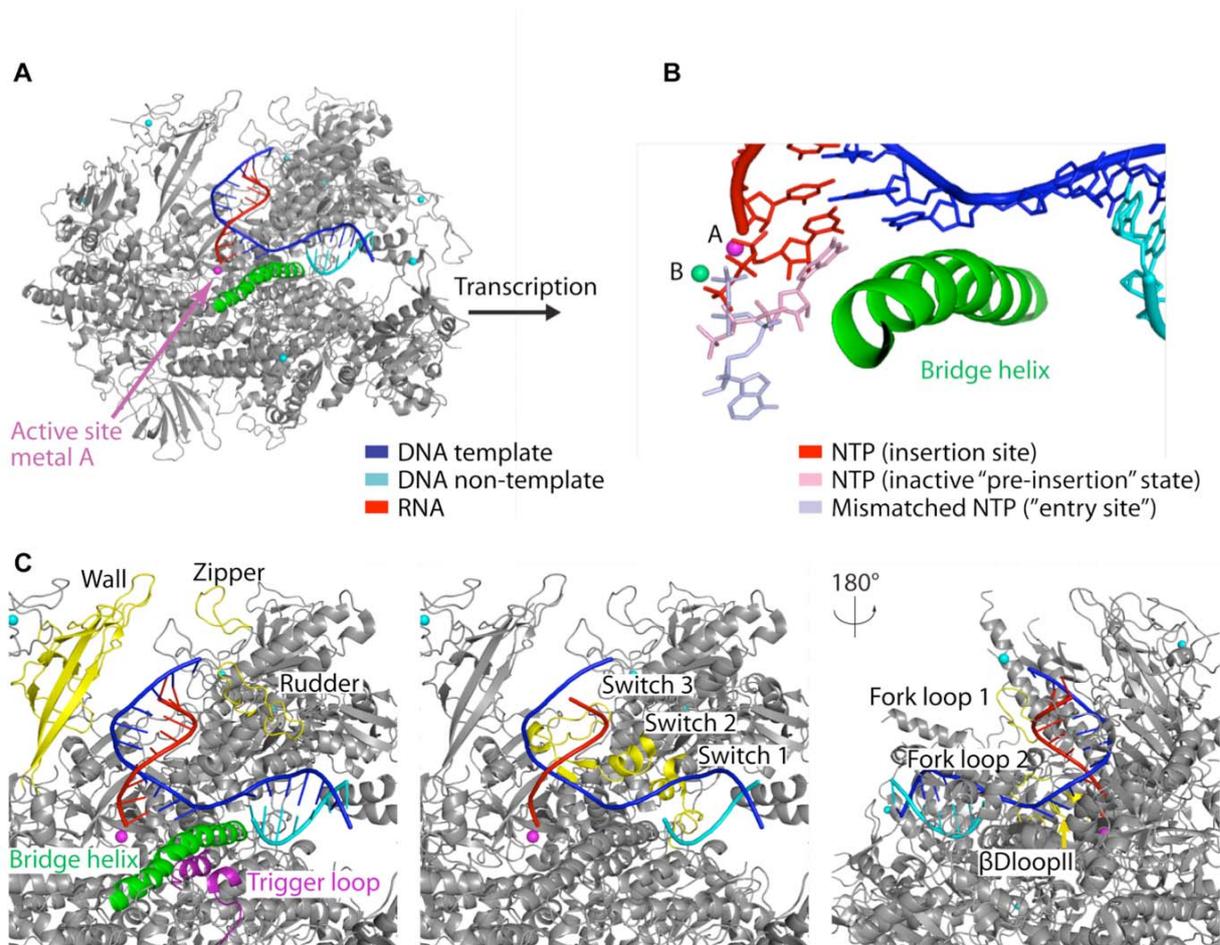


Figure 5: Structure of the Pol II elongation complex

(A) Overview of the EC structure (Kettenberger et al, 2004). The view is as in Figure 4, left.

(B) Superposition of NTP-binding sites [red, insertion site (Wang et al, 2006; Westover et al, 2004a); violet, entry site (Westover et al, 2004a); pink, inactive preinsertion-like state (Kettenberger et al, 2004)].

(C) Functional Pol II surface elements in the EC.

The EC of the single subunit T7 RNAP has been studied in great detail (Steitz, 2004; Steitz, 2006; Temiakov et al, 2004). Several important intermediates of the nucleotide addition cycle could be isolated and crystallographically analyzed at high resolution, resulting in a rather detailed mechanistic understanding of elongation in this system. Despite the structural differences between T7 RNAP and multisubunit RNAPs, there are remarkable similarities in some aspects of the elongation mechanism, including the course of nucleic acids through the active center, the two-metal ion catalytic mechanism, binding of the template-product helix from the minor groove side, passing of the transcript through an exit tunnel, and stacking of

an α -helix against the growing end of the template-product duplex (Cramer, 2002a; Steitz, 2006).

1.7 Obstacles during elongation

During active transcription, Pol II must overcome intrinsic DNA arrest sites, which are generally rich in A-T base pairs and pose a natural obstacle to transcription. At such sites, Pol II moves backward along DNA and RNA, resulting in extrusion of the RNA 3' end through the polymerase pore beneath the active site and transcriptional arrest. The RNA cleavage stimulatory factor TFIIS can rescue an arrested polymerase by creating a new RNA 3' end at the active site from which transcription can resume. The mechanism of TFIIS function was elucidated with the structures of Pol II and a Pol II EC in complex with TFIIS (Kettenberger et al, 2003; Kettenberger et al, 2004). TFIIS inserts a hairpin into the polymerase pore and complements the active site with acidic residues, changes the enzyme conformation, and repositions the RNA transcript (Kettenberger et al, 2003; Kettenberger et al, 2004). These studies supported the idea that the Pol II active site is tunable, as it can catalyze different reactions, including RNA synthesis and RNA cleavage (Kettenberger et al, 2003; Sosunov et al, 2003).

Other obstacles to transcription are bulky lesions in the DNA template strand, e.g. the UV-light induced cyclobutane pyrimidine dimer (CPD) or intrastrand cross-links induced by the anticancer drug cisplatin. Bulky DNA lesions can block transcription and replication, and lead to mutations during replication, causing cancer (Mitchell et al, 2003). Cells can eliminate bulky DNA lesions slowly by genome-wide nucleotide excision repair (NER). However, for rapid and efficient repair cells use a NER sub-pathway referred to as transcription-coupled DNA repair (TCR). TCR specifically removes lesions such as CPDs from the DNA strand transcribed by Pol II (Saxowsky & Doetsch, 2006). It is thought that only those lesions that can stably stall Pol II trigger TCR. CPDs are bulky lesions that lead to Pol II stalling, but other types of damages, such as oxidative damages, can be bypassed by Pol II, and would escape TCR (Charlet-Berguerand et al, 2006). Pol II stalling apparently triggers TCR by recruitment of a transcription-repair coupling factor (Rad26/CSB in yeast/human), and factors required for subsequent steps of nucleotide excision repair, including TFIIH, which unwinds DNA, and endonucleases, which incise the DNA strand on either side of the lesion (Mu & Sancar, 1997; Saxowsky & Doetsch, 2006; Selby et al, 1997; Svejstrup, 2002; Tremeau-Bravard et al, 2004). The obtained DNA gap is subsequently filled by DNA synthesis and ligation (Prakash & Prakash, 2000; Sancar, 1996).

1.8 Open questions and scope of this work

The structures of the Pol II elongation complexes together with numerous functional studies have given spectacular insights into the mechanism of mRNA transcription. Structural and functional studies of bacterial and T7 RNA polymerase elongation complexes allow for interesting comparisons. However, many aspects of the Pol II transcription mechanism require further investigation, some of which have been studied in this work.

Among the unresolved issues was the molecular mechanism of translocation of Pol along the template DNA after nucleotide incorporation. Therefore, the aim of this work was the isolation and structural characterization of Pol II EC states, informative on the mechanism of translocation. In particular, an intermediate of translocation was sought. In this context, also the mechanism of inhibition by the mushroom toxin α -amanitin, a proposed translocation inhibitor, was further examined. This part of the work is described in chapter 2.

In addition, the mechanism of recognition of bulky DNA lesions such as UV- and cisplatin-induced lesions by transcribing Pol II was unknown. To address this question, such lesions were introduced into the DNA template strand at several different positions around the polymerase active site and the resulting Pol II ECs were studied structurally and in RNA elongation assays. This part of the work is described in chapter 3. The analysis concerning the cisplatin lesion was mainly conducted by Gerke Damsma (Damsma et al, 2007).

All crystal structures of Pol II ECs are limited to moderate resolution ($\approx 3.5\text{-}4 \text{ \AA}$). Higher resolution would allow more detailed insight into many mechanistic aspects. Attempts to improve the resolution of Pol II EC diffraction data are summarized in chapter 5.2.

Unpublished structures of Pol II ECs with various nucleic acid scaffold, which have been obtained during the time of this thesis work, are listed in chapter 5.1. They may be a reference for future experimental design or published in the context of future research projects.

1.9 Additional contributions

This chapter lists additional contributions, which are not described in this thesis.

Several open questions on the architecture of the Pol II EC remained. Although an RNA 20-mer was used in the assembly of the complete Pol II EC, electron density was only observed for ten nucleotides from the 3'-end (Kettenberger et al, 2004). The part of the RNA further beyond the DNA-RNA hybrid was obviously mobile or disordered in this complex. Exploring the path of the RNA exiting Pol II was addressed in collaboration with the group of Jens Michaelis. Single molecule fluorescence resonance energy transfer (FRET) experiments were developed and applied in this lab. This work has been published (Andrecka et al, 2008).

Another open question on the EC architecture concerned the exact dimensions of the transcription bubble in a natural EC. Ultimate clarification of this issue by existing Pol II EC structures was precluded by the design of the nucleic acid scaffolds, containing mismatches or lacking the respective portion (Kettenberger et al, 2004; Westover et al, 2004a; Westover et al, 2004b). The question of the exact location of downstream bubble opening was addressed in collaboration with the group of Dmitry Temiakov, who applied fluorescence quenching experiments. This study is published (Kashkina et al, 2007).

Although Pol II generally uses DNA as a template, there was also evidence that Pol II can use RNA templates. Synthetic RNA scaffolds were used to investigate the RNA-dependent RNA synthesis activity *in vitro* and structurally. These experiments were mainly conducted by Elisabeth Lehmann, and the results are published (Lehmann et al, 2007).

The inhibition of Pol II transcription initiation by RNA molecules had been described in the literature. The complex of Pol II with an *in-vitro* selected inhibitory RNA molecule (aptamer) was examined using x-ray crystallography and binding assays mainly by Hubert Kettenberger in collaboration with the group of Michael Famulok. A natural RNA inhibitor from *Mus musculus* - the B2 RNA – was prepared and used in competition assays with the RNA aptamer. The attempts to obtain structural information on the inhibition mechanism by a natural RNA inhibitor were so far unsuccessful (Lehmann, 2006). Parts of the work on the RNA inhibitors have been published (Kettenberger et al, 2006).

The fidelity determinants of Pol II transcription are far less investigated than that of DNA replication. Template misalignment, an interesting mechanism affecting transcription fidelity, was examined in collaboration with the group of Dmitry Temiakov. This study has been published (Kashkina et al, 2006).

2 Molecular basis of RNA polymerase II translocation and α -amanitin inhibition

2.1 Introduction

The nucleotide addition cycle (NAC) of transcription begins with binding of a nucleoside triphosphate (NTP) substrate to the RNA polymerase (Pol) II elongation complex (EC). The EC then incorporates a nucleotide into the growing RNA, generating a pyrophosphate ion, and leading to the pre-translocation state. The polymerase subsequently translocates downstream along DNA and RNA, to generate the post-translocation state and a free substrate-binding site for the next NTP, completing the NAC. X-ray crystallography of Pol II ECs revealed how Pol II binds to DNA and RNA in the pre- (Gnatt et al, 2001) and post-translocation states (Kettenberger et al, 2004; Westover et al, 2004a), and how it binds NTP in various ways (Kettenberger et al, 2004; Wang et al, 2006; Westover et al, 2004a).

Despite this progress, the structural basis for translocation remains poorly understood. A comparison of the free crystal structures of core Pol II and a bacterial RNA polymerase initially suggested that translocation involves a conformational change in the highly conserved bridge helix within the active center (Cramer et al, 2001; Gnatt et al, 2001). The bridge helix was straight in Pol II (Cramer et al, 2001), but was bent and partially unwound at its center in the bacterial RNA polymerase (Zhang et al, 1999). The straight bridge helix stacked onto the end of the pre-translocated DNA-RNA hybrid, and its possible bending towards the hybrid was suggested to accompany nucleic acid movement during translocation (Cramer et al, 2001; Gnatt et al, 2001). In a subsequent structure of another bacterial RNA polymerase, the central region of the bridge helix was flipped-out and contacted the adjacent, previously mobile trigger loop, which was suggested to influence bridge helix movement (Vassylyev et al, 2002). Biochemical studies of bacterial RNA polymerases revealed the functional cooperation of the bridge helix and trigger loop, and suggested more detailed models of translocation (Bar-Nahum et al, 2005; Epshtein et al, 2002). Alternative bridge helix conformations were also observed within the same RNA polymerase, supporting the functional relevance of bridge helix movement (Tuske et al, 2005).

The mushroom toxin α -amanitin was suggested to interfere with bridge helix movement during translocation since it binds the free Pol II core below the bridge helix (Bushnell et al, 2002). Kinetic data supported the idea that α -amanitin inhibits translocation (Gong et al, 2004). However, a difficulty with this model was that α -amanitin bound only one residue in the bridge helix, and that this residue was separated by one turn from the residues that apparently change conformation. An alternative model for inhibition suggested that α -amanitin interferes with movement of the trigger loop, which closes over the active site during nucleotide incorporation (Wang et al, 2006).

To investigate structurally how Pol II translocation occurs, and how α -amanitin interferes with it, we formed crystals of the complete Pol II EC in which the pre- and post-translocation states co-exist. Soaking the inhibitor into such a crystal resulted in the structure of the α -amanitin-inhibited Pol II EC. This structure showed that α -amanitin traps the trigger loop in a new conformation, and stabilizes the EC in a previously unobserved state that corresponds to a translocation intermediate. The new structure leads to an extended model for the NAC and suggests how α -amanitin interferes with both nucleotide incorporation and Pol II translocation.

2.2 Results

2.2.1 Co-existence of pre- and post-translocation states in Pol II EC crystals

Previous structures of the Pol II EC were not informative on the mechanism of translocation since they always revealed the same polymerase conformation, although the nucleic acids adopted either the pre- or the post-translocation conformation. Since the EC is thought to exist in an equilibrium of pre- and post-translocation states in solution, we sought crystals of the complete Pol II EC that preserve this equilibrium. To detect the translocation state of nucleic acids within EC crystals, we labeled the DNA template strand in the hybrid with 5-bromouracil, and determined the bromine position by anomalous diffraction. With the use of a minimal synthetic nucleic acid scaffold (Figure 6A) and at low magnesium ion concentration, two bromine peaks were observed, one at position -4, indicating the post-translocation state, and one at position -3, corresponding to the pre-translocation state (Figure 6C, Table 3, chapter 2.4) (position +1 denotes the nucleotide incorporation site, and positive and negative numbers denote downstream and upstream positions, respectively). This indicates that Pol II binds the hybrid in a defined, one base pair register, consistent with movement of Pol II in discrete one base pair steps (Abbondanzieri et al, 2005). Furthermore, both translocation states co-existed within these crystals. Since nucleic acids were not fixed by crystal contacts, we assumed that the two states inter-convert in each EC of the crystal. The EC preparation used for crystallization was functional, as it supported RNA elongation, and was also sensitive to α -amanitin inhibition (Figure 6B).

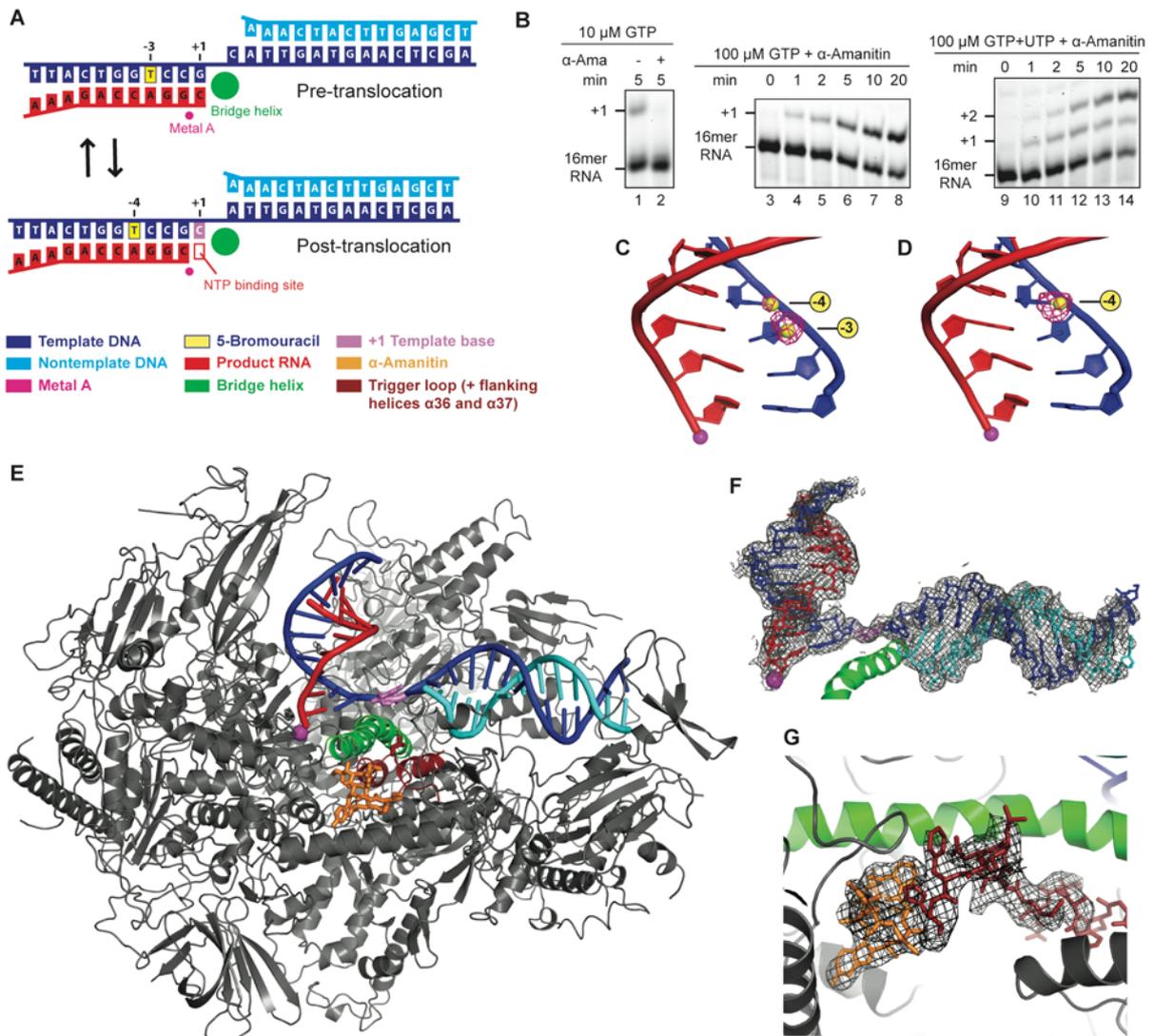


Figure 6: Structure of the α -amanitin-inhibited Pol II elongation complex (EC)

(A) Pre- and post-translocation states of the EC. The nucleic acid scaffold used in this study is depicted schematically with respect to the active site metal ion A (magenta). The color key is used throughout.

(B) Nucleotide incorporation and inhibition of the EC by α -amanitin. Inhibition of the EC was tested with a transcript extension assay (chapter 2.4.4). In the presence of α -amanitin and 10 μ M GTP, incorporation of the next complementary nucleotide was inhibited (lane 2). At higher substrate concentration (100 μ M), slow incorporation of one (lanes 3-8), or two nucleotides (lanes 9-14) is possible in the presence of α -amanitin.

(C-D) Bromine anomalous difference Fourier maps (pink net) of the free EC (C) and the α -amanitin-inhibited EC (D). The maps were calculated with phases from the Pol II model after molecular replacement and rigid body refinement and contoured at 4.4 σ . The final model of the nucleic acids and metal A in the α -amanitin-inhibited EC is superimposed. In the absence of α -amanitin, two peaks were observed, coinciding with the locations of the bromine atom (yellow sphere) in a 5-bromouracil residue at positions -3 and -4 in the template strand (Table 3). In the presence of α -amanitin, one peak was observed, coinciding with the location of the bromine atom (yellow sphere) in a 5-bromouracil residue at position -4 in the template strand (Table 3). The view is from the front (Cramer et al, 2001).

(E) Overview of the α -amanitin-inhibited Pol II EC structure. The view is from the side (as in Figure 4, left) (Kettenberger et al, 2004). α -amanitin (stick model), nucleic acids (pre-templating base as stick model), metal A, the bridge helix, and the trigger loop (leucine 1081 as stick model) are highlighted using the color key in (A). Part

of the protein is omitted for clarity.

(F) Sigmaa-weighted 2Fo-Fc electron density map for the nucleic acids in the α -amanitin-inhibited EC. The map is calculated with phases from the refined model and contoured at 1.0 σ . Nucleic acids, bridge helix and metal A of the refined model are superimposed. The view is as in (E).

(G) Unbiased initial Fo-Fc electron density map for α -amanitin and the folded trigger loop. The map is calculated with phases from Pol II alone (Rpb1 residues 1076-1081 and 1092-1096 omitted) and contoured at 2.5 σ . The refined model is superimposed, and helix α 37 is omitted for clarity.

Table 3: X-ray diffraction data and refinement statistics

Crystal	Complete Pol II EC with preserved translocation equilibrium	Complete Pol II EC with α -amanitin
<i>Data collection</i> ^{a, c}		
Space group	C222 ₁	C222 ₁
Unit cell axes (Å)	221.9, 393.1, 283.3	220.6, 394.2, 284.0
Wavelength (Å)	0.91908	0.91908
Resolution (Å)	50-3.60 (3.73-3.60) ^b	50-3.40 (3.52-3.40) ^b
Completeness (%)	99.9 (99.9)	99.9 (99.9)
Unique reflections	277,168 (28,042)	328,642 (32,526)
Redundancy	3.8 (3.6)	3.6 (3.5)
R _{sym} (%)	6.7 (36.8)	13.3 (46.7)
I/ σ I	15.9 (3.9)	8.5 (3.0)
<i>Refinement</i> ^c		
Amino acid residues		3923
Nucleic acid residues		48
Nonhydrogen atoms		32084
RMSD bonds (Å)		0.009
RMSD angles (°)		1.5
R _{cryst} (%)		25.5
R _{free} (%)		28.8
Br peaks in anomalous Fourier (σ) ^d	7.9 (at -3) , 4.9 (at -4)	7.8 (at -4)

^aDiffraction data were collected at beamline X06SA at the Swiss Light Source, Villigen, Switzerland.

^bNumbers in parenthesis refer to the highest resolution shell.

^cFriedel pairs treated as different reflections

^dCalculated using phases from Pol II alone after rigid body refinement at 3.6 Å. The observed position in the template DNA is given in parenthesis.

2.2.2 Structure of the α -amanitin-inhibited complete Pol II EC

To investigate the influence of α -amanitin on the apparent translocation equilibrium, we soaked the inhibitor into an EC crystal, and solved the structure by molecular replacement with the Pol II structure (Figure 6E, chapter 2.4.3). We achieved a resolution of 3.4 Å, which

revealed even subtle conformational changes. Only a single strong bromine peak was observed that unambiguously revealed the hybrid in the post-translocation register (Figure 6D). Strong unbiased difference electron density was observed for the hybrid, for the entire downstream DNA, for α -amanitin, and for the N-terminal region of the previously disordered trigger loop (Figure 6F-G). The structure was rebuilt and fully refined (Table 3, chapter 2.4.3).

The structure revealed that α -amanitin occupies the site observed in the free Pol II core- α -amanitin complex (Bushnell et al, 2002), but shows an altered set of contacts with Pol II (Table 4, Figure 7). In particular, the inhibitor formed two previously unobserved hydrogen bonds with the Rpb1 residue H1085 within the N-terminal part of the trigger loop (Figure 7). Since the trigger loop is disordered in the free EC, α -amanitin restricts trigger loop movement. In addition to the previously observed contact of α -amanitin to the bridge helix residue E822 (Bushnell et al, 2002), several indirect contacts to the bridge helix were observed. In particular, α -amanitin contacts residues Q767, Q768, S769 and G772 in the Rpb1 loop α 23- α 24, which in turn binds the bridge helix residues H816 and E822 with its residues Q768 and E771, respectively (Table 4, Figure 7). The OH group of the amanitin hydroxyproline residue is also near the Rpb1 residue N1082 of the trigger loop, which in turn binds D826 in the bridge helix. Thus, an intricate network of hydrogen bonds and other contacts exists between α -amanitin, the trigger loop, the bridge helix, and loop α 23- α 24.

Table 4: Hydrogen bonds between α -amanitin and Pol II residues

α -amanitin residue	α -amanitin atom (D, donor; A, acceptor)	Pol II residue (BH, bridge helix; TL, trigger loop)	Pol II atom (D, donor; A, acceptor)	Length(Å)	Present in core Pol II α -amanitin complex (Bushnell et al, 2002)
2	OD (A)	H1085 (TL)	NE2 (D)	2.9	no
2	OD (D)	E822 (BH)	OE1 (A)	3.1	yes
2	O (A)	S769	N (D)	3.1	yes
3	OD (A)	Q763	NE2 (D)	2.5	no
3	OG (A)	Q760	NE (D)	3.0	no
3	O (A)	Q768	NE2 (D)	3.4	yes
4	O (A)	R726	NH1 (D)	3.2	yes
4	O (A)	Q767	N (D)	2.8	yes
6	O (A)	N723	ND2 (D)	3.2	no
7	N (D)	H1085 (TL)	ND1 (A)	3.0	no

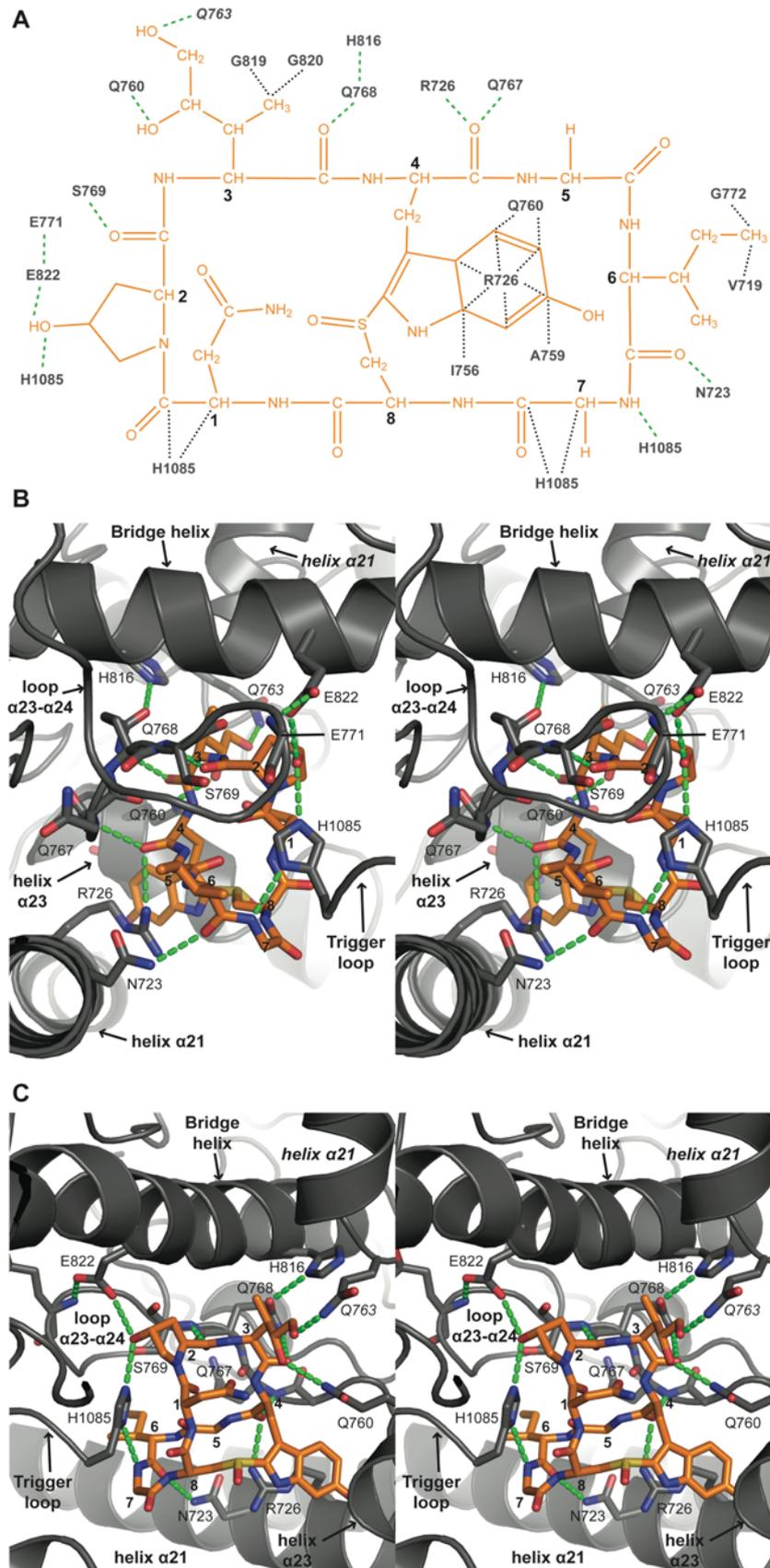


Figure 7: Contacts between α -amanitin and Pol II

(A) Schematic representation of Pol II- α -amanitin contacts. The chemical formula of α -amanitin is shown in orange. Pol II residues in contact with the inhibitor are depicted in grey. All residues are from Rpb1 except for Q763, which belongs to Rpb2 and is in italics. Hydrogen bonds are depicted as green dashed lines, hydrophobic

contacts as black dotted lines. Numbers of α -amanitin residues are shown next to their C α atoms (Wieland & Faulstich, 1991).

(B-C) Stereo views of α -amanitin in its binding pocket. α -amanitin is depicted as a stick model with carbon atoms in orange. Pol II is shown as a ribbon model in grey with selected residues of the binding pocket as stick model with carbon atoms in grey. Nitrogen, oxygen, and sulfur atoms are in blue, red and yellow, respectively. Hydrogen bonds are depicted as green dashed lines. Numbers of α -amanitin residues are shown next to their C α atoms. Pol II elements and residues that contribute to the binding pocket are labeled. Two views are depicted that are related by an approx. 180° rotation around a vertical axis. In (B), the view is from the front, similar to Figure 6G, and in (C) the view is from the back.

2.2.3 The trigger loop forms a wedge next to the central bridge helix

The refined structure showed that binding of α -amanitin locked the Pol II EC into a defined, and previously unobserved conformational state. The bridge helix remained helical along its entire length, but its central part (two turns, residues D826-E833, termed here the central bridge helix) was shifted upstream in the direction of the hybrid, as observed in the unbiased difference Fourier map (Figure 8A, E-F). The side chain of Rpb1 residue A832 moved by 2.5 Å in the direction of DNA template translocation. The trigger loop apparently stabilizes the shifted conformation of the central bridge helix, as its residue L1081 forms a wedge between residue V829 in the bridge helix and residue P1099 in helix α 37 (helix G' in bacterial RNA polymerase) (Figure 8B-D). These observations are consistent with a concerted movement of the central bridge helix and the trigger loop during translocation. The shift of the central bridge helix differs from a previously observed conformational change (Wang et al, 2006) that does not include residues T831 and A832, which exhibit the largest movements in this work.

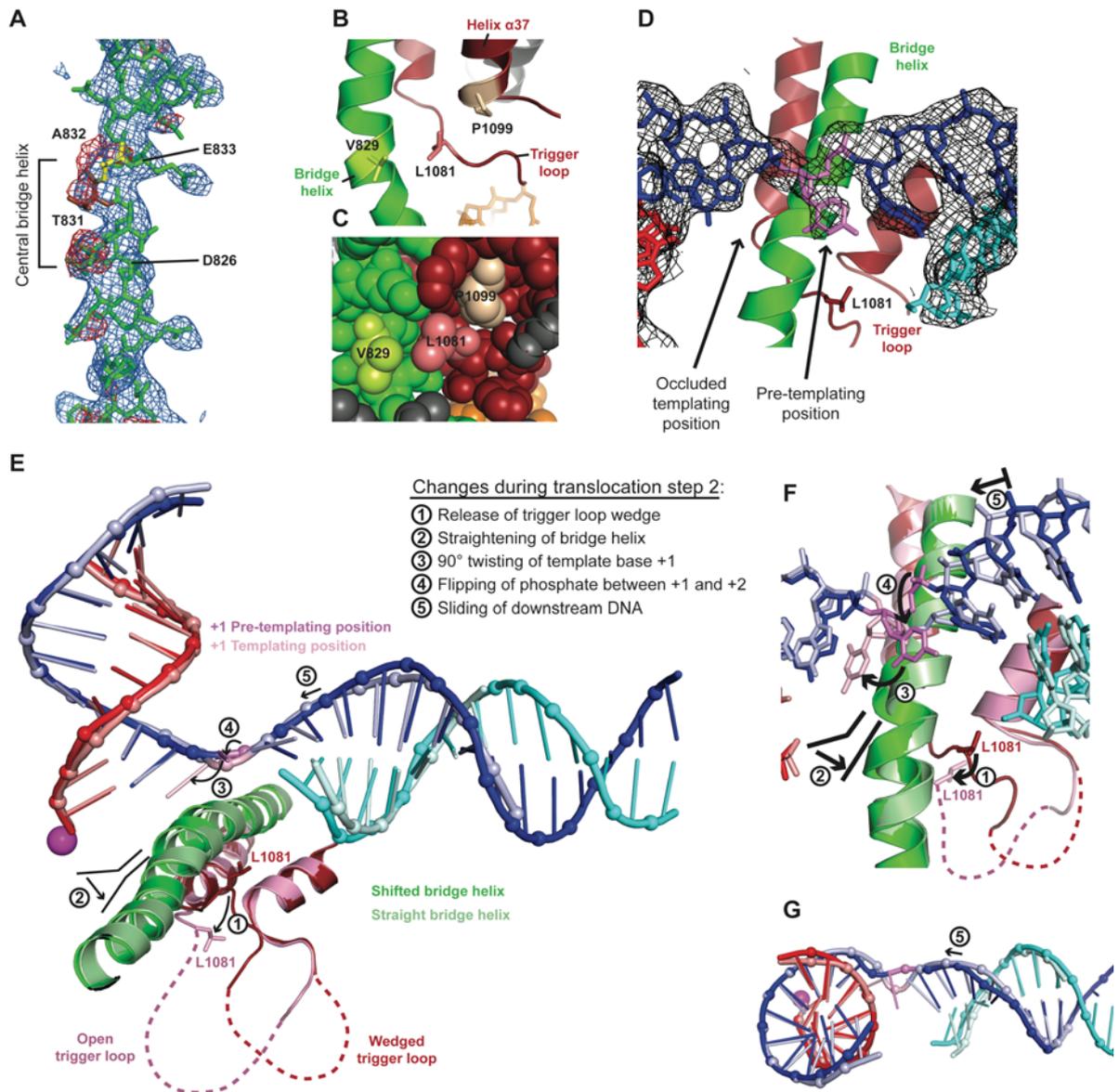


Figure 8: The α -amanitin-inhibited EC structure is a translocation intermediate

(A) The shifted central bridge helix. The Fo-Fc electron density omit map contoured at 2.5σ is shown in blue. The map was calculated with phases from the final model lacking the bridge helix (Rpb1 residues 810-845). The unbiased Fo-Fc difference map, calculated with initial phases from the Pol II model, is contoured at 2.5σ and shown in red. The final model of the bridge helix is superimposed, and residues A832 and T831 are highlighted in yellow and brown, respectively.

(B-C) The trigger loop forms a wedge between the bridge helix and helix α 37. Top views (as in Figure 4, right) of a ribbon model (B) and a space-filling model (C) are shown. The wedging trigger loop residue L1081 and the residues flanking it (V829, P1099) are shown in lighter colors.

(D) The +1 DNA template base adopts a pre-templating position. The initial unbiased Fo-Fc difference map for the nucleic acids is shown around the +1 position, contoured at 2.5σ . The +1 base in the pre-templating site is highlighted in violet. The view is from the top.

(E-G) Transition from the translocation intermediate to the post-translocation state. The structures of the translocation intermediate and the post-translocation EC (Kettenberger et al, 2004) were superimposed. Metal A, nucleic

acids, the bridge helix, and the trigger loop including its flanking helices α 36 and α 37 are shown. The elements of the post-translocated EC structure are depicted in lighter colors. Phosphates and bases are depicted as balls and sticks respectively. Conformational transitions during translocation step 2 that are required to convert the intermediate to the post-translocation EC are indicated with arrows and numbered 1-5. The view is from the side (E) as in Figure 6E or from the top (F)-(G), as in Figure 8B.

2.2.4 The new state of the EC is a translocation intermediate

The shifted central bridge helix and wedged trigger loop apparently stabilize the nucleic acids in a previously unobserved conformation intermediary between the pre- and post-translocation conformations. Whereas the hybrid adopts the post-translocation position (Figure 6D, Figure 8E, Figure 8G), the downstream DNA template adopts an intermediary position that corresponds to incomplete translocation (Figure 8E-G). The incoming DNA template base at register +1 adopts a previously unobserved position above the central bridge helix (Figure 8D-G). We refer to this new template base position as the “pre-templating” position. In the free EC, the +1 base occupies a standard “templating” position opposite the NTP-binding site. In the templating site, the +1 base is twisted by 90° with respect to the +2 base, and stacks against residues A832 and T831 in the central bridge helix (Figure 9C-D). In the new structure, the templating site is occluded by the shifted central bridge helix, in particular by the A831 side chain (Figure 9C-D). As a consequence, the +1 base is not twisted with respect to the +2 base, and retains contacts with the +2 base downstream, although the base-stacking is discontinued (Figure 8F).

This analysis identifies the state of the Pol II EC that is induced by α -amanitin binding to inter-converting pre- and post-translocation ECs as a translocation intermediate. The EC is apparently trapped after translocation of the hybrid to the post-translocation position, but before complete delivery of the +1 base to the templating position opposite the active site, and thus before full establishment of the post-translocation state.

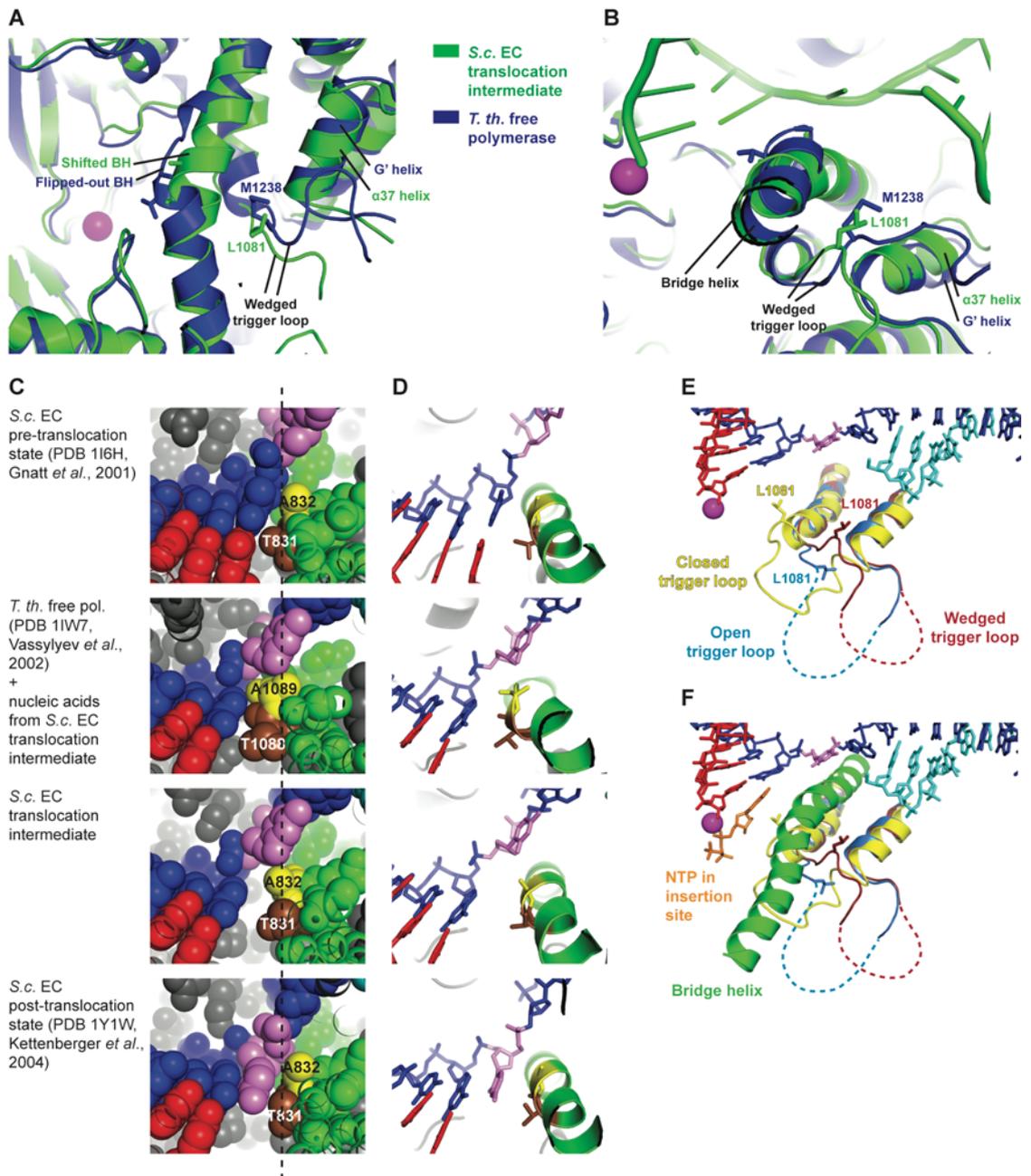


Figure 9: Comparisons with bacterial RNA polymerase and possible EC states

(A-B) Superposition of the trigger loops and bridge helices in the Pol II EC translocation intermediate and the free *Tth* RNA polymerase (Vassilyev *et al.*, 2002). The trigger loop residue L1081 (Pol II) or its homologous residue M1238 (*Tth*) forms a wedge between the bridge helix and helix α 37 in Pol II or G' in *Tth*. The views are from the top (A) or the side (B), as in Figure 8B or Figure 6E, respectively. In the Pol II EC translocation intermediate, the central bridge helix is shifted, whereas it adopts a flipped-out conformation in the bacterial holoenzyme.

(C-D) Four possible states of the EC. From top to bottom, the pre-translocation state (Gnatt *et al.*, 2001) (PDB 1I6H), a potential additional intermediate or a transition state with a modeled flipped-out bridge helix (Vassilyev *et al.*, 2002) (PDB 1IW7), the EC translocation intermediate with the shifted bridge helix (this study), and the post-translocation state (Kettenberger *et al.*, 2004) (PDB 1Y1W) are shown with space-filling models (C) or ribbon diagrams (D). The bridge helix residues A832/A1089 (Pol II/*Tth*) and T831/T1088 (Pol II/*Tth*) are highlighted in yellow and brown, respectively.

(E-F) Comparison of trigger loop conformations. Pol II EC structures in the post-translocation state (Kettenberger

et al, 2004) (PDB 1Y1W), with bound NTP substrate (Wang et al, 2006) (PDB 2E2H), and in the intermediary state are superimposed. Nucleic acids and metal A are from the translocation intermediate. The trigger loops of the three structures are depicted in dark red (wedged, translocation intermediate), light blue (open, 1Y1W) and yellow (closed, 2E2H). (F) Also depicted are the bridge helix (green, translocation intermediate) and the NTP in the insertion site (orange, 2E2H).

2.3 Discussion

2.3.1 Structure of a translocation intermediate

Here we report the structure of the Pol II EC in a new, α -amanitin-inhibited state, which is intermediary between the pre- and post-translocation states. In this intermediate structure, the DNA-RNA hybrid adopts the post-translocation position, and the downstream DNA template a new position between the pre- and post-translocation register. The +1 base in the template strand occupies a previously unobserved pre-templating site above the bridge helix. The central bridge helix is shifted upstream and partially occludes the templating site that is occupied by the +1 base in the pre- and post-translocation states. The trigger loop adopts a position where it forms a wedge that apparently stabilizes the shifted bridge helix.

2.3.2 Extended NAC model

Combining the new intermediate structure with known EC structures of Pol II (Gnatt et al, 2001; Kettenberger et al, 2004; Wang et al, 2006) and a bacterial RNA polymerase (Vassilyev et al, 2007b) leads to a model of the NAC (Figure 10) that extends previous models (Bar-Nahum et al, 2005; Cramer, 2007; Epshtein et al, 2002; Gnatt et al, 2001; Landick, 2004; Vassilyev et al, 2007b). In this model, NTP binds to the post-translocated EC in a pre-insertion state that maintains an open, flexible trigger loop (Kettenberger et al, 2004; Vassilyev et al, 2007b). Subsequent movement of the NTP to the insertion site goes along with closure of the trigger loop and formation of the catalytically active polymerase conformation (Vassilyev et al, 2007b; Wang et al, 2006). Catalytic incorporation of the nucleotide into the RNA leads to pyrophosphate formation. Catalysis *per se* does most likely not lead to a major conformational change in the active center, since the contacts that the trigger loop forms with the NTP before catalysis may be maintained with the newly incorporated nucleotide and the pyrophosphate after catalysis. Release of pyrophosphate however would enable movement of the trigger loop to the wedged position, facilitating a shift in the central bridge helix, which accompanies movement of the nucleic acids to the intermediary state (translocation step 1). Release of the wedge and relaxation of the bridge helix to the straight position

enables movement of the nucleic acids to the post-translocation state (translocation step 2). This frees the NTP-binding site for the next incoming substrate and completes the NAC.

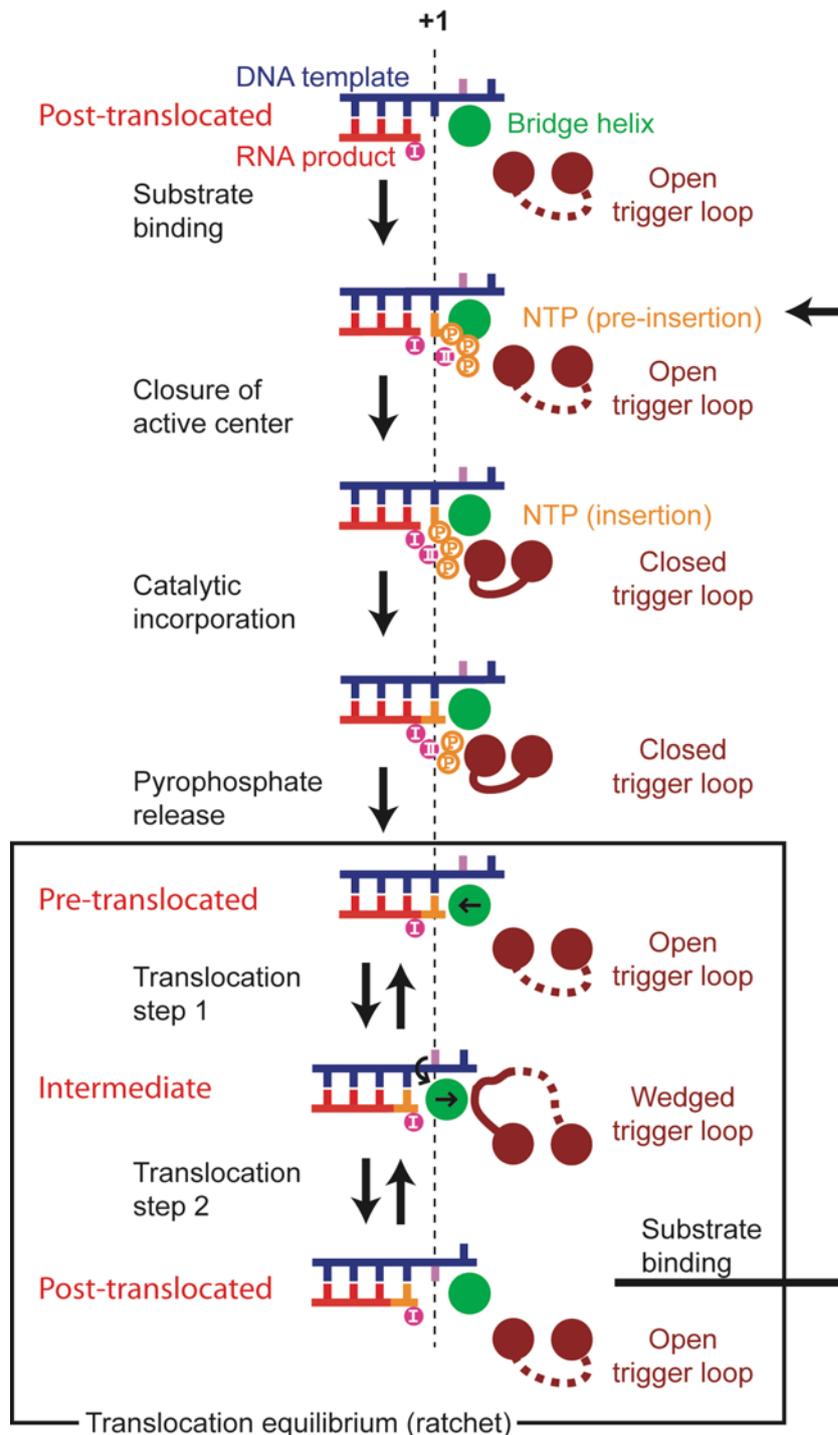


Figure 10: Model of the nucleotide addition cycle (NAC)

Schematic representation of the NAC. The vertical dashed line indicates register +1. For details refer to text.

2.3.3 Stepwise translocation and helicase activity of Pol II

Our data suggest a two-step mechanism of nucleic acid translocation via an intermediate (compare Figure 8E-G, Figure 10). During step 1, the hybrid moves from the pre- to the post-translocation position, and the downstream DNA translocates until the next DNA template base (register +2) reaches the pre-templating site. During step 2, the DNA base twists by 90° , to reach its templating position in the active center (register +1). This twisting is accompanied by a flipping of the phosphate backbone group between the DNA template bases +1 and +2, and a sliding of the downstream DNA to the post-translocation position. In the intermediate, many contacts between the polymerase and the nucleic acids are maintained. This likely preserves EC stability during translocation and decreases the energy barrier between pre- and post-translocation states, to facilitate movement of the nucleic acids over the Pol II surface (Figure 11).

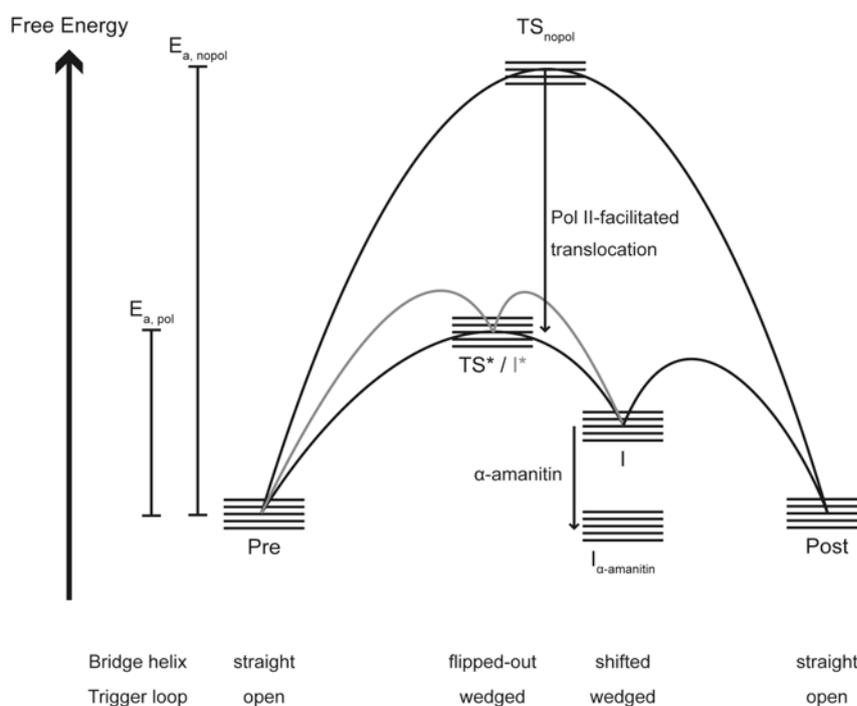


Figure 11: Hypothetical free energy diagram of translocation

The precise energy level of the pre-translocation state (Pre) and the post-translocation state (Post) and the possible intermediates depends on the nucleic acid sequence and probably other factors. The activation energy of a hypothetical non-polymerase facilitated translocation ($E_{a,nopol}$) is high, because breaking all Pol II-nucleic acid contacts is required to reach the corresponding hypothetical transition state (TS_{nopol}). The polymerase presumably lowers the activation energy by maintaining many contacts during translocation ($E_{a,pol}$). The facilitated-translocation pathway includes the translocation intermediate (I), shown in this work to be stabilized by α -amanitin ($I_{\alpha\text{-amanitin}}$) (Figure 9C-D). The translocation intermediate may be reached directly via the transition state TS^* or via another intermediate (I^* , grey graph), which may resemble the state found in free bacterial RNA polymerase holoenzyme (Vassilyev et al, 2002) (1IW7, Figure 9C-D).

The EC acts like a helicase on the incoming DNA duplex that apparently melts the base pair +2 downstream of the active site (Kashkina et al, 2007; Vassilyev et al, 2007a). Modeling

shows that the DNA base in the pre-templating site of the intermediate cannot remain base paired with its counterpart in the non-template strand, because the path of the template strand strongly deviates from the course of canonical B-DNA as suggested (Kettenberger et al, 2004), and the path of the non-template strand is blocked by fork loop 2 (Gnatt et al, 2001; Kettenberger et al, 2004; Naji et al, 2007; Wang et al, 2006). Available data thus suggest that RNA polymerase melts the downstream DNA base pair during translocation step 1.

2.3.4 Comparison with bacterial RNA polymerase

Previous structural and biochemical data for bacterial RNA polymerases had suggested the existence of a translocation intermediate and support our findings and interpretations. In the structures of the free *T. thermophilus* (*Tth*) RNA polymerase holoenzyme (Vassylyev et al, 2002) and the *T. aquaticus* (*Taq*) RNA polymerase bound to two different antibiotics (Campbell et al, 2005), the trigger loop residue M1238 intrudes between the bridge helix and helix G' ($\alpha 37$ in Pol II), similar to the corresponding Pol II trigger loop residue L1081 that forms a wedge in the intermediate (Figure 8B-C, Figure 9A-B). In these bacterial polymerase structures, the central bridge helix adopts a flipped-out conformation in the region that is shifted in the Pol II EC intermediate (*Tth/Taq* residues 1087-1092).

RNA-protein crosslinking suggested that the flipped-out conformation reflects an intermediary state between the pre- and post-translocation states, and that the bridge helix and trigger loop undergo a cooperative structural transition during nucleotide incorporation and translocation (Epshtein et al, 2002). Furthermore, crosslinking studies with *E.coli* RNA polymerases that carried mutations in the C-terminal part of the trigger loop suggested a role of the trigger loop in controlling bridge helix movement (Bar-Nahum et al, 2005). Finally, mutation of residue M932 in *E.coli* RNA polymerase, which corresponds to the wedged trigger loop residue L1081 in Pol II, strongly increases the duration of transcriptional pausing (Toulokhonov et al, 2007), consistent with an important role of this residue in translocation.

2.3.5 A general translocation mechanism

These published data and our results converge on a conserved translocation mechanism for all multisubunit RNA polymerases, which involves a trigger loop-stabilized EC intermediate with an altered structure of the central bridge helix. Due to the very high conservation of the active center in all multisubunit RNA polymerases, the observed Pol II EC intermediate conformation likely exists in the bacterial EC, although this remains to be shown, since the central bridge helix is straight in the first bacterial EC structure (Vassylyev et al, 2007a). It also remains to be demonstrated by structural analysis if the flipped-out conformation of the bridge helix can be adopted within an EC. Modeling shows that the flipped-out bridge helix can be accommodated within the Pol II EC intermediate structure (Figure 9C-D). In this

model, the *Tth* bridge helix residues T1088 and A1089 (T831 and A832 in Pol II) stack against the -1 base of the DNA template strand (Figure 9C-D). This EC conformation, if it exists, could represent a second, slightly different translocation intermediate (if it can be isolated; Figure 11, I*), or the transition state of translocation step 1 (if the central bridge helix is only transiently flipped out; Figure 11, TS*).

2.3.6 The Brownian ratchet and transcription fidelity

Whereas our studies provide insights into the structural transitions of the EC during translocation, a Brownian ratchet model may explain the directionality of translocation ((Bar-Nahum et al, 2005; Landick, 2004; Sousa, 2005) and references therein). The model assumes that the ground state of the EC is characterized by an equilibrium between rapidly inter-converting pre- and post-translocation states (the ratchet), enabled by an oscillation between different bridge helix conformations. NTP binding temporarily stops the oscillation, acting like a pawl of a ratchet. After nucleotide incorporation, oscillation resumes around the next template position. In this model, the energy for translocation comes from Brownian fluctuations, and the directionality of translocation results from trapping the forward fluctuation by NTP binding and nucleotide incorporation (Figure 10).

Our data support the Brownian ratchet mechanism. First, the proposed equilibrium between pre- and post-translocation states is observed within crystals as coexisting, inter-converting states of the EC. Second, the interruption of bridge helix oscillation by NTP binding may be rationalized by the mutually exclusive functions of the trigger loop during bridge helix shifting and NTP incorporation. The trigger loop residue L1081 (M1238 in *Tth/Taq*) forms a wedge at the shifted bridge helix in the intermediate (Figure 8B-C, Figure 9A-B), but moves by up to 13 Å to close over the nucleotide during incorporation (Figure 9E-F)(Vassylyev et al, 2007b; Wang et al, 2006). Thus swinging of the trigger loop between closed and wedged positions is apparently important for the NAC and the Brownian ratchet.

Trigger loop swinging may also be important for transcription fidelity. We have previously shown that an abasic site at template position $+1$ causes non-templated nucleotide incorporation (Damsma et al, 2007). This could also occur in the intermediate, which lacks a base in the templating position, but exhibits enough space to bind the NTP. However, nucleotide incorporation is not favored in the intermediate, since the trigger loop adopts the wedged position and is not available for active site closure.

2.3.7 Single-subunit RNA polymerase and a possible power stroke

The mechanism of substrate incorporation and translocation was very well studied for the single-subunit RNA polymerase from bacteriophage T7 (Temiakov et al, 2004; Yin & Steitz, 2004). This polymerase shares no sequence similarity with the multisubunit enzymes, but the

arrangement of nucleic acids and many mechanistic aspects are similar (Cramer, 2002a; Landick, 2004). Structures for the T7 EC include a product complex, which contained the nucleic acids in the pre-translocation state after nucleotide incorporation, and retained the pyrophosphate (Yin & Steitz, 2004). In this structure, the so-called O-helix binds the DNA template strand and the pyrophosphate with its C- and N-terminal ends, respectively. Pyrophosphate release apparently triggers a 22° rotation of the O-helix that results in translocation (Yin & Steitz, 2004). This observation suggested that the catalytic event induces a conformational change that actively drives translocation in a power stroke mechanism, an attractive alternative to the Brownian ratchet.

In T7 RNA polymerase, a single structural element, the O-helix, is involved in both pyrophosphate binding and template translocation, whereas Pol II uses two elements for these functions, the trigger loop and the bridge helix. A product complex structure is not available for Pol II, but the pyrophosphate may retain the trigger loop in a closed position, and only its release may re-mobilize the trigger loop, which could then adopt the wedged position, and re-establish the translocation equilibrium of a ratchet. Since the trigger loop passes through a mobile state between the closed and wedged positions, it cannot directly transmit a possible force generated by the catalytic event, just like a rope cannot transmit a pushing force. However, in molecular motors a power-stroke may be created by a disorder-order transition (Block, 2007; Vale & Milligan, 2000). A brownian ratchet mechanism is also supported by single-molecule experiments. Galburt et al could explain their observed translocation and pausing behavior with a diffusive model, indicative of a brownian ratchet (Galburt et al, 2007), and Abbondanzieri et al investigated the force dependence of transcription by bacterial RNAP and found that their data could be well described by brownian ratchet, but not power stroke models (Abbondanzieri et al, 2005). Taken together, it seems unlikely that in Pol II the bridge helix drives forward translocation in an active, power-stroke mechanism. We rather remain that the bridge helix accompanies nucleic acid translocation by a passive conformational change that is facilitated by the trigger loop wedge.

2.3.8 Mechanism of α -amanitin inhibition

Our data suggest that α -amanitin inhibits Pol II by trapping the wedged trigger loop and shifted bridge helix, thereby stabilizing the translocation intermediate (Figure 11). This model predicts that α -amanitin interferes with both nucleotide incorporation and translocation, and is consistent with published functional data. α -amanitin strongly reduces the polymerase elongation rate, but does not influence its NTP affinity, and does not abolish multiple nucleotide additions (Chafin et al, 1995; Rudd & Luse, 1996). Our elongation assays provided consistent results (Figure 6B). Kinetic studies distinguished three EC conformations with different sensitivities for α -amanitin (Gong et al, 2004). It appears that these EC conformations, which

showed high, medium, or no sensitivity to α -amanitin inhibition (Gong et al, 2004), correspond to the intermediate, the pre-, and the post-translocation state, respectively.

A network of direct and indirect contacts between α -amanitin, the bridge helix, and the trigger loop apparently underlies polymerase inhibition. Analysis of amanitin derivatives showed that the hydroxyl group of the hydroxyproline residue 2 of α -amanitin is crucial for inhibition (Wieland & Faulstich, 1991; Zanotti et al, 1992). The importance of this hydroxyl group can now be ascribed not only to its contact to the bridge helix residue E822 (Bushnell et al, 2002), but also to its interaction with the trigger loop residue H1085 and a possible contact to N1082, which in turn contacts the bridge helix residue E826 (Figure 7). Since α -amanitin stays attached to the active EC (Rudd & Luse, 1996), some of its contacts with Pol II, in particular with the trigger loop and bridge helix, get apparently broken during the NAC. The energy required to break these contacts may be responsible for the strong reduction in the elongation rate in the presence of the toxin.

2.4 Specific experimental procedures

2.4.1 Preparation of Pol II elongation complexes

Endogenous *S. cerevisiae* ten-subunit Pol II core enzyme and recombinant Rpb4/7 heterodimer were purified as described in chapters 4.2 and 4.4. Nucleic acid scaffolds (Figure 6A) were annealed by mixing equimolar amounts of synthetic template DNA, nontemplate DNA, and RNA in TE buffer (chapter 4.6) at a final concentration of 100 μ M, heating the mixture to 95°C for 5 min, and slow-cooling to 4°C in a thermocycler. Stoichiometric Pol II ECs were assembled by incubating core Pol II for 10 min with 2 molar equivalents of nucleic acid scaffold, followed by 20 min incubation with 5 molar equivalents of recombinant Rpb4/7 in assembly buffer (chapter 4.6) at 20°C. The complexes were purified by size exclusion chromatography (Superose 6 10/300 GL) in Pol II buffer.

2.4.2 Crystallization and crystal treatment

Purified Pol II ECs were concentrated to 3.5-4.5 mg/ml and an additional amount of the nucleic acid scaffold was added prior to crystallization to a final concentration of 2 μ M. Crystals were grown at 22 °C with the hanging drop vapor diffusion method by mixing 2 μ l of sample solution with 1 μ l of reservoir solution (NH₄-Na-acetate crystallization solution, chapter 4.7). Crystals were harvested in mother solution after 10-20 days, when they had reached their maximum size (approx. 0.4 x 0.3 x 0.2 mm). Crystals were transferred stepwise over 5 hours to mother solution containing additionally 0-22 % glycerol. For complex formation with α -amanitin (Sigma), the crystals were incubated for 4 hours in mother solution containing

additionally 22 % glycerol, 5 mM magnesium acetate and 100 μ M α -amanitin. Crystals were slowly cooled to 8 °C, and flash-cooled by plunging into liquid nitrogen.

2.4.3 X-ray structure analysis

Diffraction data were collected in 0.25 ° increments at the protein crystallography beamline X06SA of the Swiss Light Source using the new Pilatus 6M pixel detector (Broennimann et al, 2006) (Table 3). Raw data were processed with XDS (Kabsch, 1993). Structures were solved by molecular replacement with the program Phaser (McCoy et al, 2005), using the structure of the complete 12-subunit Pol II elongation complex without nucleic acids as a search model (PDB 1Y1W, (Kettenberger et al, 2004)). The molecular replacement solution was subjected to rigid body refinement with CNS version 1.2 (Brunger, 2007) using five rigid groups (core, jaw-lobe, shelf, clamp, Rpb4/7). Model building was carried out with programs Coot (Emsley & Cowtan, 2004), O (Jones et al, 1991) and Moloc (Gerber & Muller, 1995). The nucleic acids were built into the initial F_o-F_c electron density map. The register of the nucleic acids was unambiguously defined by bromine labeling as described (Brueckner et al, 2007). Shifted parts of the trigger loop and the bridge helix were manually rebuilt into omit maps (Figure 6G, Figure 8A). A model for α -amanitin (Bushnell et al, 2002) was manually placed into the initial F_o-F_c map (Figure 6G) and further adjusted manually and by restrained real space refinement with Moloc. Atomic positions and B-factors were refined with CNS version 1.2. Refinement parameters for α -amanitin were generated using Moloc. Refinement was monitored with the free R factor, calculated from the same 2 % set of excluded reflections as in refinement of complete Pol II (Armache et al, 2005) and the complete Pol II EC (Kettenberger et al, 2004), extended to the higher resolution of 3.4 Å. The increase in resolution was apparently due to differences in the crystallization protocol, use of a new, highly sensitive pixel detector with increased signal/noise ratio, and an improved processing strategy, including the use of XSCALE with zero dose extrapolation for scaling and CNS version 1.2 with bulk solvent parameter grid search for refinement and map calculation. Structural superpositions were performed with the secondary structure matching function of Coot (Emsley and Cowtan, 2004). Figures were prepared with Pymol (DeLano, 2002).

2.4.4 Transcript extension assays

Stoichiometric ECs of complete Pol II, containing the 10-subunit core and Rpb4/7, were assembled and purified as described above, but in transcription buffer without $MgSO_4$ (20 mM Hepes pH 7.6, 60 mM ammonium sulfate, 10 μ M $ZnCl_2$, 10 % glycerol, and 10 mM DTT, compare chapter 4.6). α -amanitin-inhibited Pol II ECs were prepared by incubating ECs at 20 °C with α -amanitin for 10 min. The RNAs used for extension assays were identical to those used for structural studies, except for five additional nucleotides (5'-UGCAU-3') and a fluo-

resceine label at their 5'-end. For transcript extension, ECs or α -amanitin-inhibited ECs were incubated with NTPs and 10 mM magnesium sulfate at 28 °C in transcription buffer. Reactions were stopped by incubating with an equal volume of gel loading buffer (90% (v/v) formamide, 50 mM EDTA pH 7.4) for 5 min at 95 °C. The RNA products were separated by denaturing gel electrophoresis (0.5 pmol RNA per lane, 0.4 mm 20 % polyacrylamide gels containing 7 M urea, 50-55 °C) and visualized with a Typhoon 9400 scanner (GE Healthcare).

3 CPD damage recognition by transcribing RNA polymerase II

3.1 Introduction

Ultraviolet light damages cellular DNA by inducing dimerization of adjacent pyrimidines in a DNA strand. The resulting cyclobutane pyrimidine dimer (CPD) lesions can block transcription and replication, and are a major cause of skin cancer (Mitchell et al, 2003). Cells eliminate CPDs by nucleotide excision repair (NER). A very efficient NER subpathway is transcription-coupled DNA repair (TCR), which specifically removes lesions from the DNA strand transcribed by Pol II (Saxowsky & Doetsch, 2006). Pol II stalls when a CPD in the DNA template strand reaches the enzyme active site (Mei Kwei et al, 2004; Tornaletti et al, 1997). Pol II stalling apparently triggers TCR by recruitment of a transcription-repair coupling factor (Rad26 in yeast, CSB in humans) and factors required for subsequent steps of NER, including TFIIH, which comprises helicases that unwind DNA (Saxowsky & Doetsch, 2006; Svejstrup, 2002). Endonucleases then incise the DNA strand on either side of the lesion, resulting in a 24-34 nucleotide fragment (Mu & Sancar, 1997; Selby et al, 1997; Tremeau-Bravard et al, 2004). The obtained DNA gap is subsequently filled by DNA synthesis and ligation (Prakash & Prakash, 2000; Sancar, 1996).

3.2 Results and discussion

3.2.1 The molecular mechanism of CPD damage recognition by RNA polymerase II

To elucidate CPD recognition by transcribing Pol II, we carried out a structure-function analysis of elongation complexes containing in the template strand a thymine-thymine CPD. Elongation complexes were reconstituted from the 12-subunit *Saccharomyces cerevisiae* Pol II and nucleic acid scaffolds as previously described (Kettenberger et al, 2004), except that the mobile upstream DNA and the non-template strand in the transcription bubble were omitted (minimal nucleic acid scaffold, chapter 3.2.5.1). A chemical analogue of a CPD lesion was incorporated at register +2/+3 of the template strand, directly downstream of position +1, which denotes the substrate addition site (scaffold A, Figure 12A; chapter 3.2.5.2). The crystal structure of the resulting elongation complex A was determined (Figure 12B; chapter 3.3.3), and the register of the nucleic acids was unambiguously defined by bromine labeling (Table 5; Figure 13C). The overall structure of complex A was nearly identical to the complete Pol II elongation complex (Kettenberger et al, 2004) and very similar to elongation

complex structures of the Pol II core enzyme (Gnatt et al, 2001; Westover et al, 2004a) (chapter 3.2.5.3). The template strand enters the active site, and continues into an eight base pair hybrid duplex with RNA, which occupies the upstream positions -1 to -8 (Figure 12B; chapter 3.2.5.1). In contrast to the damage-free elongation complex (Kettenberger et al, 2004), downstream DNA entering the cleft is mobile, indicating that the CPD at positions $+2/+3$ loosens the grip on downstream DNA (chapter 3.2.5.2).

Table 5: X-ray diffraction data and refinement statistics of CPD lesion containing Pol II ECs A-D.

CPD-containing elongation A complex	B	C	D	
<i>Data collection^a</i>				
Space group	C222 ₁	C222 ₁	C2	C222 ₁
Unit cell axes (Å)	222.2, 392.2, 284.0	223.6, 393.5, 283.5	394.4, 221.9, 283.1	222.6, 392.8, 282.9
Unit cell β angle (°)	90	90	90.6	90
Wavelength (Å)	0.9191	0.9190	0.9192	0.9189
Resolution range (Å)	50-3.8 (3.94-3.80) ^b	50-4.0 (4.14-4.00) ^b	50-3.8 (3.94-3.80) ^b	50-3.80 (3.94-3.80) ^b
Unique reflections	120,477 (11,664)	104,987 (10,232)	238,284 (23,709)	121,341 (11,999)
Completeness (%)	99.0 (96.8)	99.8 (98.2)	99.6 (99.6)	99.9 (99.7)
Redundancy	4.6 (3.4)	5.8 (4.5)	3.3 (3.2)	5.1 (4.0)
Mosaicity (°)	0.50	0.50	0.47	0.34
R _{sym} (%)	12.0 (45.4)	8.7 (41.1)	9.8 (43.1)	8.3 (37.3)
I/ σ (I)	10.3 (2.1)	15.8 (3.1)	10.2 (2.3)	17.1 (3.2)
<i>Refinement</i>				
Nonhydrogen atoms	31,665	32,016	63,934	32,005
RMSD bonds	0.010	0.009	0.010	0.009
RMSD angles	1.55	1.52	1.60	1.55
R _{cryst} (%)	27.0	29.2	25.7	26.9
R _{free} (%)	28.0	30.1	27.5	28.8
Br peak in anomalous Fourier (σ)	11.4	5.3	10.7, 10.6	6.4
PDB accession code	2ja5	2ja6	2ja7	2ja8

^aAll diffraction data were collected at the Swiss Light Source, Paul-Scherrer-Institut, Villigen, Switzerland.

^bNumbers in parenthesis correspond to the highest resolution shell.

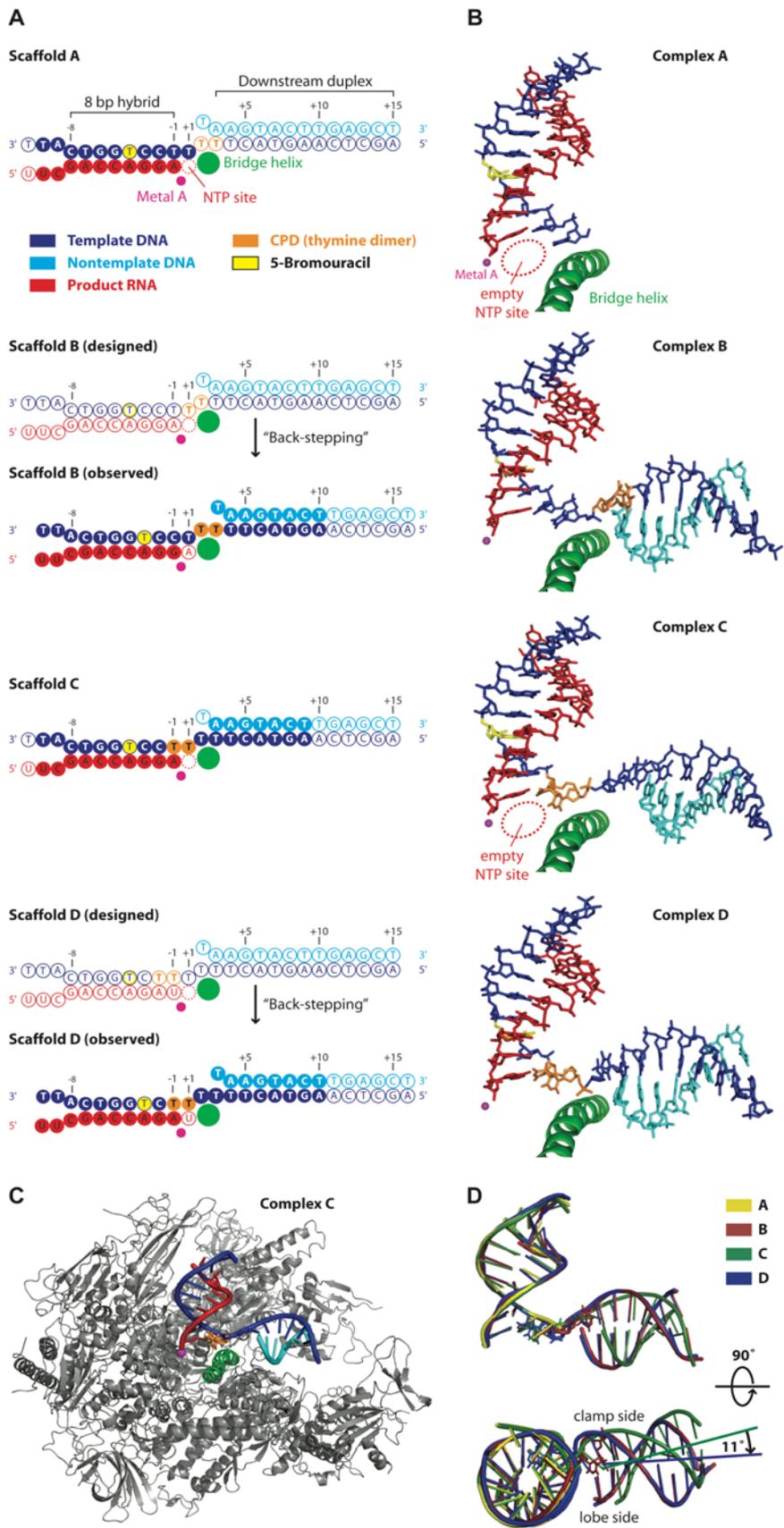


Figure 12: Pol II elongation complex structures with thymine-thymine CPD lesions in the template.

(A) Nucleic acid scaffolds A-D. The color code is used throughout. Filled circles denote nucleotides with interpret-

able electron density that were included in the structures in (B). Open circles denote nucleotides with non-interpretible or absent electron density.

(B) Structure of nucleic acids in the Pol II elongation complexes A-D. The view is from the side as in Figure 4, left. Figures prepared with PYMOL (DeLano Scientific).

(C) Overview of complex C with a CPD lesion at the active site. The view is as in (B). Protein is in grey, the bridge helix in green. The CPD is shown as a stick model in orange. A large portion of the second largest Pol II subunit was omitted for clarity.

(D) Superposition of nucleic acids in structures A-D. The protein molecules were superimposed and then omitted. The nucleic acids are depicted as ribbon models, the CPDs as stick models. Upper and lower views are related by a 90° rotation around a horizontal axis.

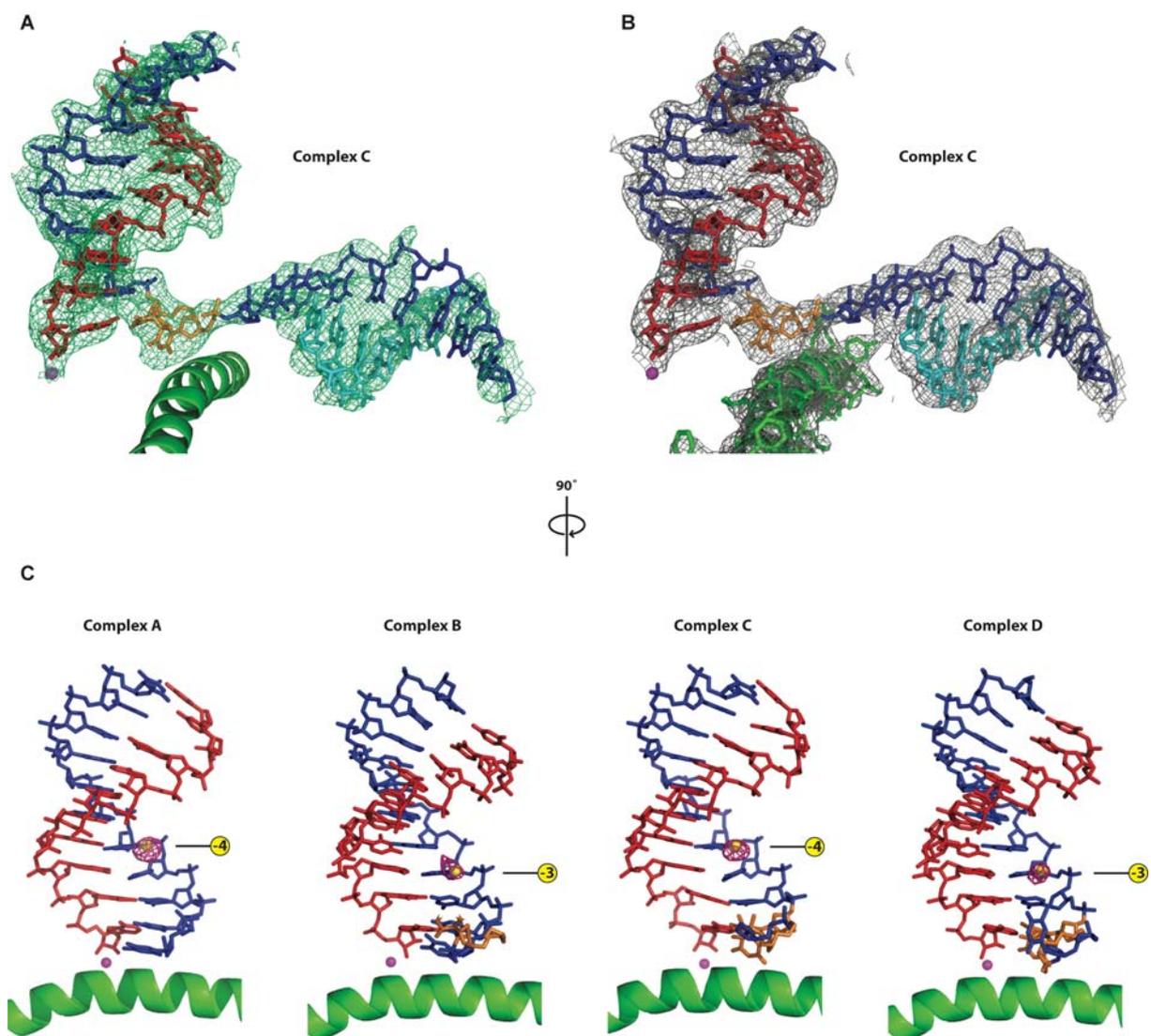


Figure 13: Electron density maps of CPD lesion containing Pol II ECs.

Electron density maps were calculated from the datasets shown in Table 5. The final model of the nucleic acids, metal A and the bridge helix are superimposed (color code as in Figure 12A).

(A) Unbiased F_o-F_c difference Fourier map (green) of complex C calculated after rigid body refinement with phases from the complete Pol II elongation complex without nucleic acids, contoured at 2.5σ . The view is as in Figure 12B (side view).

(B) $2F_o-F_c$ map (grey) of complex C calculated with phases from the refined structure C and contoured around the nucleic acids and the bridge helix (side chains shown as sticks) at 1.0σ .

(C) Bromine anomalous difference Fourier maps (pink) calculated with phases from the refined structures A, B, C, and D, contoured at 6.0σ , 3.8σ , 6.0σ , and 3.8σ , respectively. The observed peak heights are between 5.3σ and 11.4σ (Table 5), and coincide with the location of the bromine atom in a 5-bromouracil residue incorporated in the DNA template strand, thereby defining the register of nucleic acids. When the anomalous difference Fourier maps in complexes B and D were contoured at the low level of 2.5σ , there were no peaks at potential bromine positions in neighboring nucleotides. Thus there were no significantly populated alternative states with different nucleic acid registers, such as non-back-stepped states in complexes B and D. However we cannot exclude that a very minor portion of the complexes adopt a non-backstepped state. We can, however, exclude that the lower bromine peak heights in complexes B and D result from diffraction data collection at a wavelength different from the bromine anomalous peak since crystals were subjected to fluorescence scans and measured at the highly wavelength-stable SLS beamline PX1. With respect to (B), the models in (C) are rotated by 90° around a vertical axis.

To investigate Pol II stalling at the CPD lesion, we incubated complex A with nucleoside triphosphate (NTP) substrates, followed RNA extension by fluorescence-monitored capillary electrophoresis, and identified the RNA products by mass spectrometry (Figure 14; Figure 15; chapter 3.3.4). After incubation with a physiological concentration of 1 mM NTPs for one hour, the RNA was extended by three nucleotides (Figure 14A), but not any further. Thus the complex stalled after nucleotide incorporation opposite both CPD thymines. A time course showed that the first incorporation event was fast, consistent with a free substrate site in complex A (Figure 12A-B; Figure 14B; chapter 3.2.5.4). The second and third incorporation events however were progressively slower, with rate constants of approx. 16 h^{-1} and 2.4 h^{-1} , respectively (Figure 14A-B).

The second incorporation event is opposite the 3'-thymine of the CPD. Efficient incorporation required a concentration of 1 mM NTPs. Lower substrate concentrations limited RNA extension to one nucleotide after 5 min (Figure 17A). In the undamaged elongation complex template bases in positions +1 and +2 are twisted against each other by 90° (Kettenberger et al, 2004). To enter the active site the +1 template base presumably rotates by 90° from the "pre-templating" to the templating position (chapter 2.3.3). The CPD thymines are covalently linked and therefore can only rotate together. Modeling suggests that rotating the 3'-thymine to the +1 templating position would result in a severe clash of the 5'-thymine with the bridge helix in the straight conformation (not shown). Therefore, a significant conformational change in the elongation complex would be required to allow the 3'-thymine of the CPD to serve as a templating base for a nucleotide incorporation.

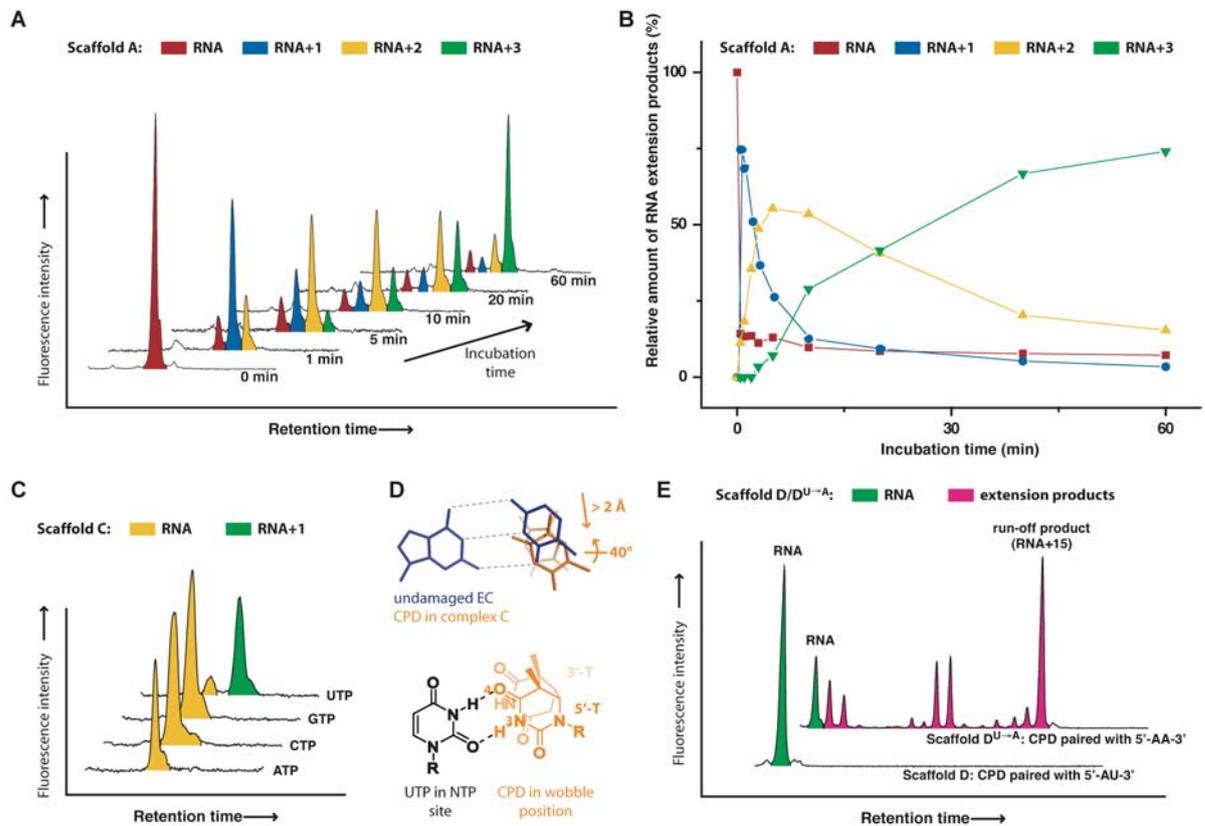


Figure 14: RNA extension assays

(A) Electropherograms of time-dependent extension of RNA in complex A. A stoichiometric complex of complete Pol II and scaffold A (Figure 12A) was incubated with 1 mM ATP, CTP, GTP and UTP. Reactions were stopped at the given time points. RNA products were subjected to fluorescence-monitored capillary electrophoresis, and identified by mass spectrometry. Signals for different RNAs are highlighted in different colors.

(B) Quantification of time-dependent extension of RNA in complex A. Electropherogram signals in (A) were integrated and the relative amount of RNA product plotted against incubation time.

(C) Specific uridine monophosphate (UMP) misincorporation opposite the CPD 5'-thymine. Stoichiometric complexes of complete Pol II and scaffold C were incubated with 1 mM of each NTP for 40 min.

(D) Model for UMP misincorporation. In the upper panel, the structure of an undamaged Pol II elongation complex with a non-hydrolysable NTP analogue (PDB 1Y77) was superposed on structure C. Depicted are the base pair at position +1 in 1Y77 (violet), and the CPD in structure C (orange). As modeled in the lower panel, the CPD 5'-thymine could form two hydrogen bonds with an incoming UTP.

(E) Lesion bypass transcription. RNA in complex D (5'-AU-3' opposite the CPD) was not extended after 20 min incubation with 1 mM of NTPs. Bypass was enabled under identical conditions by replacement of the RNA 3'-terminal UMP with adenosine monophosphate (AMP) (5'-AA-3' opposite the CPD, scaffold D^{U→A}, Figure 16).

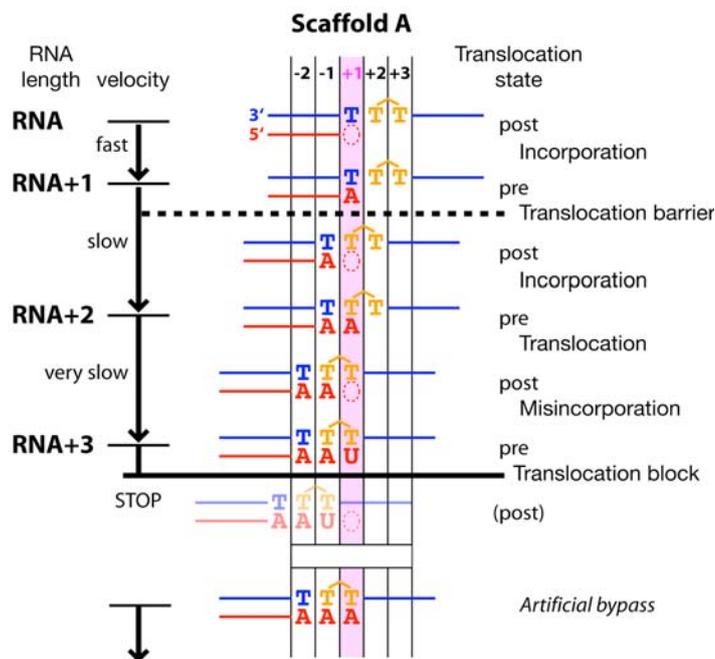


Figure 15: Mechanism of CPD recognition by transcribing Pol II.

Schematic representation of RNA extension in complex A. The initial RNA (top) corresponds to the non-extended RNA of scaffold A. The translocation barrier and the translocation block are indicated with a dashed and a solid horizontal line, respectively. The artificial situation leading to lesion bypass (Figure 14E) is depicted at the bottom.

To test this, we included a CPD at positions +1/+2 of the scaffold, and solved the crystal structure of the resulting complex B (Figure 12A-B). The CPD was observed at register +2/+3, indicating that it is not stably accommodated at positions +1/+2. Pol II had apparently “back-stepped” by one position, consistent with disfavored forward translocation (Figure 15; chapter 3.2.5.5). When complex A was crystallized under conditions with no magnesium ($\text{NH}_4\text{-Na-acetate}$ crystallization solution, chapter 4.7), a strong bromine peak was observed at position -5, indicating a hyper-translocated state with the CPD located at positions +1/+2 (Cheung, A., personal communication). Clear unbiased difference electron density for the CPD was observed above the bridge helix, overlapping with the +1 pre-templating position and the +2 template position in the translocation intermediate (chapter 2), whereas no electron density was found in the +1 templating position. In conclusion, a complex with the 3'-thymine of the CPD in the +1 templating position could not be isolated. Such a state probably does not exist, or at most only transiently. This can be rationalized by assuming that the necessary distortion of the EC is energetically disfavored. Instead, the incorporation of adenine opposite of the 3'-thymine could be a non-templated incorporation event. It was demonstrated that opposite of an abasic site Pol II prefers the incorporation of adenine over other nucleotides (Damsma et al, 2007). For another dinucleotide lesion, the 1,2-d(GpG) intrastrand crosslinks (cisplatin-lesion), preferred addition of adenine opposite the 3'-guanine was reported (chapter 3.2.4) (Damsma et al, 2007). Taken together, it is plausible that the

incorporation of an adenine opposite of the 3'-thymine of a CPD is non-templated. However, it remains to be shown directly.

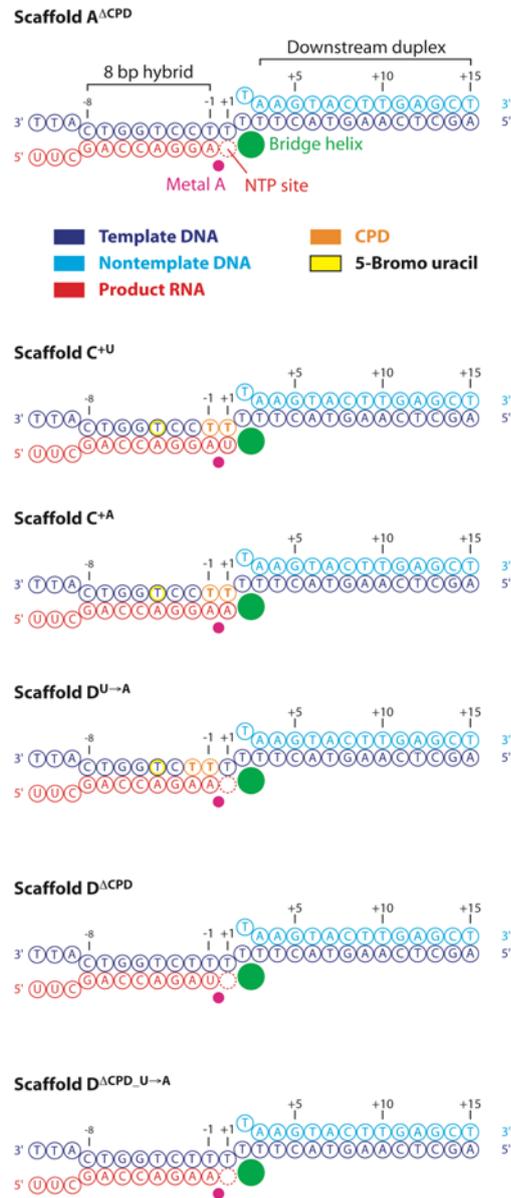


Figure 16: Additional scaffolds used in RNA extension assays.

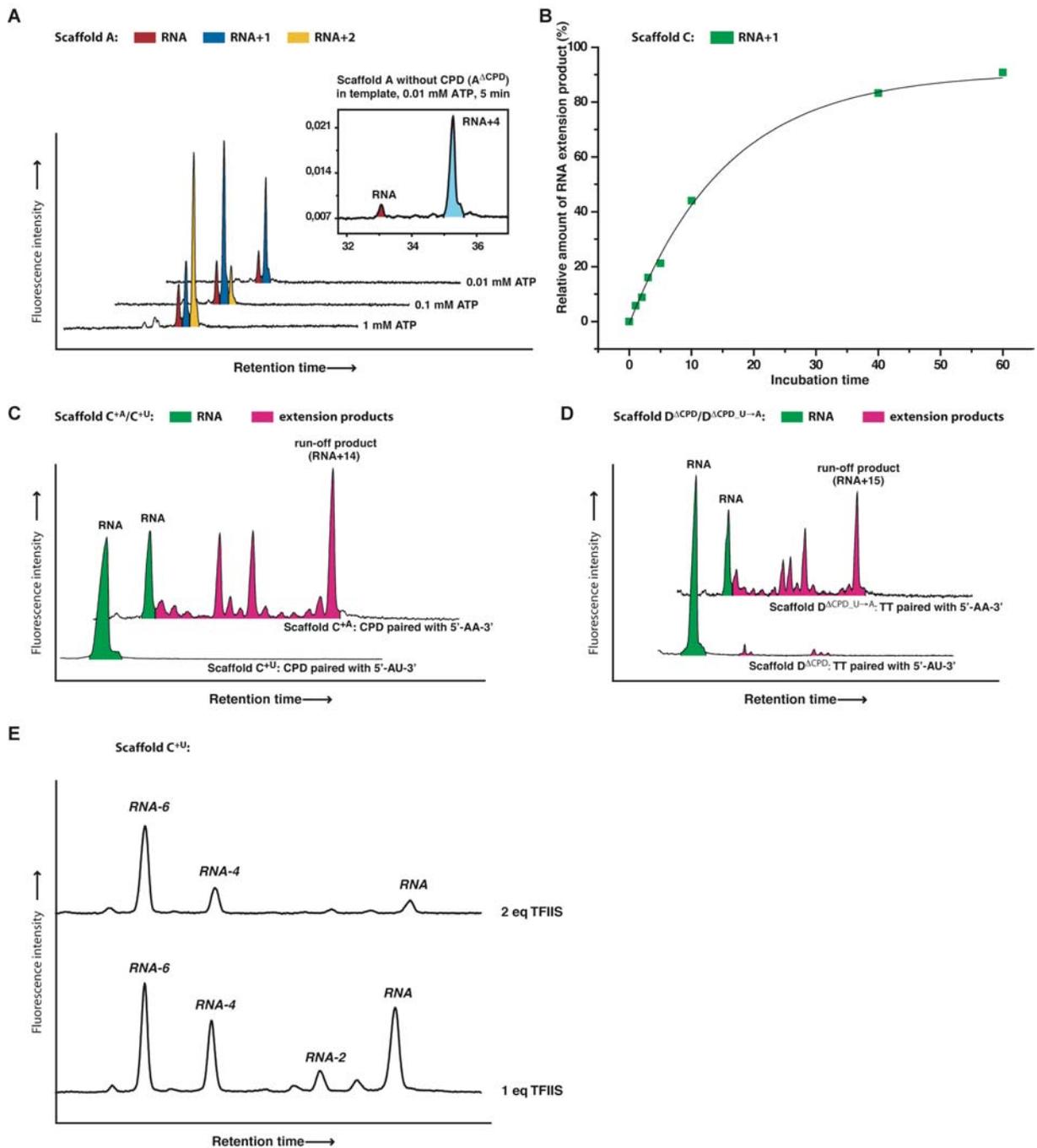


Figure 17: Additional selected RNA extension assays.

(A) RNA extension at varying NTP concentration. Stoichiometric complexes of complete Pol II and scaffold A were incubated for 5 min with different concentrations of ATP. At 1 mM ATP, the translocation barrier is overcome, but only two extension steps are possible since adenine is not incorporated opposite the CPD 5'-thymine (Figure 14A-C). At 0.01 mM ATP, the translocation barrier is not overcome and the RNA is extended by only one nucleotide, whereas an undamaged template with the same sequence (scaffold $A^{\Delta CPD}$, Figure 16) is readily extended at this ATP concentration and the same incubation time, providing a positive control (box).

(B) Kinetics of the final incorporation step. A stoichiometric complex of complete Pol II and scaffold C was incubated with 1 mM of ATP, CTP, GTP and UTP. The relative amount of RNA extension product is plotted against the incubation time (compare Figure 14B). This reaction corresponds to the third incorporation event in complex A (compare Figure 14A-C).

(C) Lesion bypass transcription using an alternative nucleic acid scaffold. To exclude construct bias (mismatch at

position -8 of scaffold D), lesion bypass was also examined using scaffold C with an additional adenine at the RNA 3'-end (5'-AA-3' opposite the CPD, scaffold C^{+A}, Figure 16). Pol II could efficiently bypass the lesion in this scaffold after 20 min incubation with 1 mM of ATP, CTP, GTP and UTP. When scaffold C with an additional UMP was used (5'-AU-3' opposite the CPD, scaffold C^{+U}, Figure 16), no bypass was observed after identical conditions.

(D) Highly inefficient T-U mismatch extension. To examine bypass of a T-U mismatch, the CPD in the template DNA of scaffold D was replaced by two thymines (Scaffold D^{A^{CPD}}, Figure 16). After 20 min incubation with 1 mM of ATP, CTP, GTP and UTP, a small fraction of elongation complexes had bypassed the mismatch (approx. 5 %). Without the mismatch (Scaffold D^{A^{CPD-U→A}}, Figure 16) efficient RNA extension to the run-off product was observed under identical conditions, providing a positive control.

(E) TFIIIS-induced transcript cleavage in dinucleotide increments. Backtracking and transcript cleavage by TFIIIS was analyzed using scaffold C with an additional UMP at the RNA 3'-end (scaffold C^{+U}, Figure 16). The resulting elongation complex corresponds to a complex naturally stalled at a CPD lesion. Incubation time was 10 min, and 1 or 2 molar equivalents of TFIIIS were used. Shortened RNA products are indicated (RNA-2, RNA-4, RNA-6, initial RNA shortened by 1, 2, and 3 dinucleotides, respectively).

After the non-templated adenine addition a subsequent translocation step may allow rotation of the whole CPD into the active site, where it would occupy the template positions -1/+1. To test this and to investigate the very slow incorporation of the third nucleotide into complex A, we included a CPD at positions -1/+1, and solved the structure of the resulting complex C. The CPD was seen stably accommodated in the active site, with a Watson-Crick base pair between the CPD 3'-thymine and the adenine at the 3' end of the product RNA. The NTP-binding site opposite the CPD 5'-thymine was free (Figure 12A-C; Figure 18). We therefore used complex C to monitor incorporation of different NTPs. Only UTP led to nucleotide incorporation opposite the 5'-thymine (Figure 14C), generally consistent with data for human Pol II (Mei Kwei et al, 2004). This misincorporation was very slow, with a rate constant of approx. 2.9 h⁻¹, comparable with the rate determined for the third nucleotide incorporation into complex A (Figure 17B; chapter 3.2.5.4). Since translocation is not required for nucleotide incorporation in complex C, the rate-limiting step in reaching the stalled state is the slow uridine monophosphate (UMP) misincorporation, not CPD translocation from positions +1/+2 to -1/+1. Consistent with misincorporation opposite the CPD 5'-thymine, RNA in complex A was extended by only two residues, instead of three, when only ATP was present (Figure 17A, Figure 14A).

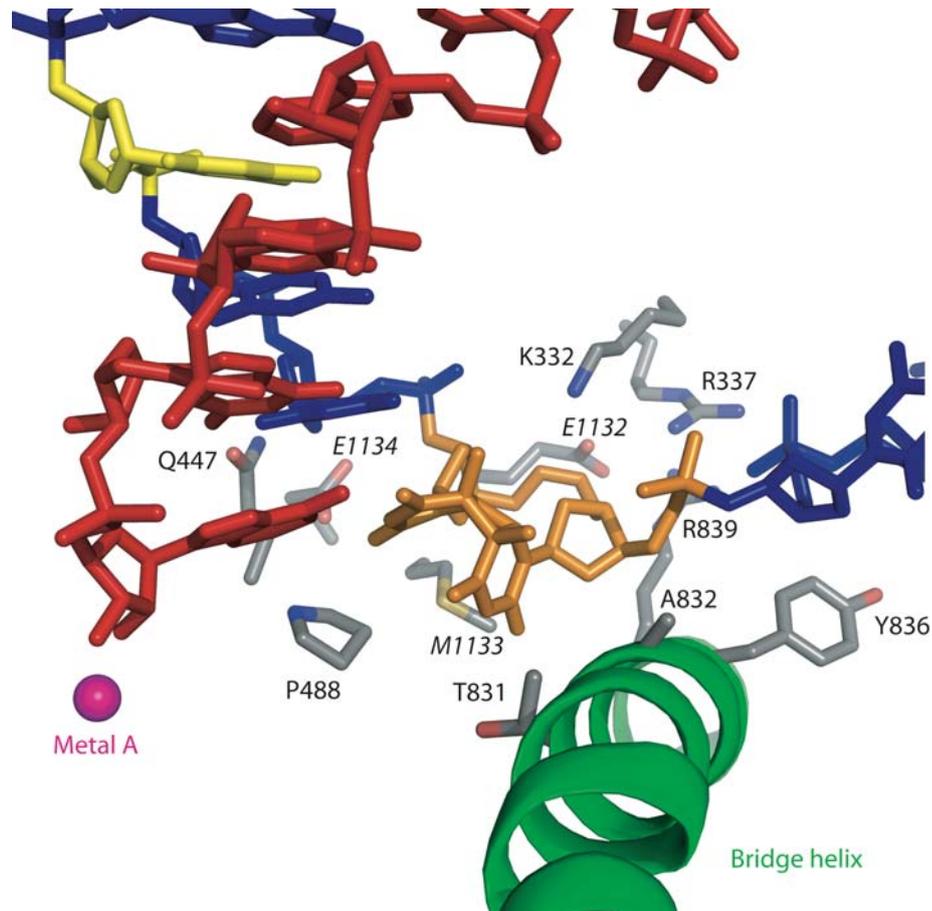


Figure 18: Active site of Pol II in complex C

View as in Figure 12B. Color code as in Figure 12A. Protein side chains within 5 Å of the CPD are shown as grey sticks. Italics are used to distinguish Rpb2 from Rpb1 residues.

Specific UMP misincorporation may be explained with the complex C structure. Whereas the CPD 3'-thymine occupies the same position as in the undamaged elongation complex (Kettenberger et al, 2004), the CPD 5'-thymine is tilted by approx. 40°, and is shifted downwards by more than 2 Å into a wobble position, with the O4 atom at the location normally occupied by the N3 atom (Figure 14D). Provided that binding of the incoming NTP (Kettenberger et al, 2004; Westover et al, 2004a) is unaffected, the wobbled 5'-thymine could form two hydrogen bonds with UTP, whereas only one hydrogen bond would be possible with other NTPs (Figure 14D). Attempts to visualize the CPD 5'-thymine-uracil mismatch crystallographically were unsuccessful (chapter 3.2.5.7).

These results suggested that Pol II stalls because translocation of the CPD 5'-thymine-uracil mismatch from position +1 to position -1 is strongly disfavored. This translocation event would move the damage-containing mismatch into the DNA-RNA hybrid, and the resulting distortion of the hybrid would destabilize the elongation complex (Kireeva et al, 2000) (chapter 3.2.5.8). To test this model, we incorporated the CPD at positions -2/-1 of the scaffold, including a UMP residue opposite the 5'-thymine, and solved the structure of the resulting complex D (Figure 12A-B). In this structure, Pol II had apparently back-stepped by one posi-

tion, and the CPD was again located to positions $-1/+1$ in the active site, consistent with disfavored translocation (chapter 3.2.5.5).

Disfavored translocation of the CPD from position $-1/+1$ to $-2/-1$ may result from distortions due to the CPD and/or due to the mismatch. To distinguish these possibilities, we tested if Pol II extends the RNA in a variant of complex D with a matched CPD 5'-thymine-adenine base pair (Scaffold D^{U→A}, Figure 16). Surprisingly, this RNA was extended to the run-off transcript (Figure 14E; chapter 3.2.5.9). Thus Pol II would bypass a CPD lesion if it could incorporate adenine opposite the CPD 5'-thymine. To test whether a T-U mismatch base pair alone is sufficient to stall Pol II, we used complex D, but without CPD, in RNA extension assays (Scaffold D^{ΔCPD}, Figure 16). Only a small portion of the RNA was extended (Figure 17D). Taken together, Pol II stalling does not result from CPD-induced distortions *per se*, but from CPD-directed misincorporation. In contrast, DNA polymerases can correctly incorporate adenine opposite both CPD thymines, and, dependent on the type of polymerase, this can lead to stalling or lesion bypass (Li et al, 2004; Ling et al, 2003).

3.2.2 A topological model for transcription-coupled DNA repair

In all CPD-containing structures, the polymerase conformation is unchanged. This argues against allosteric models of TCR, which assume an incoming lesion causes a conformational change in Pol II that triggers recruitment of repair factors. In complexes B and D, downstream DNA is repositioned in the polymerase cleft (Figure 12D). However, DNA repositioning cannot support an allosteric mechanism, since it occurs only in back-stepped complexes, which would not form when NTPs are present. A damage-stalled complex could alternatively be detected via exposure of the lesion by Pol II backtracking (Donahue et al, 1994). The transcript cleavage factor TFIIIS induces backtracking of a CPD-stalled complex (Figure 17E; chapter 3.2.5.10), but TFIIIS is not required for TCR *in vivo* (Verhage et al, 1997). The lesion could also be exposed after polymerase bypass or dissociation from DNA. The latter mechanism underlies bacterial TCR, which involves the ATPase Mfd (Deaconescu et al, 2006). However, the related eukaryotic ATPase CSB does neither trigger polymerase dissociation nor bypass (Selby & Sancar, 1997b).

An alternative model for eukaryotic TCR that combines and extends previous models (Sarker et al, 2005; Svejstrup, 2003; Tremeau-Bravard et al, 2004) can explain recognition of the stalled complex without allostery or exposure of the lesion (Figure 19). Complexes that stall at an arrest site are rescued by TFIIIS (Wind & Reines, 2000). Complexes that stall at a non-bulky lesion are rescued by CSB-induced lesion bypass (Lee et al, 2002). In both cases, transcription resumes. At a CPD lesion, however, CSB counteracts TFIIIS-induced backtracking (Laine & Egly, 2006; Selby & Sancar, 1997a), resulting in a stably stalled complex, and opening a time window for assembly of the repair machinery. TFIIH catalyzes extension of the transcription bubble. TFIIH contains two subunits with helicase activity, XPB and XPD.

XPB is required for transcription and repair and is located at the downstream part of the cleft during transcription initiation (Kim et al, 2000; Miller & Hahn, 2006; Tantin, 1998). The observed repositioning of downstream DNA in the CPD-containing structures B and D shows that downstream DNA in the elongation complex can occupy different positions, and a change in the position of downstream DNA may be required for TFIIH action during TCR. There is evidence that the upstream incision site is rendered accessible by TFIIH action (Sarker et al, 2005) and this may be due to the second helicase of TFIIH, XPD. Extension of the transcription bubble permits dual incision of the template strand on the Pol II surface (Selby et al, 1997; Tremereau-Bravard et al, 2004). The lesion-containing DNA fragment and the RNA transcript are removed together with Pol II, although this requires more than dual incision (Tremereau-Bravard et al, 2004). The remaining gapped DNA is repaired. Pol II may be recycled, circumventing its ubiquitination and destruction (Woudstra et al, 2002). In conclusion, our data establish the molecular mechanism of CPD recognition by a cellular RNA polymerase, and provide a structural framework for further analysis of eukaryotic TCR.

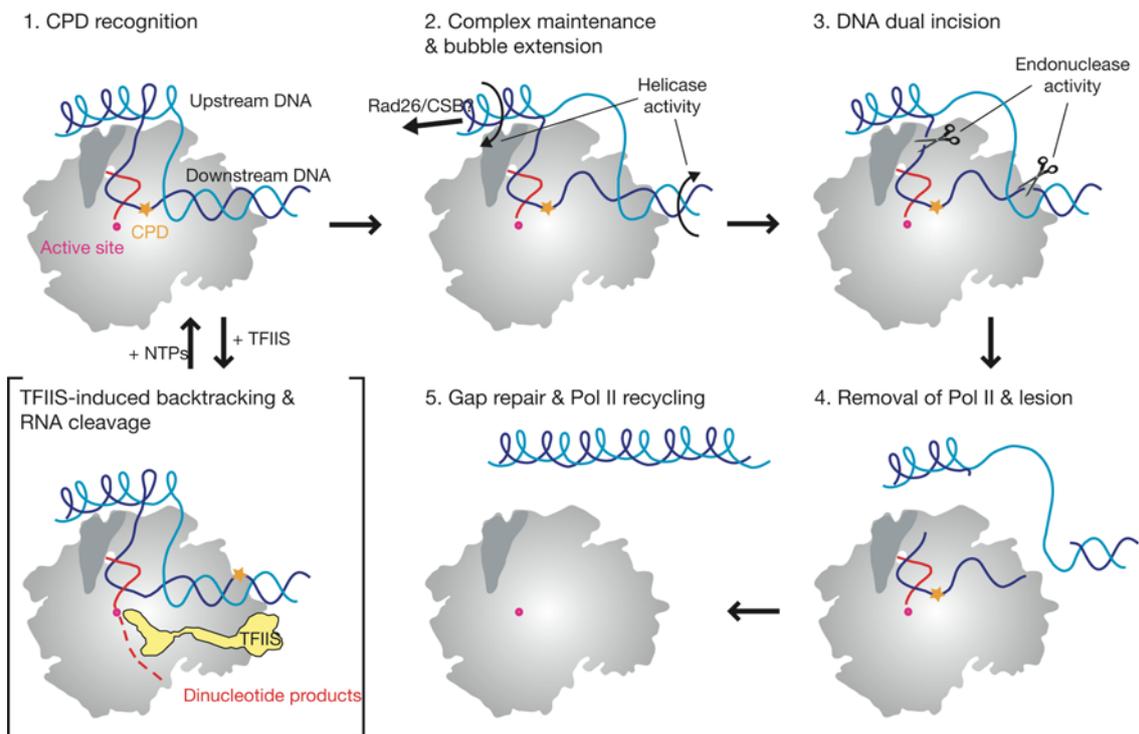


Figure 19: Topological model for transcription-coupled DNA repair.

The lesion-containing Pol II EC is shown schematically. The view and color code are as in Figure 12A-C. An orange star and a magenta sphere indicate the CPD lesion and the active site, respectively. The upstream DNA was placed on Pol II on the basis of its location in the bacterial RNA polymerase-promoter complex (Murakami et al, 2002). The nontemplate strand in the bubble region was modeled. TFIIH helicase subunits XPD and XPB may be involved in extending the transcription bubble in the upstream and downstream direction, respectively, to enable dual incision. Modeling shows that incisions could occur at the two edges of an extended bubble around 10-15 nucleotides upstream (3') and around 15-20 nucleotides downstream (5') of the lesion.

3.2.3 Comparison of CPD recognition by different polymerases

The mechanism of CPD recognition by Pol II is likely the same in Pol I and Pol III, since their active centers are conserved (Jasiak et al, 2006; Kettenberger et al, 2004). However, DNA polymerases recognize and process a CPD differently. The DNA polymerase of bacteriophage T7 and the lesion-bypass polymerase Dpo4 incorporate adenosine monophosphate (AMP) opposite the CPD 5'-thymine (Li et al, 2004; Ling et al, 2003), in contrast to Pol II, which incorporates UMP. In addition, dATP binds differently opposite the CPD 5'-thymine in these two DNA polymerases (Li et al, 2004; Ling et al, 2003). Therefore, detailed structure-function studies of each type of DNA lesion in the context of each polymerase family are required to unravel the mechanisms of lesion recognition and processing.

3.2.4 Comparison with the mechanism of cisplatin damage recognition

The anticancer drug cisplatin (cis-diamminedichloroplatinum(II)) forms 1,2-d(GpG) DNA intrastrand cross-links (cisplatin lesions) that stall Pol II and trigger transcription-coupled DNA repair (Corda et al, 1991; Corda et al, 1993; Jung & Lippard, 2006; Kartalou & Essigmann, 2001; Tornaletti et al, 2003; Wang & Lippard, 2005). Whereas in the CPD lesion two neighboring thymine bases are covalently linked with a cyclobutane ring including the C5 and C6 atoms, in a cisplatin lesion the platinum atom coordinates the N7 atoms of two adjacent guanines in a DNA strand (Figure 20). The cisplatin lesion can be stably accommodated in a Pol II EC at position +2/+3 of the template strand, but translocation to position +1/+2 is disfavored (Figure 21)(Damsma et al, 2007), which is both also the case for the CPD lesion. There is strong evidence that adenine is incorporated in a non-templated fashion opposite of the cisplatin 3'-guanine, as it is also proposed for the CPD lesion (chapter 3.2.1). However, unlike the CPD, the cisplatin lesion cannot be stably accommodated in the active site (positions -1/+1). There are two possible causes. First, the cisplatin lesion is a more bulky dinucleotide lesion than the CPD. The maximum lateral dimension is 7.2 Å (N2-N2 distance) compared to 5.3 Å (O2-O2 distance) for the CPD (Figure 20). Modeling suggested, that a conformational change of the bridge helix would be required to accommodate the lesion in the active site (Figure 22). Second, a G•A mismatch base pair would be formed at position -1 in contrast to a stabilizing T-A base pair in the case of the CPD.

In conclusion, the mechanism of recognition by transcribing Pol II is different for both dinucleotide lesions (Figure 23). At a cisplatin lesion, Pol II stalls because the lesion cannot be delivered to the active site, whereas it stalls at a CPD lesion after delivery to the active site and specific UMP misincorporation opposite the 5'-thymine. Bypass of the CPD lesion is only possible by artificially replacing the resulting T•U mismatch by a T-A match. Bypass of the cisplatin lesion is possible by artificially providing a starting transcript, extending at least up to the 3'-guanine (Figure 23). Remarkably, bypass is possible in presence of a G•A mismatch with the 3'-guanine of the cisplatin dimer.

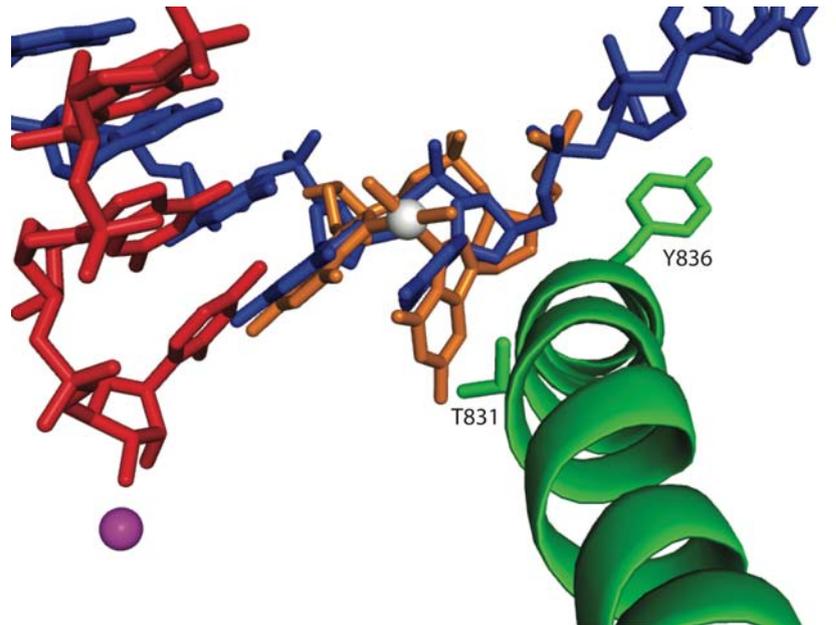


Figure 22: Model of the cisplatin lesion in the active site of Pol II

Modeling of the cisplatin lesion at positions +1/-1. A cisplatin lesion was manually placed such that it superimposes with template nucleotides at the active center positions +1 and -1 of complex A. Owing to some close contacts, in particular with Thr831 in the bridge helix, slight conformational changes are required to accommodate the lesion.

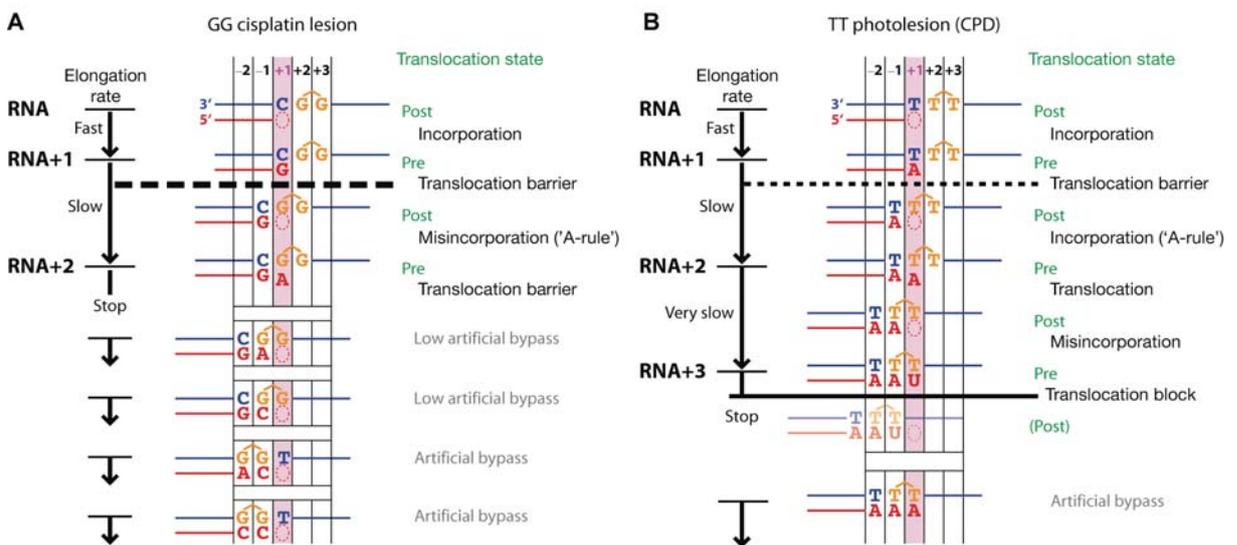


Figure 23: Different mechanisms of Pol II stalling at dinucleotide lesions.

(A) Pol II stalling at the cisplatin lesion, shown as a schematic representation of RNA extension in complex A. The initial RNA (top) corresponds to the non-extended RNA of scaffold A. Dashed line represents translocation barrier. The artificial conditions leading to lesion bypass are depicted at the bottom.

(B) Pol II stalling at a CPD lesion (compare Figure 15, page 38).

3.2.5 Additional detailed information

3.2.5.1 Minimal nucleic acid scaffolds and hybrid length in elongation complexes.

The structure of the complete Pol II elongation complex showed seven DNA-RNA base pairs at positions -1 to -7 , with Watson-Crick edges of the bases at position -8 at the upstream end of the hybrid about 4 Å apart (Kettenberger et al, 2004). This slightly increased base distance may have resulted from a bulky A-A DNA-RNA mismatch at position -9 . The CPD lesion containing Pol II EC structures A and C contain a smaller A-C mismatch at position -9 , and show a hybrid length of eight base pairs (positions -1 to -8). To test if the scaffold design influences the outcome of the biochemical results, we tested complete scaffolds that included a nontemplate strand in biochemical assays. The results were essentially not influenced by the nontemplate strand, showing that the use of minimal scaffolds in the functional studies is justified. The use of the minimal scaffolds in structural studies is justified by the highly similar nucleic acid structure observed in the complete elongation complex structure (Kettenberger et al, 2004).

3.2.5.2 Use of a CPD analogue.

A thymine-thymine CPD formacetal analogue was used that carries a methylene group instead of a phosphate group between the two thymines (Butenandt et al, 2000) (chapter 3.3.1). As a consequence there is a missing negative charge in the DNA template strand backbone between positions $+2$ and $+3$, which may be partially responsible for downstream DNA mobility in complex A. The analogue was used because it can be synthesized in large quantities, as required for crystallographic studies. All current evidence argues that the CPD analogue is structurally and functionally equivalent to the natural CPD. First, high-resolution small molecule crystallographic data of the CPD formacetal analogue show that it adopts the same structure and conformation than the normal CPD with a phosphate backbone (Butenandt et al, 2000; Cadet et al, 1985). Second, the base pairing properties of the lesion and the lesion analogue, as judged from melting point studies, are the same (Butenandt, 1985; Taylor et al, 1990). Third, the formacetal linkage is a nearly exact bioisostere of the natural phosphate (Gao et al, 1992; Veal et al, 1993). Fourth, the small differences, which are in the range of only around 0.1 Å are without significance for our results and conclusions. The change to the formacetal does neither affect the nucleobases nor the sugar units, the only difference is the missing charge of the backbone. Finally, the CPD analogue is well-accepted as an analogue for recognition by a photolyase (Mees et al, 2004).

3.2.5.3 Structural similarity of lesion-containing and lesion-free elongation complexes.

The RMS deviation of C α atoms between the lesion-free complete Pol II elongation complex (PDB code 1Y1W) and the CPD-containing complexes is 0.4 Å in all four cases. The RMS deviation of C α atoms between a lesion-free core Pol II elongation complex (PDB code 1R9S) and the CPD-containing complexes is 0.9 Å in all four cases.

3.2.5.4 Rate constants and kinetic measurements.

The rate constant for normal elongation *in vitro* is greater than 20 s⁻¹ (Edwards et al, 1991), too fast to be monitored by the kinetic methods used here. This is why the first incorporation into RNA of complex A appears to be instantaneous.

3.2.5.5 Detection of apparently “back-stepped” complexes.

The observed "back-stepping" in complexes B and D leads to the pre-translocated state, and occurs although it results in an A-C mismatch at position -8 at the upstream end of the hybrid. When the anomalous difference Fourier maps in complexes B and D (Figure 13C) were contoured at the low level of 2.5 σ , there were no peaks at potential bromine positions in neighboring nucleotides. Thus there were no significantly populated alternative states with different nucleic acid registers, such as non-back-stepped states. In all pre-translocated complexes, electron density for the 3'-terminal RNA nucleotide in position +1 was fragmented or absent, presumably since the phosphodiester bond between positions -1 and +1 was cleaved by the intrinsic 3'-exonuclease activity of Pol II under crystallization conditions (initially 50 mM magnesium ions, 10-20 days incubation time at 22°C).

3.2.5.6 Pol II binds the hybrid in a defined register.

In the "back-stepped" structures B and D, the bromine peak in the template strand is observed precisely at register -3, instead of -4 (Figure 12A-B, Figure 13C). This confirms that Pol II binds the hybrid in a defined, one base pair register, consistent with movement of Pol II in discrete one base pair steps (chapter 2.2.1) (Abbondanzieri et al, 2005).

3.2.5.7 Attempts to visualize the CPD 5'-thymine-uracil mismatch crystallographically.

We included a UMP at the RNA 3'-end of complex C (Scaffold C^{+U}, Figure 16), and re-solved the structure, but the resulting electron density for the UMP was fragmented (not shown). Soaking UTP into complex C crystals also resulted in poor density for the UMP, and did not result in translocation, as seen from an unaltered location of the bromine label in the DNA template (not shown). The modeling of the CPD thymine-uracil wobble base pair must be taken with care because of the limited resolution of the new data and of previous data with

bound nucleotide substrates. In structure D the uracil is present at position +1 but could not be modeled due to fragmented electron density (see chapter 3.2.5.5).

3.2.5.8 A DNA-RNA hybrid is destabilized by a mismatch at the CPD.

Consistent with the model that a CPD 5'-thymine-uracil mismatch destabilizes the DNA-RNA hybrid, the melting point of a 16 base-pair hybrid duplex with a CPD 5'-thymine-uracil mismatch is substantially lower than that of the same duplex with a 5'-thymine-adenine base pair (50.1°C versus 60.0°C). The complete melting point data are as follows. Undamaged DNA-RNA heteroduplex (5'-TT-3'): 3'-AA-5', 63.8°C; 3'-AU-5', 53.5°C; 3'-UA-5', 54.0°C. Damaged DNA-RNA heteroduplex (5'-CPD-3'): 3'-AA-5', 60.0°C, 3'-AU-5', 47.5°C; 3'-UA-5', 50.1°C. Thus a CPD alone only leads to a minor destabilization of the DNA-RNA duplex, consistent with earlier data (Butenandt, 1999). However, a T-U mismatch leads to substantial destabilization.

3.2.5.9 Additional support of lesion-bypass transcription.

When an additional AMP is included at the RNA 3'-end of scaffold C (scaffold C^A, Figure 16), the RNA was readily extended to the run-off transcript (Figure 17C). As a control, scaffold C with an additional UMP at the RNA 3'-end was used (Scaffold C^U, Figure 16) and no extension of the RNA could be observed (Figure 17C).

3.2.5.10 TFIIS induces backtracking of the CPD-stalled complex and RNA cleavage in discrete dinucleotide steps.

When complex C with an additional UMP at the RNA 3'-end (scaffold C^U, Figure 16) was incubated with TFIIS, repeated dinucleotide cleavage from the RNA 3'-end was observed (Figure 17E). This indicates Pol II backtracking and movement of the CPD lesion out of the active site into the downstream direction.

3.3 Specific experimental procedures

3.3.1 Preparation of Pol II-nucleic acid complexes.

Endogenous *S. cerevisiae* ten-subunit Pol II core enzyme and recombinant Rpb4/7 heterodimer were purified as described in chapter 4.2 and 4.4. The thymine-thymine CPD formacetal analogue-containing DNA oligonucleotides were synthesized as described (Butenandt et al, 1998). 5'-bromouracil phosphoamidite was purchased from Glen Research and incorporated as described by the manufacturer. Nucleic acid scaffolds (Figure 12A; Figure 16) were annealed by mixing equimolar amounts of synthetic template DNA, nontem-

plate DNA, and RNA in TE buffer (chapter 4.6) at a final concentration of 100 μ M, heating the mixture to 95°C for 2 min, and slow-cooling to room temperature. Stoichiometric transcribing Pol II-nucleic acid complexes were assembled by incubating core Pol II for 10 min with 1.5 molar equivalents of nucleic acid scaffold, followed by 20 min incubation with 5 molar equivalents of recombinant Rpb4/7 in assembly buffer (chapter 4.6) at 20°C. The complexes were purified by size exclusion chromatography (Superose 6 10/300 GL) in Pol II buffer (chapter 4.6).

3.3.2 Crystallization and crystal treatment.

Purified Pol II-nucleic acid complexes were concentrated to 3.5-4.5 mg/ml and an additional amount of the nucleic acid scaffold was added prior to crystallization to a final concentration of 2 μ M. Crystals were grown at 22 °C with the hanging drop vapor diffusion method by mixing 2 μ l of sample solution with 1 μ l of reservoir solution (NH₄-Mg-acetate crystallization solution, chapter 4.7). Crystals were harvested in mother solution after 10-20 days of growth, when they had reached their maximum size (0.4 x 0.3 x 0.2 mm). Subsequently the crystals were transferred stepwise to mother solution containing additionally 0-20 % glycerol over 5 h, slowly cooled to 8°C, and flash-cooled by plunging into liquid nitrogen.

3.3.3 X-ray structural analysis.

Diffraction data were collected in 0.25° increments at the protein crystallography beamline X06SA of the Swiss Light Source (Table 5). Raw data were processed with Denzo and Scalepak (Otwinowski & Minor, 1996). Structures were solved by molecular replacement with the program Phaser (McCoy et al, 2005), using the structure of the complete 12-subunit Pol II elongation complex without nucleic acids as a search model (PDB 1Y1W (Kettenberger et al, 2004)). The molecular replacement solution was subjected to rigid body refinement with CNS (Brunger et al, 1998). Model building was done using the program O (Jones et al, 1991). The nucleic acids were built into the initial F_o-F_c electron density maps (Figure 13A). The register of the nucleic acids was unambiguously defined by bromine labeling. A thymine residue in the template strand was replaced for 5-bromouracil, diffraction data of the resulting complex were recorded at the wavelength of the bromine K absorption edge, and the resulting anomalous difference Fourier maps revealed single peaks demarking the positions of the bromine atom (Figure 13C). After minor adjustments to the protein model, atomic positions and B-factors were refined with CNS. Refinement was monitored carefully with the free R factor, calculated from the same set of excluded reflections as in refinement of complete Pol II (Armache et al, 2005) and the complete Pol II elongation complex (Kettenberger et al, 2004).

3.3.4 Transcript extension assays.

Stoichiometric transcribing complexes of complete Pol II, containing the 10-subunit core and Rpb4/7, were assembled and purified as described above, but in transcription buffer (chapter 4.6). Full-length yeast TFIIIS was purified as described in chapter 4.5. Transcribing complexes of ten-subunit core Pol II were assembled by incubating core Pol II for 20 min with 1.5-2 molar equivalents of the respective nucleic acid scaffold in transcription buffer at 20°C. Core Pol II elongation complexes were used for measurements presented in Figure 14E and Figure 17C-D. In these assays, core Pol II is generally indistinguishable from the complete Pol II. For all other measurements, stoichiometric elongation complexes of complete Pol II were used. The RNAs used for extension assays were identical to the ones used for structural studies, except five additional nucleotides (5'-UGCAU-3') and a fluoresceine label at their 5'-end. For the transcript extension and cleavage reactions, 0.2-0.4 µM of Pol II-nucleic acid complex were incubated with NTPs or TFIIIS, respectively, at 28°C in transcription buffer. The reaction was stopped by adding 50 mM EDTA pH 7.4. Complexes were disassembled and protein was denatured by gradually heating to 95 °C in 12 min and then incubating at 95 °C for 4 min. After removal of protein by centrifugation, the nucleic acid samples were desalted using C18 ZipTips (Millipore). Fluoresceine-labeled RNA extension products were analyzed using a Beckmann Coulter PACE-MDQ DNA system with laser-induced fluorescence detection (excitation and detection wavelengths were 488 nm and 520 nm, respectively). Separations were carried out using fused silica capillaries filled with polyacrylamide gel (Beckmann Coulter) and Tris borate buffer (0.1 M Tris borate, 1 mM EDTA, 7 M urea, pH 8.4). Samples were injected by a voltage pulse (10 kV for 10 s) and separated at 8 kV for 50 min. The nucleic acids of the transcript extension reactions were additionally analyzed by mass spectrometry using a Bruker Daltonics autoflex 2 MALDI-TOF spectrometer. This confirmed the identity of the RNA extension products.

4 Common experimental procedures

4.1 Measurement of protein concentration

For the determination of protein concentrations the Bradford protein assay was used (Bradford, 1976). Dye reagent was purchased from Biorad and the assay was performed according to the manufacturer's instructions. For each new batch of dye reagent a calibration curve was generated using Bovine serum albumin (Fraktion V, Roth).

4.2 Isolation of 10-subunit core Pol II from yeast

4.2.1 Yeast fermentation

Core RNA polymerase II was isolated from the *Saccharomyces cerevisiae* strain CB010ΔRpb4 (MATa *pep4::HIS3/prb1::LEU2, prc1::HISG, can1, ade2, trp1*) (Edwards et al, 1990; Fu et al, 1999). This strain carries knockouts of several cellular proteases and of RPB4. In the absence of Rpb4, Rpb7 dissociates from core Pol II during purification, giving rise to homogeneous 10-subunit core Pol II. Two types of fermenters were available for producing up to 1.5 kg of yeast pellet per batch. The small fermenter (ISF200, Infors) has a nominal volume of 20 l and is ideally run with up to 15 l of media. The large fermenter (ABEC, Infors) has a nominal volume of 200 l and can be filled with up to 160 l of media. Table 6 shows the media composition and the culture parameters of both fermenters:

Table 6: Conditions for the fermentation of yeast

	small fermenter (20 l)	large fermenter (200 l)
YPD media	300 g peptone	3200 g peptone
	300 g glucose	3200 g glucose
	222 g yeast extract	2370 g yeast extract
	15 l desalted water	160 l desalted water
Antibiotics ¹	0.75 g ampicillin	8.0 g ampicillin
	0.15 g tetracycline · HCl	1.6 g tetracycline · HCl
typical inoculate volume	0.3 l @ OD600 ≈ 2	4-5 l @ OD600 ≈ 2
air flow	8 l/min	20 l/min
stirrer speed	800 rpm	200 rpm
typical growth time	12 – 15 hours	12 – 15 hours

¹added after sterilization, prior to inoculation

4.2.2 Purification of 10-subunit core RNA Pol II

4.2.2.1 Buffers

100 x protease inhibitor mix (p.i.)

1.42 mg Leupeptin

6.85 mg Pepstatin A

850 mg PMSF

1650 mg benzamidine

dry ethanol to 50 ml

stored at -20°C ; added immediately before use

3x freezing buffer

150 mM Tris-Cl, pH 7.9 @ 4°C

3 mM EDTA

30 % glycerol

30 μM ZnCl_2

3 % DMSO

30 mM DTT

3 x protease inhibitor mix

1 x HSB150 buffer

50 mM Tris-Cl, pH 7.9 @ 4°C

150 mM KCl

1 mM EDTA

10 % glycerol

10 μM ZnCl_2

10 mM DTT

1 x protease inhibitor mix

1 x HSB600 buffer

50 mM Tris-Cl, pH 7.9 @ 4 °C

600 mM KCl

1 mM EDTA

10 % glycerol

10 μ M ZnCl₂

10 mM DTT

1 x protease inhibitor mix

TEZ buffer

50 mM Tris-Cl, pH 7.5 @ 20 °C

1 mM EDTA

10 μ M ZnCl₂

1 mM DTT

1 x protease inhibitor mix

UnoQ buffer

50 mM Tris-Cl, pH 7.5 @ 20 °C

1 mM EDTA

10 μ M ZnCl₂

10 % (v/v) glycerol

10 mM DTT

no protease inhibitors

1 x Pol II buffer

5 mM Hepes pH 7.25 @ 20 °C

40 mM ammonium sulfate

10 μ M ZnCl₂

10 mM DTT

Acetate buffer

100 mM sodium acetate pH 4.0

500 mM sodium chloride

PBS

4.3 mM Na₂HPO₄

1.4 mM KH₂PO₄

137 mM sodium chloride

2.7 mM potassium chloride

pH 7.4

Coupling buffer

100 mM sodium bicarbonate pH 8.3

500 mM sodium chloride

4.2.2.2 Harvesting and storage of yeast

Yeast was harvested at late logarithmic/early stationary phase, monitored by OD600 measurement. Cells were pelleted by centrifugation for 20 min. at 5000 rpm in a SLC6000 rotor (small fermenter) or by a continuous flow centrifuge (Padberg Z41G, 20000 rpm) in case of the large fermenter. The cell pellet was resuspended in 330 ml of 3x freezing buffer per kg cells and stirred at 4 °C for 30 min, before shock-freezing in liquid nitrogen and stored at -80 °C.

4.2.2.3 Purification - day 1 (lysis and heparin column)

For three bead-beaters (BioSpec), up to 600 ml of cell suspension were thawed in warm water. Each bead-beater was filled with 200 ml of borosilicate glass beads (0.45-0.50 mm diameter), 1 ml of protease inhibitor mix and 200 ml of the cell suspension. HSB150 was added to fill the bead-beater completely, taking care to avoid any remaining air bubbles. Lysis was achieved within 60-75 min of bead-beating (30 s on/90 s off) while the beater chambers were submersed in a salt/ice mixture.

Glass beads were removed by filtration through a mesh funnel. The beads were washed with HSB150 until the flowthrough was clear. The lysate was cleared by two rounds of centrifugation (45 min at 9000 rpm in a GS3 rotor or 30 min at 12000 rpm in a SLA1500 rotor). Lipids were then removed by filtration of the supernatant through two layers of paper filter discs underneath a dressing cloth.

The cleared lysate was applied onto a column packed with 250 ml of Heparin Sepharose 6 FF (GE Healthcare) (flow rate: 6-8 ml/min), pre-equilibrated with 750 ml of HSB150. Elution was accomplished with 500 ml of HSB600 (flow rate: 6-8 ml/min). Proteins in the eluate were precipitated by adding 291 g of fine-ground ammonium sulfate per litre of eluate (= 50 % saturation), followed by 60 min of stirring at 4 °C, over-night incubation at 4 °C and finally centrifugation (45 min. at 12000 rpm in a SLA1500 rotor). The heparin column was restored

by washing with 1 l of 6 M urea, followed by water, and stored in 5 mM potassium acetate and 20 % (v/v) ethanol. Every five runs, the heparin column was regenerated by a brief wash with 500 ml of 0.1 M NaOH, followed by water and 5 mM potassium acetate in 20 % (v/v) ethanol.

4.2.2.4 Purification - day 2 (immunoaffinity column)

The ammonium sulfate pellet of day 1 was dissolved in 50 ml of buffer TEZ. More TEZ was added to adjust the conductivity below the conductivity of TEZ containing additionally 400 mM ammonium sulfate (TEZ400). This sample was centrifuged (15 min at 14000 rpm in a SLA1500 rotor) to remove undissolved constituents and then loaded by gravity flow onto the immunoaffinity column at 4 °C (see chapter 4.2.2.6, flow rate: 0.5-1.0 ml/min). The column was pre-equilibrated with 20 ml of TEZ containing 250 mM ammonium sulfate (TEZ250). The flowthrough was loaded onto a second column to increase the yield of Pol II. The columns were brought to room temperature, washed with 25 ml of TEZ500 at room temperature and Pol II was eluted in 1 ml fractions with TEZ500 containing additionally 50 % (v/v) glycerol (ca. 15 ml). Directly afterwards, 9 mM DTT was added to the elution fractions containing Pol II (monitored with the Bradford assay), and they were stored at 4 °C over night. The columns were washed with 5 ml of TEZ500 containing 70 % (v/v) ethylene glycol but no DTT, and re-equilibrated with 25 ml of TEZ250 containing 0.02 % sodium azide. Generally, the recovery of Pol II decreased with each use of the column starting already from the first use. One reason is probably the sensitivity of the antibody towards DTT. However, it was reported before that these kind of columns could be used > 10 times before recovery drops (Kettenberger, 2005).

4.2.2.5 Purification - day 3 (anion exchange chromatography or buffer exchange)

Peak fractions from day 2 were diluted six-fold and loaded onto a UnoQ column (BioRad, column volume 1.35 ml), pre-equilibrated with buffer UnoQ containing 60 mM ammonium sulfate (UnoQ-A). The column was washed with 3 column volumes of this buffer, and Pol II was eluted with a linear gradient over 10 column volumes from 0-50 % buffer UnoQ containing 1 M ammonium sulfate (UnoQ-B). Pol II elutes at about 25 % buffer UnoQ-B. The column was restored by washing with 5 column volumes of UnoQ-B.

To increase the final yield of Pol II the anion exchange step was replaced by a simple buffer exchange procedure, which didn't affect the suitability of the purified Pol II for structural or functional experiments. The buffer in the antibody column elution fractions was exchanged for 1x Pol II buffer using centrifugal ultrafiltration devices (MWCO 100,000 Da, Millipore Amicon Ultra-15). Completeness of the buffer exchange was monitored by measuring the conductivity of the flowthrough. Finally, Pol II was concentrated to 1-2 mg/ml.

The Pol II sample after anion exchange chromatography or after buffer exchange was split into aliquots of 100-500 µg Pol II. The aliquots after anion exchange chromatography were mixed with an equal volume of ammonium sulfate solution saturated at room-temperature, the aliquots in 1 x Pol II buffer with 1.13 times the volume. The mixture was incubated for > 1 hour at 4 °C and centrifuged for 45 min. at 4 °C in a table-top centrifuge at 13000 rpm. Most of the supernatant was decanted so that the pellet was still covered with supernatant, before it was shock-frozen in liquid nitrogen and stored at -80 °C. Pol II stored this way is stable for about 3 months. From 600 g yeast pellet, a yield of 0.5-4 mg of highly purified Pol II was achieved, although higher yields up to 8 mg have been described (Kettenberger, 2005).

4.2.2.6 Preparation of Pol II immunoaffinity resin

The monoclonal antibody 8WG16 (NeoClone, Madison/USA), described in (Thompson & Burgess, 1996) is specific for the unphosphorylated CTD of Pol II and optimized to release Pol II upon treatment with glycerol or ethylene glycol at room temperature ("polyol responsive antibody"). The antibodies were purified from mouse ascites and immobilized on activated chromatography media according to the following procedure:

Lyophilized ascites was dissolved in PBS to its original volume and filtered through 0.2 µm membrane filters. The solution was passed > 3 times through a protein-A sepharose column (5 ml column volume, Sigma), pre-equilibrated in PBS. The column was washed with 50 ml PBS and antibodies were eluted with 20 ml of 0.75 M acetic acid. Fractions of 1 ml were collected into tubes containing 200 µl of 2 M HEPES (pH 7.9) to neutralize the acid. Peak fractions were pooled and the protein-A sepharose column was regenerated by washing for 5 min. with 1 M acetic acid, followed by PBS with 0.02 % sodium azide.

The matrix for immunoaffinity columns was cyanogen bromide (CNBr)-activated sepharose 4 B (Sigma), which reacts with free amines, e.g. accessible -NH₂ groups on proteins. Care was taken to avoid other sources of free amines (e.g. Tris) and to use a sealed bottle of activated sepharose. For each immunoaffinity column, 5 ml of gel was prepared by suspending 1.43 g of CNBr-sepharose in several ml of 1 mM HCl in a disposable gravity-flow column. The suspended CNBr-sepharose was first washed with 100 ml of 1 mM HCl, then with 20 ml of coupling buffer. Coupling was performed with 10 mg of purified antibodies per column for 2 hours at 20 °C or over night at 4 °C. When the coupling reaction was completed, no protein was detectable in the supernatant. The column was then washed with 25 ml of 1 M Tris, pH 8 and incubated for 2 hours at room temperature or over night at 4 °C. Finally, the column was washed with 20 ml of coupling buffer, followed by acetate buffer and coupling buffer. Columns were stored at 4 °C in TEZ60 with 0.02 % sodium azide. The columns could be used several times, if DTT exposure was reduced to a minimum, but in general a decrease in Pol II yield was observed already after the first use with further decrease after subsequent uses.

4.3 Purification of His-tagged RNA polymerase II

Due to the low and unreproducible yield of Pol II and the high costs and fast decay of the antibody column an alternative purification procedure was established in the lab (Kireeva et al, 2003) (Cheung, A., Damsma, G.E., Lehmann, E., Sydow., J., personal communication). This procedure involves purification of Pol II from a yeast strain with a hexahistidine tag at the N-terminus of Rpb3, which is exploited in a Ni-affinity chromatography step. For the experiments described in this thesis, the protocol described in chapter 4.2.2 still had been used.

4.4 Purification of the subcomplex Rpb4/7

Buffer 1

150 mM NaCl

5 % (v/v) glycerol

50 mM Tris pH 7.5

10 mM β -mercaptoethanol

protease inhibitors (see chapter 4.5)

Buffer 2

50 mM Tris pH 7.5

5 mM DTT

1 mM EDTA

Recombinant yeast Rpb4/7 was expressed in *E. coli* BL21(DE3) RIL (Stratagene) using a bicistronic vector (Armache, 2005; Sakurai et al, 1999). Cells were grown in 2 x 2 L of culture in auto-induction medium (Studier, 2005). After approx. 4 h, when OD₆₀₀ \approx 0.6 was reached, the temperature was shifted from 30 °C to 20 °C. After 11 h, the cells were harvested by centrifugation (15 min at 5000 rpm in a SLC6000 rotor), resuspended in buffer 1 and lysed by means of a French Press. The lysate was cleared by centrifugation (30 min at 15000 rpm in a SS34 rotor) and applied onto a NiNTA column (Quiagen; 1 ml column volume). The column was washed subsequently with 3 ml of buffer 1, 3 ml buffer 1 containing additionally 10 mM imidazole and 3 ml of buffer 1 containing additionally 20 mM imidazole. Elution was performed subsequently with 3 ml of buffer 1 containing additionally 50 mM imidazole and 3 ml of buffer 1 containing additionally 200 mM imidazole. Peak fractions were pooled, diluted 1:3 with buffer 2 and applied on a ResourceQ column (GE Healthcare, 6 ml column volume), pre-equilibrated in buffer 2. Rpb4/7 was eluted with a linear gradient from 0-1000 mM NaCl in buffer 2. Peak fractions were concentrated and applied on a Superose12 10/300 GL gel

filtration column (GE Healthcare), pre-equilibrated in Pol II buffer (see chapter 4.2.2.1). The purified Rpb4/7 heterodimer was concentrated to 10 mg/ml and aliquots were stored at -80°C .

4.5 Purification of elongation factor TFIIS

Buffer A

50 mM Hepes pH 7.5

300 mM NaCl

5% glycerol

10 μM ZnCl₂

1 x protease inhibitors

10 mM β -mercaptoethanol

10 x protease inhibitors

3 mg/L leupeptin

14 mg/L pepstatin A

1.7 g/L PMSF

3.3 g/L benzamidine

dissolved in ethanol

For the expression of full-length TFIIS, transformed BL21(DE3) RIL (Stratagene) cells were grown in LB medium, supplemented with chloramphenicol (30 mg/l) and kanamycin (30 mg/l), to an OD₆₀₀ of 0.7 and cooled down to 20 $^{\circ}\text{C}$ before expression was induced with 1 mM IPTG, and continued over night.

Cells from 1 l of culture were lysed by sonication in 50 ml of buffer A. The lysate was cleared by centrifugation (45 min, 13000 rpm, SS34 rotor) and was applied to a His Trap HP column (1 ml; GE Healthcare). The column was washed with buffer A containing 500 mM NaCl, and the protein was eluted over a total of 20 ml with a gradient of 0 mM to 500 mM imidazole in buffer A containing 500 mM NaCl. Peak fractions were diluted 5-fold in buffer A and loaded onto a Mono-S 5/5HR anion exchange column (GE Healthcare), pre-equilibrated with buffer A containing 100 mM NaCl. The TFIIS variant was eluted over a total of 15 column volumes with a gradient of 100-500 mM NaCl in buffer A. Peak fractions were pooled and subsequently applied onto a Superose 12 10/300GL gel filtration column (GE Healthcare), previously equilibrated in 1 x Transcription buffer (chapter 4.6). Elution fractions (2 mg/ml protein) were shock-frozen in liquid nitrogen and stored at -80°C .

4.6 Buffers for Pol II and Pol II ECs

TE buffer

10 mM Tris pH 7.4

1 mM EDTA

Assembly buffer

50 mM Hepes pH 7.5

40 mM $(\text{NH}_4)_2\text{SO}_4$

10 μM ZnCl_2

5 % (v/v) glycerol

10 mM DTT

Pol II buffer

5 mM Hepes pH 7.25 at 20 °C

40 mM $(\text{NH}_4)_2\text{SO}_4$

10 μM ZnCl_2

10 mM DTT

Transcription buffer

20 mM Hepes pH 7.6

60 mM $(\text{NH}_4)_2\text{SO}_4$

8 mM MgSO_4

10 μM ZnCl_2

10 % (v/v) glycerol

10 mM DTT

4.7 Crystallization solutions for Pol II and Pol II ECs

Tartrate crystallization solution

100 mM Hepes pH 7.5

750 -950 mM ammonium-sodium-tartrate

5 mM TCEP

for cryocooling additionally 22 % (v/v) glycerol

NH₄-Mg-acetate crystallization solution (= Natrix #38)

50 mM Hepes pH 7.0

140-200 mM ammonium-acetate

150 mM magnesium-acetate

3.5-5.5 % (v/v) PEG 6000

5 mM TCEP

for cryocooling additionally 20-22 % glycerol

NH₄-Na-acetate crystallization solution (= Natrix #38woMg)

50 mM Hepes pH 7.0

300 mM sodium-acetate

150 mM magnesium-acetate

4.0-7.0 % (v/v) PEG 6000

5 mM TCEP

for cryocooling additionally 20-22 % glycerol

5 Appendix

5.1 Unpublished RNA polymerase II elongation complex structures

5.1.1 Nucleic acid scaffolds

T: 5-bromouracil

AGCT: mismatch between template and nontemplate DNA

TTAA: self complementary region

Mismatch bubble scaffold 14d

```

HK-RNA4:      5'-      UAU AUG CAU AAA
                                     GAC CAG GC
HK-temC>G:    3'-  CCG TCA TGA TCA TTA CTG gTC CgC ATT CAT gAA CTC GAA CC-5'
HK-DNA6:      5'-CCGGC AGT ACT AGT AAA CTA GTA TTG AAA GTA CTT GAG CTT -3'
  
```

Mismatch bubble scaffold KB-2

```

KB-RNA2:      5'-      UAU AUG CAU AAA
                                     GAC CAG GCG
HK-temC>G-Br: 3'-  CCg TCA TgA TCA TTA CTg gTC CgC ATT CAT gAA CTC gAA CC-5'
HK-DNA6:      5'-CCgGc AgT ACT AgT AAA CTA gTA TTg AAA gTA CTT gAg CTT -3'
  
```

Mismatch bubble scaffold KB-3

```

KB-RNA1:      5'-      AAA
                                     GAC CAG GC
KB-tem3-Br:   3'-  CCg TCA TgA TCA TTA CTg gTC CgT TTT CAT gAA CTC gAA CC-5'
HK-DNA6:      5'-CCgGc AgT ACT AgT AAA CTA gTA TTg AAA gTA CTT gAg CTT -3'
  
```

Minimal scaffold KB-4

```

KB-RNA1:      5'-      AAA
                                     GAC CAG GC
HK-DNA4:      3'-      TTA CTg gTC CgT ATT CAT gAA CTC gA -5'
KB-ntem4:     5'-      AAA gTA CTT gAg CT -3'
F: FaPy-G
  
```

Natural bubble scaffold FB-2

```

FB-RNA2:      5'-      UG CAU UUC
                                     GAC CAG GC
HK-temC>G-Br: 3'-  CCg TCA TgA TCA TTA CTg gTC CgC ATT CAT gAA CTC gAA CC -5'
FB-ntem2:     5'-CCgGc AgT ACT AgT AAT GAC CAG GCG TAA gTA CTT gAg CTT GG -3'
  
```

Minimal scaffold FB-3

```

FB-3RNA:      5'-          UG CAU UUC
                GAC CAG GC
FB-3tem:      3'-          TTA CTG GTC CGC ATT GAT GAA CTC GA -3'
FB-3ntem:     5'-          TAA CTA CTT GAG CT -3'

```

Minimal scaffold FB-5

```

FB-3RNA:      5'-          UG CAU UUC
                GAC CAG GC
FB-3tem:      3'-          TTA CTG GBC CGC ATT GAT GAA CTC GA -3'
FB-5ntem:     5'-          AAA CTA CTT GAG CT -3'

```

Minimal scaffold FB-7

```

KB-RNA1:      5'-          AAA
                GAC CAG GC
FB-3tem:      3'-          TTA CTG GBC CGC ATT GAT GAA CTC GA -3'
FB-5ntem:     5'-          AAA CTA CTT GAG CT -3'

```

5.1.2 Crystallographic results

Bubble scaffold 14d

This scaffold was used for the structure of the complete 12-subunit Pol II EC structure by Hubert Kettenberger (Kettenberger et al, 2004). It is listed here for the purpose of comparison, because all other scaffold used in this work were designed based on this scaffold.

Mismatch bubble scaffold KB-2

Mismatch bubble scaffold KB-2 Pol II EC	
<i>Assembly</i>	
Protein	Pol II Δ Rpb4 recombinant Rpb4/7
<i>Crystallization</i>	
Crystallization solution	NH ₄ -Mg-acetate crystallization solution (chapter 4.7)
Cocrystallization	15.7 μ M scaffold KB-2, 12-subunit Pol II was assembled without scaffold
<i>Special crystal treatment</i>	
Soak (in final cryo solution)	2 μ M bubble scaffold KB-2
Dataset(s)	fb239
<i>Data collection</i>	
Detector	MAR225 (SLS PX1)
Program for integration/scaling	Denzo/Scalepack
Space group	C222 ₁
Unit cell axes a, b, c (Å)	222.8, 393.3, 284.9
Unit cell angles α , β , γ (°)	90.0, 90.0, 90.0
Wavelength (Å)	0.91883
Resolution (Å)	50-4.0 (4.14-4.0)
Completeness (%)	99.8 (99.9)
Redundancy	4.7 (4.5)
R _{sym} (%)	10.5 (43.0)
I/ σ	12.3 (3.6)
<i>Refinement</i>	
Program for refinement	CNS version 1.1
Type of refinement	Rigid body refinement, using 1WCM as a model
R _{cryst} (%)	32.8 (Rigid body refinement), 29.2 (Model map)
R _{free} (%)	-
Br peaks in anomalous Fourier (σ)	7 (-5)
Translocation state (from Br peak)	Post-translocation
Observations	<ul style="list-style-type: none"> • NTP site empty, no base pair at register +1 • bases at register +2 apparently form no base pair local lack of electron density at register -8 • probably there is a base pair present at register -9

Mismatch bubble scaffold KB-3

Mismatch bubble scaffold KB-3 Pol II EC	
<i>Assembly</i>	
Protein	Pol II Δ Rpb4 recombinantRpb4/7
<i>Crystallization</i>	
Crystallization solution	NH ₄ -Mg-acetate crystallization solution (chapter 4.7)
Cocrystallization	15.7 μ M scaffold KB-3, 12-subunit Pol II was assembled without scaffold
<i>Special crystal treatment</i>	
Soak (in final cryo solution)	2 μ M bubble scaffold KB-3
Dataset(s)	fb222
<i>Data collection</i>	
Detector	MAR225 (SLS PX2)
Program for integration/scaling	Denzo/Scalepack
Space group	P2 ₁
Unit cell axes a, b, c (Å)	223.1, 284.3, 225.0
Unit cell angles α , β , γ (°)	90.0, 119.6, 90.0
Wavelength (Å)	0.91881
Resolution (Å)	50-4.0 (4.14-4.0)
Completeness (%)	99.6 (98.1)
Redundancy	3.6 (3.2)
R _{sym} (%)	10.8 (44.5)
I/ σ	11.1 (2.8)
<i>Refinement</i>	
Type of refinement	Rigid body refinement, using 1WCM as a model
Program for refinement	CNS version 1.1
R _{cryst} (%)	35.0 (Rigid body refinement), 32.6 (Model map)
R _{free} (%)	-
Br peaks in anomalous Fourier (σ)	7.5 (-4), 8.5 (-4)
Translocation state (from Br peak)	Post-translocation
Observations	NTP site empty, no base pair at register +1 → Post-translocation state

Mismatch bubble scaffold KB-3 Pol II EC with 3'-dATP (soaked)	
<i>Assembly</i>	
Protein	Pol II Δ CTD Δ Rpb4 recombinantRpb4/7
Nucleic acids	Bubble scaffold KB-3
<i>Crystallization</i>	
Crystallization solution	NH ₄ -Mg-acetate crystallization solution (chapter 4.7)
Cocrystallization	1 μ M bubble scaffold KB-3
<i>Special crystal treatment</i>	
Soak (in final cryo solution)	1 μ M bubble scaffold KB-3 2 mM 3'-dATP (cordycepin)
Dataset(s)	fb41
<i>Data collection</i>	
Detector	MAR225 (SLS PX1)
Program for integration/scaling	HKL2000
Space group	C222 ₁
Unit cell axes a, b, c (Å)	220.2, 394.5, 283.8
Unit cell angles α , β , γ (°)	90.0, 90.0, 90.0
Wavelength (Å)	0.91977
Resolution (Å)	50-4.0 (4.14-4.0)
Completeness (%)	98.9 (93.1)
Redundancy	4.6 (3.4)
R _{sym} (%)	11.5 (38.9)
I/ σ	11.8 (2.3)
<i>Refinement</i>	
Type of refinement	Rigid body refinement, using 1WCM as a model
Program for refinement	CNS version 1.1
R _{cryst} (%)	30.4 (Rigid body refinement)
R _{free} (%)	-
Translocation state (from Br peak)	Pre-translocation/back-tracked
Br peaks in anomalous Fourier (σ)	5.5 (-3)
Observations	<ul style="list-style-type: none"> • difference density below active center towards funnel • weak electron density in NTP site
Conclusions	<p>possible scenario 1:</p> <ul style="list-style-type: none"> • incorporation of 3'-dAMP • backstepping by one base pair • fraying/dinucleotide cleavage <p>possible scenario 2:</p> <ul style="list-style-type: none"> • incorporation of 3'-dAMP • translocation • backstepping by two base pairs • fraying/dinucleotide cleavage

Minimal scaffold KB-4

Minimal scaffold KB-4 Pol II EC	
<i>Assembly</i>	
Protein	Pol II Δ CTD Δ Rpb4 recombinantRpb4/7
Nucleic acids	Minimal scaffold KB-4
<i>Crystallization</i>	
Crystallization solution	600 mM ammonium acetate 150 mM magnesium acetate 50 mM HEPES, pH 7.0 3.5-4.5 % PEG 6000 5 mM TCEP
Cocrystallization	1 μ M bubble scaffold KB-4
<i>Special crystal treatment</i>	
Soak (in final cryo solution)	2 μ M bubble scaffold KB-4
Dataset(s)	fb102
<i>Data collection</i>	
Detector	MAR225 (SLS PX1)
Program for integration/scaling	HKL2000
Space group	C222 ₁
Unit cell axes a, b, c (Å)	220.8, 393.5, 282.4
Unit cell angles α , β , γ (°)	90.0, 90.0, 90.0
Wavelength (Å)	0.91947
Resolution (Å)	50-4.0 (4.14-4.0)
Completeness (%)	95.3 (71.2)
Redundancy	6.5 (4.4)
R _{sym} (%)	11.4 (36.0)
I/ σ I	14.6 (2.8)
<i>Refinement</i>	
Type of refinement	Rigid body refinement, using 1WCM as a model
Program for refinement	CNS version 1.1
R _{cryst} (%)	31.6 (Rigid body refinement)
R _{free} (%)	-
Br peaks in anomalous Fourier (σ)	6.0 (-3)
Translocation state (from Br peak)	Pre-translocation
Observations	<ul style="list-style-type: none"> • no density in NTP site • weak electron density for base at expected position for FaPydG (+1) • missing electron density for template backbone between register +1 and +2
Conclusions	To decide, whether the EC is pre-translocated (back-stepped) because FaPydG cannot be accommodated at position -1, a control experiment is required using the same scaffold replacing FaPydG with guanine

Natural bubble scaffold FB-2

Natural bubble scaffold FB-2 Pol II EC					
<i>Assembly^a</i>					
Protein	Pol II Δ Rpb4 recombinantRpb4/7				
Nucleic acids	Natural bubble scaffold FB-7				
<i>Crystallization</i>					
Crystallization conditions	NH ₄ -Mg-acetate crystallization solution (chapter 4.7)				
Cocrystallization	-				
<i>Special crystal treatment</i>					
Soak (in final cryo solution)					
Dataset(s)	fb334	fb337			
<i>Data collection</i>					
Detector	MAR225 (SLS PX1)				
Program for integration/scaling	Denzo/Scalepack				
Space group	C222 ₁	C2			
Unit cell axes a, b, c (Å)	221.7 394.6 283.6	222.1 393.6 283.2			
Unit cell angle β (°)	90.0	90.0			
Wavelength (Å)	0.91909	0.91901			
Resolution (Å)	50.0-4.1 (4.25-4.10)	50.0-4.0 (4.14-4.10)			
Completeness (%)	96.6 (88.4)	99.7 (99.8)			
Redundancy	5.0 (3.6)	6.1 (5.1)			
R _{sym} (%)	10.5 (45.9)	11.0 (49.9)			
I/ σ	11.0 (2.5)	14.4 (2.9)			
<i>Refinement</i>					
Type of refinement	Rigid body refinement, using 1WCM as a model				
Program for refinement	CNS version 1.1				
R _{cryst} (%)	32.3	32.1			
R _{free} (%)	-	-			
Br peaks in anomalous Fourier (σ)	-	-			
Translocation state (from Br peak)	?	?			

Natural bubble scaffold FB-2 Pol II EC	
<i>Observations (for all datasets)</i>	
only partial/fragmented electron density for nucleic acids electron density for protein is ok	
<i>Observations for the individual datasets</i>	
Dataset	Observation
fb334	<ul style="list-style-type: none"> • DNA/RNA hybrid: <ul style="list-style-type: none"> ○ strong electron density for bases , but not well defined ○ weak density for the backbone ○ probably different states/conformations exist ○ strongest density in Fo-Fc map: template positions -2 to -5 • downstream DNA: fragmented electron density
fb337	<ul style="list-style-type: none"> • DNA/RNA hybrid: <ul style="list-style-type: none"> ○ fragmented electron density ○ strongest density in Fo-Fc and 2Fo-Fc map: template positions -3 to -5 and RNA positions -1 to -4 • downstream DNA: <ul style="list-style-type: none"> ○ fragmented electron density ○ strongest density in 2Fo-Fc map: nontemplate positions +3 to +6
<i>Conclusions</i>	
<p>The fact that that electron density for the nucleic acids of the EC was fragmented and weak on the one hand, especially for the backbone, and undefined on the other hand may have two causes:</p> <p>1) Low occupancy of nucleic acids One possible cause may be that no additional nucleic acid scaffold could not be added for cocrystallization and soaking in the final cryo solution (usually 1-2 μM are added). In the case of the natural bubble scaffold FB-2 assembly of the complete scaffold is not possible since the completely complementary nontemplate DNA would presumably displace the RNA. To achieve better occupancy of the nucleic acids in the EC an alternative assembly strategy might be tried: 12-subunit Pol II would be assembled and cocrystallized with annealed template DNA/RNA; in the final cryo solution an additional amount of annealed template DNA/RNA would be added before the nontemplate DNA would be soaked into the crystals to complete the ECs.</p> <p>2) Mixture of states/conformations The natural bubble scaffold FB-2 may not bind in a unique way, different translocational states may coexist.</p>	

^aAssembly protocol:

Template DNA (HK-temC>G-Br) and RNA (FB-RNA2) of the natural bubble scaffold FB-2 were annealed by mixing equimolar amounts in TE buffer (chapter 4.6) at a final concentration of 200 μ M, heating the mixture to 95°C for 5 min, and slow-cooling to 4 °C in a thermocycler. Pol II ECs were assembled by incubating core Pol II for 10 min at 20 °C with 2 molar equivalents of annealed template DNA/RNA, followed by 10 min incubation at 25 °C with 4 molar equivalents of nontemplate DNA (FB-ntem2) and then 10 min incubation at 25 °C with 5 molar equivalents of recombinant Rpb4/7 in assembly buffer (chapter 4.6). The complex was then purified by size exclusion chromatography (Superose 6 10/300 GL) in Pol II buffer (chapter 4.6).

Minimal scaffold FB-3

Minimal scaffold FB-3 Pol II EC without/with GMPCPP and/or α -amanitin					
<i>Assembly</i>					
Protein	Pol II Δ Rpb4 recombinantRpb4/7				
Nucleic acids	Minimal scaffold FB-3				
<i>Crystallization</i>					
Crystallization conditions	NH ₄ -Mg-acetate or NH ₄ -Na-acetate crystallization solution (chapter 4.7)				
	150 mM Mg-acetate	150 mM Mg-acetate	300 mM Na-acetate	300 mM Na-acetate	300 mM Na-acetate
Cocrystallization	2 μ M bubble scaffold FB-3				
<i>Special crystal treatment</i>					
Soak (in final cryo solution)	2 μ M bubble scaffold FB-3				
	-	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 3 mM GMPCPP	5 mM Mg-acetate 3 mM GMPCPP	5 mM Mg-acetate 100 μ M α -amanitin 3 mM GMPCPP
Dataset(s)	fb404	fb369	fb396	fb367	fb382
<i>Data collection</i>					
Detector	PILATUS 6M (SLS PX1)	MAR225 (SLS PX1)	PILATUS 6M (SLS PX1)	MAR225 (SLS PX1)	MAR225 (SLS PX1)
Program for integration/scaling	XDS/XSCALE with 0-dose extrapolation				
Space group	C222 ₁	C222 ₁	C2	C222 ₁	C2
Unit cell axes a, b, c (Å)	223.5	221.9	395.1	221.7	393.6
	394.7	392.4	222.7	393.0	221.5
	283.5	282.7	284.7	283.3	282.2
Unit cell angle β (°)	90.0	90.0	90.7	90.0	90.5
Wavelength (Å)	0.91909	0.91891	0.91909	0.91909	0.91891
Resolution (Å)	50.0-3.8 (3.94-3.80)	50.0-3.6 (3.73-3.60)	50.0-3.6 (3.73-3.60)	50.0-3.5 (3.60-3.50)	50.0-3.4 (3.52-3.40)
Completeness (%)	98.0 (96.7)	99.9 (100)	96.1 (98.2)	98.2 (94.5)	98.3 (90.5)
Redundancy	2.8 (2.7)	4.3 (4.3)	1.8 (1.8)	2.6 (2.5)	2.1 (2.1)
R _{sym} (%)	11.1 (46.8)	10.3 (36.2)	7.3 (43.1)	9.2 (44.3)	6.0 (32.2)
I/ σ I	11.2 (3.4)	12.5 (4.3)	10.4 (2.3)	10.5 (2.4)	12.8 (2.8)
<i>Refinement</i>					
Type of refinement	Rigid body refinement, using 1WCM as a model				
Program for refinement	CNS version 1.2				
R _{cryst} (%)	25.9	29.5	26.5	31.6	27.7
R _{free} (%)	-	-	25.6	32.5	-
Br peaks in anomalous Fourier (σ) (as displayed by Coot (Emsley & Cowtan, 2004))	10.4 (-3)	13.3 (-3)	11.3 (-3/-4)	8.5 (-3)	6.5e (-3/-4)
	8.3 (-4)	15.3 (-4)	6.6 (-3)	9.4 (-4)	6.6e (-3)
			6.6 (-4)		5.2e (-4)
Translocation state (from Br peak)	mixed	mixed	mixed	mixed	mixed

Minimal scaffold FB-3 Pol II EC without/with GMPCPP and/or α-amanitin			
Observations (for all datasets)			
<ul style="list-style-type: none"> • Mixtures between pre- and post-translocation • blob between +1 and +2 template phosphate positions of 1Y1W • some difference density below the active site (conformational change? fraying?) • hybrid in > 1 conformational state (not so clear for fb369) • density for base in NTP-site present, more or less strong (except fb369: no density) 			
<i>Observations for the individual datasets</i>			
Dataset	Soaking	[Mg ²⁺]/mM	Observation
fb404	only FB-3	150	<ul style="list-style-type: none"> • mainly pre-translocated • changed and weakly defined hybrid conformation • weak density in NTP-site
fb369	100 μ M α -amanitin	150	<ul style="list-style-type: none"> • mixture of pre- and post-translocated ca. 1:1 • hybrid conformation similar to 1Y1W • density for amanitin and bound trigger loop • downstream mixture of conformations
fb396	3 mM GMPCPP	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocated • changed and weakly defined hybrid conformation • very weak density in NTP-site
fb367	3 mM GMPCPP	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation • hybrid in >1 conformational state, with one state similar to 1Y1W • density for base in NTP-site • strange repetitive noise
fb382	100 μ M α -amanitin + 3 mM GMPCPP	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation • changed and weakly defined hybrid conformation • weak density in NTP-site • density for amanitin and bound trigger loop

Minimal scaffold FB-5

Minimal scaffold FB-5 Pol II EC without/with GMPCPP and/or α -amanitin					
<i>Assembly</i>					
Protein	Pol II Δ Rpb4 recombinant Rpb4/7				
Nucleic acids	Minimal scaffold FB-5				
<i>Crystallization</i>					
Crystallization conditions	NH ₄ -Na-acetate crystallization solution (chapter 4.7)				
Cocrystallization	2 μ M bubble scaffold FB-3				
<i>Special crystal treatment</i>					
Soak (in final cryo solution)	2 μ M bubble scaffold FB-5				
	-	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin 15 mM GMPCPP
Dataset(s)	fb457	fb460	fb461	fb463	fb465
<i>Data collection</i>					
Detector	PILATUS 6M (SLS PX1)				
Program for integration/scaling	XDS/XSCALE with 0-dose extrapolation				
Space group	C222 ₁	C2	C2	C222 ₁	C2
Unit cell axes a, b, c (Å)	222.7 393.5 283.6	393.3 221.8 282.9	393.4 222.5 283.6	222.1 394.4 283.1	392.9 223.3 283.1
Unit cell angle β (°)	90.0	90.2	90.3	90.0	90.7
Wavelength (Å)	0.91909	0.91909	0.91909	0.91909	0.91909
Resolution (Å)	50.0-3.6 (3.73-3.60)	50.0-3.6 (3.73-3.60)	50.0-3.8 (3.94-3.80)	50.0-3.6 (3.73-3.60)	50.0-3.8 (3.94-3.80)
Completeness (%)	99.2 (100)	95.6 (92.8)	95.9 (91.8)	99.5 (99.9)	95.8 (94.8)
Redundancy	3.8 (3.9)	1.8 (1.8)	1.8 (1.7)	3.9 (3.9)	1.8 (1.7)
R _{sym} (%)	11.2 (52.2)	7.9 (55.2)	7.6 (66.2)	7.6 (54.2)	5.4 (39.1)
I/ σ	9.3 (2.7)	10.4 (1.9)	11.8 (2.3)	13.8 (2.9)	12.3 (2.5)
<i>Refinement</i>					
Type of refinement	Rigid body refinement, using 1WCM as a model				
Program for refinement	CNS version 1.2				
R _{cryst} (%)	27.6	28.4	25.6	27.1	26.2
R _{free} (%)	29.1	26.9	24.8	28.7	25.2
Br peaks in anomalous Fourier (σ) (as displayed by Coot (Emsley & Cowtan, 2004))	12.4 (-3/-4)	7.2 (-3/-4) 4.5 (-3) 7.6 (-3) 5.8 (-4)	6.5 (-3) 5.7 (-4) 7.7 (-3/-4)	9.1 (-3/-4)	7.8 (-3) 7.6 (-4) 7.2 (-4)
Translocation state (from Br peak)	mixed	mixed	mixed	mixed	mixed

Minimal scaffold FB-5 Pol II EC without/with GMPCPP and/or α-amanitin			
<i>Observations (for all datasets)</i>			
<ul style="list-style-type: none"> • Mixtures between pre- and post-translocation • clear mixtures for templating base between “template insertion” and “template pre-insertion” states (regardless of soaking !!!) • blob between +1 and +2 template phosphate positions of 1Y1W • some difference density below the active site (conformational change? fraying?) • hybrid in > 1 conformational state • no density for base in NTP-site present, which seems to be base pairing with the templating base 			
<i>Observations for the individual datasets</i>			
Dataset	Soaking	[Mg ²⁺]/mM	Observation
fb457	only FB-5	0	<ul style="list-style-type: none"> • mainly pre-translocation • changed and weakly defined hybrid conformation
fb460	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation • changed and weakly defined hybrid conformation • density for amanitin and bound trigger loop, the latter weak • blobs for Br in anomalous and fofc do not lie perfectly on top of each other
fb461	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, more pre • changed and weakly defined hybrid conformation • density for amanitin and bound trigger loop, the latter weak
fb463	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, more pre • changed and weakly defined hybrid conformation • density for amanitin and bound trigger loop, the latter weak
fb465	100 μ M α -amanitin + 15 mM GMPCPP	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation (weak Br-signal) • changed and weakly defined hybrid conformation • no density in NTP-site • density for amanitin and bound trigger loop, the latter fragmented and weak

Minimal scaffold FB-7

Minimal scaffold FB-7 Pol II EC without/with α -amanitin					
<i>Assembly</i>					
Protein	Pol II Δ Rpb4 recombinant Rpb4/7				
Nucleic acids	Minimal scaffold FB-7				
<i>Crystallization</i>					
Crystallization conditions	NH ₄ -Na-acetate crystallization solution (chapter 4.7)				
Cocrystallization	2 μ M bubble scaffold FB-7				
<i>Special crystal treatment</i>					
Soak (in final cryo solution)	2 μ M bubble scaffold FB-7				
	-	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin	5 mM Mg-acetate 100 μ M α -amanitin
Dataset(s)	fb444^a	fb437^a	fb438	fb440	fb441
<i>Data collection</i>					
Detector	PILATUS 6M (SLS PX1)				
Program for integration/scaling	XDS/XSCALE with 0-dose extrapolation				
Space group	C222 ₁	C222 ₁	C2	C2	C222 ₁
Unit cell axes a, b, c (Å)	221.9 393.1 283.1	220.6 394.2 284.0	394.0 221.2 283.6	393.9 222.1 283.4	222.6 394.2 283.4
Unit cell angle β (°)	90.0	90.0	90.3	90.6	90.0
Wavelength (Å)	0.91909	0.91909	0.91909	0.91909	0.91901
Resolution (Å)	50.0-3.6 (3.73-3.60)	50.0-3.4 (3.52-3.40)	50.0-3.4 (3.52-3.40)	50.0-3.4 (3.52-3.40)	50.0-3.4 (3.52-3.40)
Completeness (%)	99.9 (99.9)	99.9 (99.9)	94.4 (96.1)	92.6 (98.0)	99.7 (99.9)
Redundancy	3.8 (3.6)	3.6 (3.5)	1.8 (1.8)	1.8 (1.8)	3.9 (3.9)
R _{sym} (%)	6.7 (36.8)	13.3 (46.7)	5.4 (41.0)	7.3 (41.7)	7.1 (52.4)
I/ σ	15.9 (3.9)	8.5 (3.0)	13.1 (2.2)	8.9 (2.3)	15.9 (3.0)
<i>Refinement</i>					
Type of refinement	Rigid body refinement, using 1WCM as a model	Full model building and refinement (positional and individual B-factor)	Rigid body refinement, using 1WCM as a model		
Program for refinement	CNS version 1.2				
R _{cryst} (%)	26.2	25.5	27.2	28.9	27.4
R _{free} (%)	-	28.8	26.0	27.2	29.4
Br peaks in anomalous Fourier (σ) (for fb438, fb440, and fb441 as displayed by Coot (Emsley & Cowtan, 2004))	7.9 (-3) 4.9 (-4)	7.8 (-4)	10.8 (-3) 5.6 (-4) 12.6 (-3) 10.5 (-4)	10.2 (-3/-4) 6.5 (-3) 5.0 (-3) 8.9 (-4)	13.8 (-3) 15.6 (-4)
Translocation state (from Br peak)	mixed	post-transl.	mixed	mixed	mixed

^asame datasets as in Table 3

Minimal scaffold FB-7 Pol II EC with α-amanitin and GMPCPP					
<i>Assembly</i>					
Protein	Pol II Δ Rpb4 recombinant Rpb4/7				
Nucleic acids	Minimal scaffold FB-7				
<i>Crystallization</i>					
Crystallization conditions	NH ₄ -Na-acetate crystallization solution (chapter 4.7)				
Cocrystallization	2 μ M bubble scaffold FB-7				
<i>Special crystal treatment</i>					
Soak (in final cryo solution)	2 μ M bubble scaffold FB-7				
	5 mM Mg-acetate 100 μ M α -amanitin 3 mM GMPCPP	5 mM Mg-acetate 3 mM GMPCPP 100 μ M α -amanitin			
Dataset(s)	fb425	fb433			
<i>Data collection</i>					
Detector	PILATUS 6M (SLS PX1)				
Program for integration/scaling	XDS/XSCALE with 0-dose extrapolation				
Space group	C222 ₁	C2			
Unit cell axes a, b, c (Å)	222.3 393.5 283.5	394.8 223.1 284.2			
Unit cell angle β (°)	90.0	90.3			
Wavelength (Å)	0.91908	0.91908			
Resolution (Å)	50.0-3.6 (3.73-3.60)	50.0-3.5 (3.63-3.50)			
Completeness (%)	99.9 (100)	98.4 (98.0)			
Redundancy	3.9 (3.8)	1.7 (1.7)			
R _{sym} (%)	6.7 (34.2)	6.4 (2.7)			
I/ σ	16.6 (4.5)	9.6 (2.8)			
<i>Refinement</i>					
Type of refinement	Rigid body refinement, using 1WCM as a model				
Program for refinement	CNS version 1.2				
R _{cryst} (%)	26.0	26.9			
R _{free} (%)	-	-			
Br peaks in anomalous Fourier (σ) (as displayed by Coot (Emsley & Cowtan, 2004))	21.2 (-3/-4)	13.4 (-3) 10.3 (-4) 10.2 (-3) 11.9 (-4)			
Translocation state (from Br peak)	mixed	mixed			

Minimal scaffold FB-7 Pol II EC with α-amanitin and GMPCPP			
<i>Observations (for all datasets)</i>			
<ul style="list-style-type: none"> • Mixtures between pre- and post-translocation (except FB437) • blob between +1 and +2 template phosphate positions of 1Y1W • some difference density below the active site (conformational change? fraying?) (except FB437) • hybrid in > 1 conformational state (except FB 437, also one state seems to be prominent for FB440) • mostly weak density for base in NTP-site present, which seems to be base pairing with the templating base (no density observed in FB440, FB441, FB437) • templating base mainly in "template insertion" site (mixtures apparent e.g. in FB425, FB433, FB440; FB437 is "pre-insertion"!) 			
<i>Observations for the individual datasets</i>			
Dataset	Soaking	[Mg ²⁺]/mM	Observation
fb444	only FB-7	0	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, more pre • changed and weakly defined hybrid conformation • downstream mixture of conformations
fb437	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • post-translocation • template base in pre-insertion site • hybrid density similar to 1Y1W and well defined • density for amanitin and bound trigger loop
fb438	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, little more pre • changed and weakly defined hybrid • conformation density for amanitin and bound trigger loop • part of fork loop 2 ordered • difference density peaks on upstream side of bridge helix
fb440	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, more post • hybrid conformation more similar to 1Y1W • density for amanitin and bound trigger loop • part of fork loop 2 ordered • mixture between +1 template insertion and pre-insertion • difference density peaks on upstream side of bridge helix
fb441	100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation • changed and weakly defined hybrid conformation density for amanitin and bound trigger loop • part of fork loop 2 ordered
fb425	100 μ M α -amanitin + 3 mM GMPCPP	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, ca. 1:1 • changed and weakly defined hybrid conformation • weak density in NTP-site • density for amanitin and bound trigger loop
fb433	3 mM GMPCPP + 100 μ M α -amanitin	5	<ul style="list-style-type: none"> • mixture of pre- and post-translocation, ca. 1:1 • changed and weakly defined hybrid conformation • density for amanitin and bound trigger loop

5.2 Attempts to improve the resolution of diffraction data from complete 12-subunit RNA polymerase II crystals

5.2.1 Capillary mounting and measuring at room temperature

To find out, whether cryo-cooling of 12-subunit Pol II EC crystals influences the crystal parameters and diffraction quality, crystals of an EC with a minimum bubble scaffold (Scaffold C^U, Figure 16, page 39) were measured at room temperature. Crystallization was performed as described in chapter 3.3.2, crystallization solution contained 170 mM ammonium acetate, 150 mM magnesium acetate, 50 mM Hepes pH 7.0, 3.5-5.5 % PEG 6000, 5 mM TCEP (NH₄-Mg-acetate crystallization solution, see chapter 4.7). For room temperature measurement, crystals were mounted in Quartz capillaries (0.5-1.0 mm diameter, Hampton research) at room temperature. The crystals were sucked into the capillaries, surrounding mother liquor was removed and a reservoir of mother liquor was placed at both ends of the capillary. Finally the capillaries were sealed with wax at both ends. To test additionally, whether adding the cryoprotectant glycerol changes crystal and/or diffraction properties, crystals with and without glycerol were measured. Four crystals were measured, for which the standard cryoprotectant treatment was applied (0-20 % glycerol in 5 steps), except that incubation over night was at room temperature instead of 4 °C. Three crystals were measured without the cryoprotectant treatment applied, i.e. glycerol added. The capillaries containing the crystals were then transferred carefully in a styrofoam box to the synchrotron beamline PX1 at the SLS. For each crystal several 0.25 ° oscillation images with 45 ° spacing were recorded (room temperature, exposure time: 1 s, readout time: approx. 4 s, no filter, ring current 350 mA, detector MAR225).

Signs of radiation damage could in general already be observed after the first image (1 s exposure), in particular a reduction in the diffraction intensity at higher resolution. On the first image of all tested capillary-mounted crystals, the spots on the detector were particularly small and round-shaped, as compared to cryo-cooled crystals. However the maximum resolution, estimated visually, was somewhat lower (5-8Å), but there was no significant difference between crystals with and without cryoprotectant treatment. Diffraction data was indexed with the program MOSFLM (Leslie, 2006) from the first one to three images to test whether cryo-cooling and/or addition of glycerol have an effect on unit cell parameters (Table 7 and Table 8). No significant change of unit cell parameters could be observed.

Table 7: Maximum resolution, space group and unit cell parameters of capillary-mounted Pol II EC crystals with cryoprotectant treatment.

Crystal #	Maximum resolution	Space group	Unit cell axes		
			a	b	c
fb310	6	C222 ₁	223	395	285
fb311	6	C222 ₁	224	395	284
fb313	5	C222 ₁	223	395	285
fb314	6.5	C222 ₁	223	395	285

Table 8: Maximum resolution, space group and unit cell parameters of capillary-mounted Pol II EC crystals without cryoprotectant treatment

Crystal #	Maximum resolution	Space group	Unit cell axes		
			a	b	c
fb316	8	C222 ₁	222	405	284
fb317	6.5	C222 ₁	222	395	284
fb318	7.5	C222 ₁	223	394	284

5.2.2 Crystallization of 12-subunit Pol II without CTD

The CTD of Pol II was removed, using a newly introduced TEV cleavage site (Kettenberger, 2005). The resulting 12-subunit Pol II without CTD crystallized under the same conditions as Pol II with CTD. The crystals were crystallographically indistinguishable from the 12-subunit Pol II crystals, however large crystals apparently formed slightly more readily and reproducibly.

5.2.3 Further attempts in table form

Approach	C ^a	Reference	Observation/Conclusion
New crystallization conditions: 300 mM Na-acetate (NH ₄ -Na-acetate crystallization solution, chapter 4.7) replace 150 mM Mg-acetate (NH ₄ -Mg-acetate crystallization solution, chapter)	3	-	Pol II EC apparently crystallizes more readily and reproducibly to form large three-dimensional crystals, from which apparently diffraction data to slightly higher resolution can be collected.

Approach	C ^a	Reference	Observation/Conclusion
Crosslinking with glutaraldehyde before cryoprotectant treatment	1,2	(Lusty, 1999)	crystals became prone to dissolving in pure water; no significant change of diffraction properties
Using oils (Paraffin oil, PFO) as cryoprotectants: removal of adhering mother liquor from crystal by dragging through the oil.	1,3	-	no improvement of diffraction properties
Removal of the cold gas layer above liquid nitrogen for by blowing with a dry nitrogen gas stream for cryo-cooling by plunging into liquid nitrogen	1,2,3	(Warkentin et al, 2006)	no significant change of diffraction properties
Cryo-cooling by plunging into liquid Ethane	2,3	-	no significant change of diffraction properties
Crystal annealing by simply blocking the cryostream		(Heras & Martin, 2005) and references therein	diffraction lost completely
Crystal annealing by blocking the cryostream, transferring the crystal back to the mother liquor, incubating several minutes in the mother liquor and re-cooling in the cryostream		(Heras & Martin, 2005) and references therein	diffraction intensity and maximum resolution significantly lowered, but diffraction is not completely lost
New pixel detector PILATUS 6M	1,2,3	(Broennimann et al, 2006)	
XDS/XSCALE with 0-dose extrapolation for indexing, integration/scaling instead of DENZO/SCALEPACK	1,2,3	(Kabsch, 1993) (Diederichs et al, 2003)	better data statistics, using 0-dose extrapolation
CNS version 1.2 instead of CNS version 1.1 for structure refinement and map calculation	1,2,3	(Brunger, 2007)	Improved bulk solvent correction with parameter grid search improves R-factor

^aCrystallization solution used for crystallization and during crystal treatment: (1) Tartrate (2) NH₄-Mg-acetate (3) NH₄-Na-acetate crystallization solution (see chapter 4.7)

5.2.4 Ideas for the future

- Screen for new crystal form of 12-subunit Pol II without CTD

-
- Screen for new crystal form of 12-subunit Pol II EC using Pol II without CTD and/or nucleic acid scaffolds forming crystal contacts
 - Controlled dehydration of 12-subunit Pol II crystals using the free mounting system (Kiefersauer et al, 2000).
 - Try radioprotectants to reduce radiation damage during data collection (Kauffmann et al, 2006; Southworth-Davies & Garman, 2007)
 - Structure refinement with TLS for improving R-factor and electron density maps (Re-
fmac, phenix.refine)

6 Conclusions and Outlook

Results presented in chapter 2 of this thesis have led to a more detailed mechanistic understanding of translocation by RNA polymerase II. Together with the other crystallographic snapshots of the nucleotide addition cycle, a movie of elongating Pol II can now be assembled. Further insights into the coupling of catalysis and translocation could be obtained by trapping the EC right after catalysis but before the release of pyrophosphate, and structurally analyzing the resulting product complex.

The studies, described in chapter 3 of this thesis unveiled distinct molecular recognition mechanisms of transcribing Pol II for two different bulky DNA lesions. Because of the apparent absence of a general lesion recognition mechanism, the effect of other kinds of DNA lesions on transcribing Pol II will have to be examined individually. This also includes lesions, which can be bypassed by Pol II but may cause transcriptional mutagenesis, like some oxidative lesions.

In the absence of a DNA lesion, mismatch nucleotide incorporation and extension of the mismatch lead to transcriptional mutagenesis, affecting the fidelity of gene transcription. For a more comprehensive understanding of transcriptional fidelity, it is necessary to survey the events causing errors in the mRNA and analyze the mechanisms preventing them.

Many more questions on Pol II elongation remain unanswered. Structural and biochemical characterization of functional Pol II ECs can provide extremely valuable and detailed information on the mechanism of transcription, but it is a formidable challenge because of the large size and transient nature of these complexes. Further investigation of Pol II could focus on more natural elongation complexes with a fully complementary nontemplate strand, complexes with additional protein factors involved in transcription, and complexes of Pol II with natural RNA inhibitors, to name just a few.

7 Abbreviations

AMP	adenosine monophosphate
ATP	adenosine triphosphate
CMP	cytosine monophosphate
CPD	cyclobutane pyrimidine dimer
CTD	C-terminal domain of Rpb1 of Pol II
CTP	cytosine triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DPE	downstream promoter element
DTT	dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
EC	elongation complex
EDTA	ethylene diamine tetraacetic acid
EM	electron microscopy
GMP	guanosine monophosphate
GMPCPP	Guanosine-5'-[(α,β)-methylene]triphosphate
GTF	general transcription factor
GTP	guanosine triphosphate
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
INR	initiator element
IPTG	isopropyl β -D-1-thiogalactopyranoside
MALDI-TOF	matrix-assisted laser desorption ionization with time-of-flight analysator mass spectrometry
mRNA	messenger RNA
MWCO	molecular weight cutoff
NAC	nucleotide addition cycle
NER	nucleotide excision repair
NTP	nucleotide triphosphate
PDB	protein data bank
PEG	polyethylene glycol (number indicates average molecular weight in Da)
PFO	perfluoropolyether
PIC	preinitiation complex
Pol	eukaryotic DNA-dependent RNA polymerase

RMSD	root mean square deviation
RNA	ribonucleic acid
RNAP	DNA-dependent RNA polymerase
Rpb	subunit of Pol II (=RNA polymerase B)
<i>S.cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
TAF	TBP-associated factor
<i>Taq</i>	<i>Thermus aquaticus</i>
TBP	TATA binding protein
TCEP	tris(2-carboxyethyl)phosphine
TCR	transcription coupled repair
TFII	transcription factor of Pol II transcription
Tris	trishydroxymethylaminomethane
<i>Tth</i>	<i>Thermus thermophilus</i>
UMP	uridine monophosphate
UTP	uridine triphosphate

8 References

- Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**: 460-465
- Alberts B, Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002) *Molecular biology of the cell*, 4 edn. New York: Garland Science.
- Andrecka J, Lewis R, Bruckner F, Lehmann E, Cramer P, Michaelis J (2008) Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc Natl Acad Sci U S A* **105**: 135-140
- Armache K-J, Kettenberger H, Cramer P (2003) Architecture of the initiation-competent 12-subunit RNA polymerase II. *Proc Natl Ac Sc USA* **100**: 6964-6968
- Armache K-J, Mitterweger S, Meinhart A, Cramer P (2005) Structures of complete RNA polymerase II and its subcomplex Rpb4/7. *J Biol Chem* **280**: 7131-7134
- Armache KJ (2005) Crystal structures of the complete 12-subunit RNA polymerase II and its subcomplex Rpb4/7, and modeling of RNA polymerases I and III. *Dissertation, LMU München: Fakultät für Chemie und Pharmazie*
- Armstrong JA (2007) Negotiating the nucleosome: factors that allow RNA polymerase II to elongate through chromatin. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **85**: 426-434
- Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E (2005) A ratchet mechanism of transcription elongation and its control. *Cell* **120**: 183-193
- Berg JM, Tymoczko, J. L., Stryer, L. (2001) *Biochemistry*, 5 edn. New York: W. H. Freeman and Company.
- Block SM (2007) Kinesin motor mechanics: binding, stepping, tracking, gating, and limping. *Biophys J* **92**: 2986-2995
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254
- Broennimann C, Eikenberry EF, Henrich B, Horisberger R, Huelsen G, Pohl E, Schmitt B, Schulze-Briese C, Suzuki M, Tomizaki T, Toyokawa H, Wagner A (2006) The PILATUS 1M detector. *J Synchrotron Radiat* **13**: 120-130
- Brueckner F, Hennecke U, Carell T, Cramer P (2007) CPD damage recognition by transcribing RNA polymerase II. *Science* **315**: 859-862
- Brunger AT (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**: 2728-2733

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**: 905-921

Buratowski S (2003) The CTD code. *Nat Struct Biol* **10**: 679-680

Bushnell DA, Cramer P, Kornberg RD (2002) Structural basis of transcription: alpha-amanitin-RNA polymerase II cocystal at 2.8 Å resolution. *Proc Natl Acad Sci U S A* **99**: 1218-1222

Bushnell DA, Kornberg RD (2003) Complete RNA polymerase II at 4.1 Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A* **100**: 6969-6972

Bushnell DA, Westover KD, Davis RE, Kornberg RD (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms. *Science* **303**: 983-988

Butenandt J, Burgdorf, L. T., Carell, T. (1985) "Base Flipping": Photodamaged DNA-RNA Duplexes Are Poor Substrates for Photoreactivating DNA-Repair Enzymes. *Angew Chem Int Ed* **38**: 708-711

Butenandt J, Burgdorf, L. T., Carell, T. (1999) "Base Flipping": Photodamaged DNA-RNA Duplexes Are Poor Substrates for Photoreactivating DNA-Repair Enzymes. *Angew Chem Int Ed* **38**: 708-711

Butenandt J, Eker APM, Carell T (1998) *Chemistry-a European Journal*: 642-654

Butenandt J, Epple R, Wallenborn E-U, Eker APM, Gramlich V, Carell T (2000) A Comparative Repair Study of Thymine- and Uracil-Photodimers with Model Compounds and a Photolyase Repair Enzyme. *Chemistry - A European Journal* **6**: 62-72

Cadet J, Voituriez L, Hruska FE, Grand A (1985) Crystal structure of the cis-syn photodimer of thymidylyl (3'-5') thymidine cyanoethyl ester. *Biopolymers* **24**: 897-903

Campbell EA, Pavlova O, Zenkin N, Leon F, Irschik H, Jansen R, Severinov K, Darst SA (2005) Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *Embo J* **24**: 674-682

Chafin DR, Guo H, Price DH (1995) Actions of α -amanitin during pyrophosphorylysis and elongation by RNA polymerase II. *J Biol Chem* **270**: 19114-19119

Charlet-Berguerand N, Feuerhahn S, Kong SE, Ziserman H, Conaway JW, Conaway R, Egly JM (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J* **25**: 5481-5491

Chen H-T, Warfield L, Hahn S (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription initiation complex. *Nat Struct Mol Biol* **8**: 696-703

Corda Y, Job C, Anin MF, Leng M, Job D (1991) Transcription by eucaryotic and procaryotic RNA polymerases of DNA modified at a d(GG) or a d(AG) site by the antitumor drug cis-diamminedichloroplatinum(II). *Biochemistry* **30**: 222-230

Corda Y, Job C, Anin MF, Leng M, Job D (1993) Spectrum of DNA--platinum adduct recognition by prokaryotic and eukaryotic DNA-dependent RNA polymerases. *Biochemistry* **32**: 8582-8588

Cramer P (2002a) Common structural features of nucleic acid polymerases. *Bioessays* **24**: 724-729

Cramer P (2002b) Multisubunit RNA polymerases. *Curr Op Struct Biol* **12**: 89-97

Cramer P (2007) Gene transcription: extending the message. *Nature* **448**: 142-143

Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD (2000) Architecture of RNA polymerase II and implications for the transcription mechanism [see comments]. *Science* **288**: 640-649

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876.

Crick F (1970) Central dogma of molecular biology. *Nature* **227**: 561-563

Damsma GE, Alt A, Brueckner F, Carell T, Cramer P (2007) Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* **14**: 1127-1133

Darst SA, Edwards AM, Kubalek EW, Kornberg RD (1991) Three-dimensional structure of yeast RNA polymerase II at 16Å resolution. *Cell* **66**: 121-128

Deaconescu AM, Chambers AL, Smith AJ, Nickels BE, Hochschild A, Savery NJ, Darst SA (2006) Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124**: 507-520

DeLano WL (2002) The PyMOL Molecular Graphics System. *DeLano Scientific*: San Carlos, CA, USA

Diederichs K, McSweeney S, Ravelli RBG (2003) Zero-dose extrapolation as part of macromolecular synchrotron data reduction. *Acta Crystallographica Section D* **59**: 903-909

Donahue BA, Yin S, Taylor JS, Reines D, Hanawalt PC (1994) Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc Natl Acad Sci U S A* **91**: 8502-8506

Edwards AM, Darst SA, Feaver WJ, Thompson NE, Burgess RR, Kornberg RD (1990) Purification and lipid layer crystallization of yeast RNA polymerase II. *Proc Natl Acad Sci USA* **87**: 2122-2126

Edwards AM, Kane CM, Young RA, Kornberg RD (1991) Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter *in vitro*. *J Biol Chem* **266**: 71-75

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

Epshtein V, Mustaev A, Markovtsov V, Bereshchenko O, Nikiforov V, Goldfarb A (2002) Swing-gate model of nucleotide entry into the RNA polymerase active center. *Mol Cell* **10**: 623-634

Fernandez-Tornero C, Bottcher B, Riva M, Carles C, Steuerwald U, Ruigrok RW, Sentenac A, Muller CW, Schoehn G (2007) Insights into Transcription Initiation and Termination from the Electron Microscopy Structure of Yeast RNA Polymerase III. *Mol Cell* **25**: 813-823

Fu J, Gnatt AL, Bushnell DA, Jensen G, J., Thompson NE, Burgess RR, David PR, Kornberg RD (1999) Yeast RNA polymerase II at 5 Å resolution. *Cell* **98**: 799-810

Galbur EA, Grill SW, Wiedmann A, Lubkowska L, Choy J, Nogales E, Kashlev M, Bustamante C (2007) Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature* **446**: 820-823

Gao X, Brown FK, Jeffs P, Bischofberger N, Lin KY, Pipe AJ, Noble SA (1992) Probing structural factors stabilizing antisense oligonucleotide duplexes: NMR studies of a DNA.DNA duplex containing a formacetal linkage. *Biochemistry* **31**: 6228-6236

Gerber PR, Muller K (1995) MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J Comput Aided Mol Des* **9**: 251-268

Gilmour DS, Fan R (2008) Derailing the locomotive: transcription termination. *J Biol Chem* **283**: 661-664

Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876-1882.

Gong XQ, Nedialkov YA, Burton ZF (2004) Alpha-amanitin blocks translocation by human RNA polymerase II. *J Biol Chem* **279**: 27422-27427

Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394-403

Heras B, Martin JL (2005) Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallographica Section D* **61**: 1173-1180

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118-120

Hirata A, Klein BJ, Murakami KS (2008) The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**: 851-854

Hirose Y, Manley JL (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev* **14**: 1415-1429

Hurwitz J, Bresler A, Diring R (1960) The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. *Biochemical and Biophysical Research Communications* **3**: 15-19

Jasiak AJ, Armache KJ, Martens B, Jansen RP, Cramer P (2006) Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Mol Cell* **23**: 71-81

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst* **A47**: 110-119

Jung Y, Lippard SJ (2006) RNA polymerase II blockage by cisplatin-damaged DNA. Stability and polyubiquitylation of stalled polymerase. *J Biol Chem* **281**: 1361-1370

Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **26**: 795-800

Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* **37**: 761-765

Kartalou M, Essigmann JM (2001) Recognition of cisplatin adducts by cellular proteins. *Mutat Res* **478**: 1-21

Kashkina E, Anikin M, Brueckner F, Lehmann E, Kochetkov SN, McAllister WT, Cramer P, Temiakov D (2007) Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site. *J Biol Chem* **282**: 21578-21582

Kashkina E, Anikin M, Brueckner F, Pomerantz RT, McAllister WT, Cramer P, Temiakov D (2006) Template misalignment in multisubunit RNA polymerases and transcription fidelity. *Mol Cell* **24**: 257-266

Kauffmann B, Weiss MS, Lamzin VS, Schmidt A (2006) How to avoid premature decay of your macromolecular crystal: a quick soak for long life. *Structure* **14**: 1099-1105

Kettenberger H (2005) Structure of the Complete RNA Polymerase II Elongation Complex and its Interaction with the Elongation Factor TFIIS. *Dissertation, LMU München: Fakultät für Chemie und Pharmazie*

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

Kettenberger H, Armache K-J, Cramer P (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* **16**: 955-965

Kettenberger H, Eisenfuhr A, Brueckner F, Theis M, Famulok M, Cramer P (2006) Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by non-coding RNAs. *Nat Struct Mol Biol* **13**: 44-48

Kiefersauer R, Than ME, Dobbek H, Gremer L, Melero M, Strobl S, Dias JM, Soulimane T, Huber R (2000) A novel free-mounting system for protein crystals: transformation and improvement of diffraction power by accurately controlled humidity changes. *Journal of Applied Crystallography* **33**: 1223-1230

Kim TK, Ebright RH, Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**: 1418-1422

Kireeva ML, Komissarova N, Waugh DS, Kashlev M (2000) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J Biol Chem* **275**: 6530-6536

Kireeva ML, Lubkowska L, Komissarova N, Kashlev M (2003) Assays and affinity purification of biotinylated and nonbiotinylated forms of double-tagged core RNA polymerase II from *Saccharomyces cerevisiae*. *Methods Enzymol* **370**: 138-155

Kuhn CD, Geiger SR, Baumli S, Gartmann M, Gerber J, Jennebach S, Mielke T, Tschochner H, Beckmann R, Cramer P (2007) Functional architecture of RNA polymerase I. *Cell* **131**: 1260-1272

Kulaeva OI, Gaykalova DA, Studitsky VM (2007) Transcription through chromatin by RNA polymerase II: histone displacement and exchange. *Mutat Res* **618**: 116-129

Kusser AG, Bertero MG, Naji S, Becker T, Thomm M, Beckmann R, Cramer P (2008) Structure of an archaeal RNA polymerase. *J Mol Biol* **376**: 303-307

Laine JP, Egly JM (2006) Initiation of DNA repair mediated by a stalled RNA polymerase II. *Embo J* **25**: 387-397

- Landick R (2004) Active-site dynamics in RNA polymerases. *Cell* **116**: 351-353
- Lee SK, Yu SL, Prakash L, Prakash S (2002) Yeast RAD26, a homolog of the human CSB gene, functions independently of nucleotide excision repair and base excision repair in promoting transcription through damaged bases. *Mol Cell Biol* **22**: 4383-4389
- Lehmann E (2006) CocrySTALLISATION of yeast RNA polymerase II with mouse B2 RNA and studies of RNA-directed RNA synthesis by RNA polymerase II. *Masterarbeit, LMU München: Fakultät für Chemie und Pharmazie*
- Lehmann E, Brueckner F, Cramer P (2007) Molecular basis of RNA-dependent RNA polymerase II activity. *Nature* **450**: 445-449
- Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**: 48-57
- Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* **128**: 707-719
- Li Y, Dutta S, Doublie S, Bdour HM, Taylor JS, Ellenberger T (2004) Nucleotide insertion opposite a cis-syn thymine dimer by a replicative DNA polymerase from bacteriophage T7. *Nat Struct Mol Biol* **11**: 784-790
- Ling H, Boudsocq F, Plosky BS, Woodgate R, Yang W (2003) Replication of a cis-syn thymine dimer at atomic resolution. *Nature* **424**: 1083-1087
- Lusty C (1999) A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. *Journal of Applied Crystallography* **32**: 106-112
- McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**: 458-464
- Mees A, Klar T, Gnau P, Hennecke U, Eker AP, Carell T, Essen LO (2004) Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. *Science* **306**: 1789-1793
- Mei Kwei JS, Kuraoka I, Horibata K, Ubukata M, Kobatake E, Iwai S, Handa H, Tanaka K (2004) Blockage of RNA polymerase II at a cyclobutane pyrimidine dimer and 6-4 photoproduct. *Biochem Biophys Res Commun* **320**: 1133-1138
- Meinhart A, Kamenski T, Hoepfner S, Baumli S, Cramer P (2005) A structural perspective of CTD function. *Genes Dev* **19**: 1401-1415
- Miller G, Hahn S (2006) A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat Struct Mol Biol* **13**: 603-610

Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebricht RH, Severinov K (2001) Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci U S A* **98**: 892-897.

Mitchell JR, Hoeijmakers JH, Niedernhofer LJ (2003) Divide and conquer: nucleotide excision repair battles cancer and ageing. *Curr Opin Cell Biol* **15**: 232-240

Mu D, Sancar A (1997) Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans. *J Biol Chem* **272**: 7570-7573

Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285-1290.

Murakami Y (2007) Crystal structure of an archaeal RNA polymerase. *to be published*

Naji S, Bertero MG, Spitalny P, Cramer P, Thomm M (2007) Structure function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res*

Otwinowski Z, Minor W (1996) Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzym* **276**: 307-326

Poglitsch CL, Meredith GD, Gnatt AL, Jensen GJ, Chang WH, Fu J, Kornberg RD (1999) Electron crystal structure of an RNA polymerase II transcription elongation complex. *Cell* **98**: 791-798

Prakash S, Prakash L (2000) Nucleotide excision repair in yeast. *Mutat Res* **451**: 13-24

Roeder RG, Rutter WJ (1969) Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234-237

Rudd MD, Luse DS (1996) Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. *J Biol Chem* **271**: 21549-21558

Sakurai H, Mitsuzawa H, Kimura M, Ishihama A (1999) The Rpb4 subunit of fission yeast *Schizosaccharomyces pombe* RNA polymerase II is essential for cell viability and similar in structure to the corresponding subunits of higher eukaryotes. *Mol Cell Biol* **19**: 7511-7518.

Sancar A (1996) DNA excision repair. *Annu Rev Biochem* **65**: 43-81

Sarker AH, Tsutakawa SE, Kostek S, Ng C, Shin DS, Peris M, Campeau E, Tainer JA, Nogales E, Cooper PK (2005) Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. *Mol Cell* **20**: 187-198

Saxowsky TT, Doetsch PW (2006) RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem Rev* **106**: 474-488

Schwartz LB, Sklar VE, Jaehning JA, Weinmann R, Roeder RG (1974) Isolation and partial characterization of the multiple forms of deoxyribonucleic acid-dependent ribonucleic acid polymerase in the mouse myeloma, MOPC 315. *J Biol Chem* **249**: 5889-5897

Selby CP, Drapkin R, Reinberg D, Sancar A (1997) RNA polymerase II stalled at a thymine dimer: footprint and effect on excision repair. *Nucleic Acids Res* **25**: 787-793

Selby CP, Sancar A (1997a) Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proc Natl Acad Sci U S A* **94**: 11205-11209

Selby CP, Sancar A (1997b) Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J Biol Chem* **272**: 1885-1890

Sims RJ, 3rd, Belotserkovskaya R, Reinberg D (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437-2468

Sosunov V, Sosunova E, Mustaev A, Bass I, Nikiforov V, Goldfarb A (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *Embo J* **22**: 2234-2244

Sousa R (2005) Machinations of a maxwellian demon. *Cell* **120**: 155-156

Sousa R, Chung YJ, Rose JP, Wang B-C (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3Å resolution. *Nature* **364**: 593-599

Southworth-Davies RJ, Garman EF (2007) Radioprotectant screening for cryocrystallography. *Journal of Synchrotron Radiation* **14**: 73-83

Steitz TA (1998) A mechanism for all polymerases [news; comment]. *Nature* **391**: 231-232

Steitz TA (2004) The structural basis of the transition from initiation to elongation phases of transcription, as well as translocation and strand separation, by T7 RNA polymerase. *Curr Opin Struct Biol* **14**: 4-9

Steitz TA (2006) Visualizing polynucleotide polymerase machines at work. *Embo J* **25**: 3458-3468

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**: 207-234

- Svejstrup JQ (2002) Mechanisms of transcription-coupled DNA repair. *Nat Rev Mol Cell Biol* **3**: 21-29
- Svejstrup JQ (2003) Rescue of arrested RNA polymerase II complexes. *J Cell Sci* **116**: 447-451
- Svejstrup JQ (2004) The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta* **1677**: 64-73
- Tantin D (1998) RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62. *J Biol Chem* **273**: 27794-27799
- Taylor JS, Garrett DS, Brockie IR, Svoboda DL, Telser J (1990) ¹H NMR assignment and melting temperature study of cis-syn and trans-syn thymine dimer containing duplexes of d(CGTATTATGC).d(GCATAATACG). *Biochemistry* **29**: 8858-8866
- Temiaikov D, Patlan V, Anikin M, McAllister WT, Yokoyama S, Vassilyev DG (2004) Structural basis for substrate selection by t7 RNA polymerase. *Cell* **116**: 381-391
- Thieffry D, Sarkar S (1998) Forty years under the central dogma. *Trends Biochem Sci* **23**: 312-316
- Thompson NE, Burgess RR (1996) Immunoaffinity purification of RNA polymerase II and transcription factors using polyol-responsive monoclonal antibodies. *Methods Enzymol* **274**: 513-526
- Till S, Ladurner AG (2007) RNA Pol IV plays catch with Argonaute 4. *Cell* **131**: 643-645
- Tornaletti S, Donahue BA, Reines D, Hanawalt PC (1997) Nucleotide sequence context effect of a cyclobutane pyrimidine dimer upon RNA polymerase II transcription. *J Biol Chem* **272**: 31719-31724
- Tornaletti S, Patrick SM, Turchi JJ, Hanawalt PC (2003) Behavior of T7 RNA polymerase and mammalian RNA polymerase II at site-specific cisplatin adducts in the template DNA. *J Biol Chem* **278**: 35791-35797
- Toulokhonov I, Zhang J, Palangat M, Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* **27**: 406-419
- Tremeau-Bravard A, Riedl T, Egly JM, Dahmus ME (2004) Fate of RNA polymerase II stalled at a cisplatin lesion. *J Biol Chem* **279**: 7751-7759
- Tuske S, Sarafianos SG, Wang X, Hudson B, Sineva E, Mukhopadhyay J, Birktoft JJ, Leroy O, Ismail S, Clark AD, Jr., Dharia C, Napoli A, Laptenko O, Lee J, Borukhov S, Ebright RH,

Arnold E (2005) Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* **122**: 541-552

Vale RD, Milligan RA (2000) The way things move: looking under the hood of molecular motor proteins. *Science* **288**: 88-95

Vassylyev DG, Sekine S, Laptenko O, Lee J, Vassylyeva MN, Borukhov S, Yokoyama S (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712-719.

Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I (2007a) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162

Vassylyev DG, Vassylyeva MN, Zhang J, Palangat M, Artsimovitch I, Landick R (2007b) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**: 163-168

Veal JM, Gao X, Brown FK (1993) A comparison of DNA oligomer duplexes containing formacetal and phosphodiester linkers using molecular dynamics and quantum mechanics. *J Am Chem Soc* **115**: 7139-7145

Verhage RA, Heyn J, van de Putte P, Brouwer J (1997) Transcription elongation factor S-II is not required for transcription-coupled repair in yeast. *Mol Gen Genet* **254**: 284-290

Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**: 941-954

Wang D, Lippard SJ (2005) Cellular processing of platinum anticancer drugs. *Nat Rev Drug Discov* **4**: 307-320

Warkentin M, Berejnov V, Hussein NS, Thorne RE (2006) Hyperquenching for protein cryocrystallography. *Journal of Applied Crystallography* **39**: 805-811

Weiss SB, Nakamoto T (1961) Net Synthesis of Ribonucleic Acid with a Microbial Enzyme Requiring Deoxyribonucleic Acid and Four Ribonucleoside Triphosphates. *J Biol Chem* **236**: PC18-20

Westover KD, Bushnell DA, Kornberg RD (2004a) Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**: 481-489

Westover KD, Bushnell DA, Kornberg RD (2004b) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**: 1014-1016

Wieland T, Faulstich H (1991) Fifty years of amanitin. *Experientia* **47**: 1186-1193

Wind M, Reines D (2000) Transcription elongation factor SII. *Bioessays* **22**: 327-336

Woudstra EC, Gilbert C, Fellows J, Jansen L, Brouwer J, Erdjument-Bromage H, Tempst P, Svejstrup JQ (2002) A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature* **415**: 929-933

Yin YW, Steitz TA (2004) The structural mechanism of translocation and helicase activity in t7 RNA polymerase. *Cell* **116**: 393-404

Zanotti G, Petersen G, Wieland T (1992) Structure-toxicity relationships in the amatoxin series. *Int J Peptide Protein Res* **40**: 551-558

Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**: 811-824

9 Curriculum vitae – Florian Brückner

9.1 Personal Details



Name	Florian Brückner
Private address	Saalburgstr. 9 81375 München
Institute address	Department of Chemistry and Biochemistry Gene Center University of Munich Feodor-Lynen-Str. 25 D-81377 München
Private/institute phone	+49-89-71668913 / +49-89-2180-76955
Private mobile phone	+49-151-17878779
Email	brueck@lmb.uni-muenchen.de
Date of birth	March 12, 1978
Place of birth	Munich, Germany
Nationality	German
Marital status	married
Foreign languages	English, Russian

9.2 Curriculum vitae

1988-1997	Grammar school in Landsberg/Lech
June 27, 1997	Qualification of university matriculation (Abitur)
1997-1998	Military service (German Airforce)
1998-2003	Diploma course in Biology at the Technical University of Munich. <i>Main subjects: Biochemistry, Organic chemistry, Genetics</i>
2003-2004	Diploma thesis at the Technical University of Munich, Institute for Biological Chemistry (Prof. Dr. Arne Skerra) under the supervision of Dr. Ingo Korndörfer <i>Topic: Production and preliminary crystallographic analysis of crystals of the human S100A8/S100A9 heterodimer</i>
April 1, 2004	Diploma degree in Biology (passed with distinction)
since June, 2004	PhD thesis at the Gene Center of the University of Munich in the research group of Prof. Dr. Patrick Cramer (Gene transcription and gene regulation) <i>Topic: Structural and functional studies of the RNA polymerase II elongation complex</i>
since 2005	Member of the self organized pilot project international graduate program “Protein Dynamics in Health and Disease” as part of the Elite Network of Bavaria (ENB)
December, 2007	Winner of the Römer prize from the Department of Chemistry and Biochemistry, Ludwig-Maximilians-University Munich in the category dissertation
May, 2008	PhD thesis defense

9.3 Presentations at conferences

- September 2004 **EMBO Young Investigator Programme PhD course “spotlights on current biology” in Heidelberg, Germany.**
Poster title: “Structural Biology of RNA Polymerase II - Nucleic Acid Complexes”.
- June 2006 **ESF-EMBO Symposium on Gene Transcription in Yeast in Sant Feliu de Guixols, Spain.**
Poster title: “Recognition of a DNA Damage by RNA Polymerase II”.
- Oktober 2006 **Mitgliedertagung des Elitenetzwerks Bayern (ENB) in Garching, Germany.**
Title of poster and talk: “Recognition of a DNA damage by RNA polymerase II”.
- August 2007 **24th European Crystallographic Meeting in Marrakech, Morocco.**
Poster title: “CPD Damage Recognition by Transcribing RNA Polymerase II”.
- January 2008 Invitation for: **CCP4 Study Weekend 2008 in Leeds, UK.**
Title of talk: “Structural and Functional Studies of the RNA Polymerase II Elongation Complex”.



Florian Brückner