

Genexpression

10 Jahre RNA-Polymerase-Struktur – von Transkription zu Genregulation

JASMIN F. SYDOW, DIRK KOSTREWA, DIETMAR MARTIN, PATRICK CRAMER
GENZENTRUM DER LMU MÜNCHEN

Zehn Jahre nach Entschlüsselung der RNA-Polymerase-Struktur ist der Mechanismus der mRNA-Synthese während der Transkription verstanden. Um die Regulation der Transkription zu studieren, muss die Strukturbiologie nun mit funktionaler und computerbasierter Genomik kombiniert werden.

Ten years after the determination of the RNA polymerase structure we understand the mechanism of mRNA synthesis during transcription. The regulation of transcription must now be studied with a combination of structural biology and functional and computational genomics.

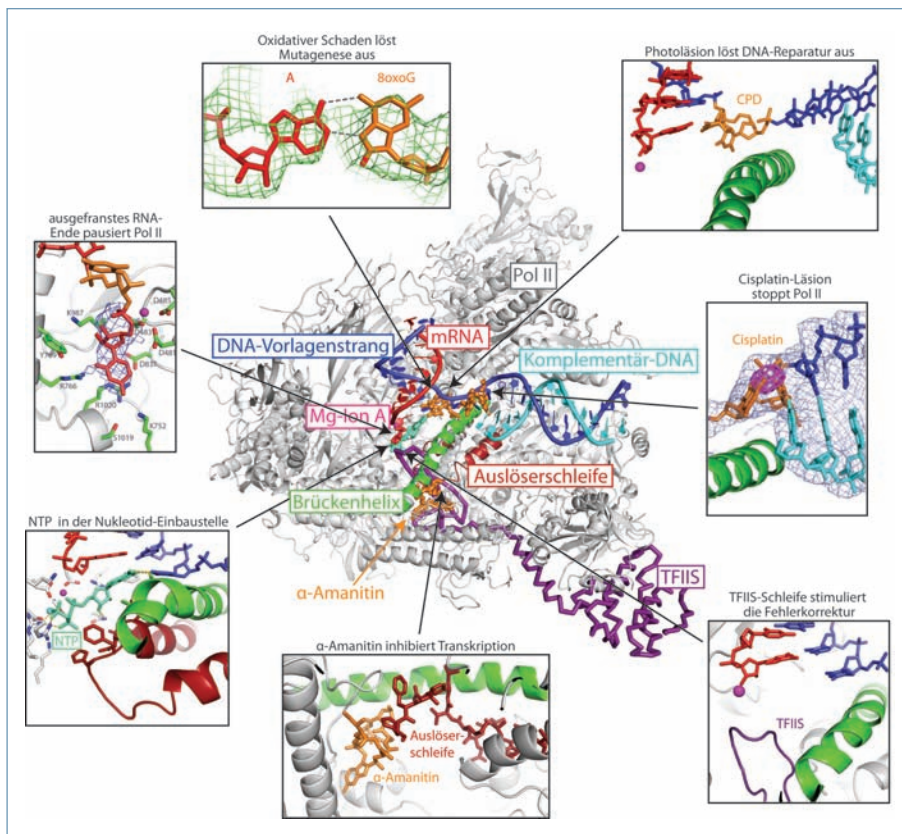
Die Transkription ist der erste Schritt der Genexpression und wird umfassend reguliert. Während der Transkription synthetisiert das Enzym RNA-Polymerase II (Pol II) die mRNA. Pol II steht im Zentrum einer dynamischen

Multiprotein-Maschinerie, deren Zusammensetzung variiert. Zur Promotor-Erkennung assemblieren Initiationsfaktoren mit Pol II. Ein Multiproteinkomplex, der Mediator, wandelt regulatorische Information von Dutzenden

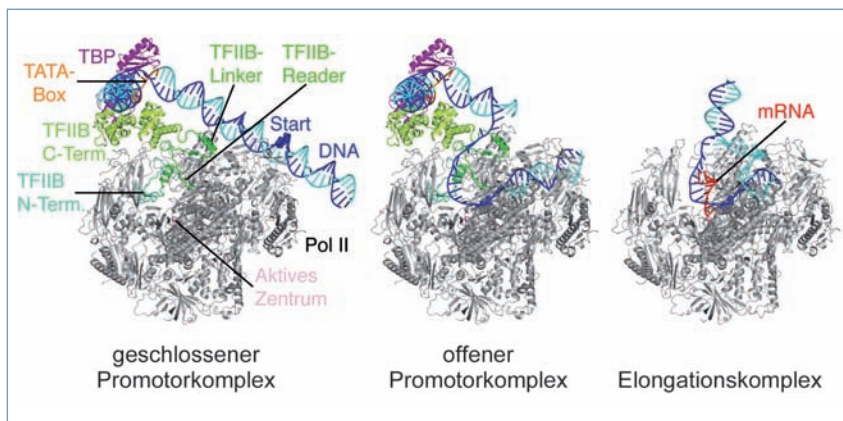
den Aktivatoren und Repressoren um. Weitere Proteinkomplexe binden Pol II während der Elongation, um die Prozessierung der RNA und Änderungen in der Chromatinstruktur zu bewirken.

Die Elongation im Film

Der Startschuss für eine detailliertere Untersuchung der Transkription erfolgte vor zehn Jahren, als die Kristallstruktur des Pol II-Kernenzym, das aus zehn Untereinheiten besteht und knapp ein halbes Mega-Dalton Molekulargewicht aufweist, gelöst wurde [1–2]. Ebenso wurde die Struktur eines ersten Elongationskomplexes beschrieben, der aus Pol II, DNA und RNA bestand [3]. Diese Strukturen offenbarten die Architektur der Pol II; am Boden einer zentralen Spalte des Enzyms liegt das aktive Zentrum, das DNA und RNA bindet (**Abb. 1**). Diese Durchbrüche im Verständnis der molekularen Grundlagen eukaryotischer Transkription waren maßgebend für die Verleihung des Chemie-Nobelpreises 2006 an Prof. Dr. Roger D. Kornberg, Stanford University. In den folgenden Jahren wurden zahlreiche Strukturen funktionaler Komplexe der RNA-Polymerase aufgeklärt [3–7], die zusammen mit funktionellen Daten zu einem ersten Film der Transkriptions-Elongation



◀ **Abb. 1:** Aspekte der Transkriptions-Elongation. Pol II fügt während dieser wiederholt ein NTP (blau-grün) zum wachsenden 3'-Ende der mRNA hinzu, wobei sich die Auslöserschleife (dunkelrot) über dem NTP-Substrat schließt, das sich in der Nukleotid-Einbaustelle befindet [5]. Das Schließen der Auslöserschleife wird durch das Gift α-Amanitin (orange) blockiert, was zu einer Hemmung der Transkription führt [6]. Entstehende Fehler in der RNA können mithilfe von TFIS (violett) effizienter korrigiert werden [4, 9]. Das hierzu notwendige Pausieren von Pol II kann durch ein Ausfransen des RNA-Endes erreicht werden [10]. Schäden in der DNA (CPD-Schaden, Cisplatin-Läsion, orange) können zu einer Blockade der Transkription und zum Auslösen von DNA-Reparatur [11, 12] oder zu transkriptioneller Mutagenese führen (8oxoG, orange) [13].



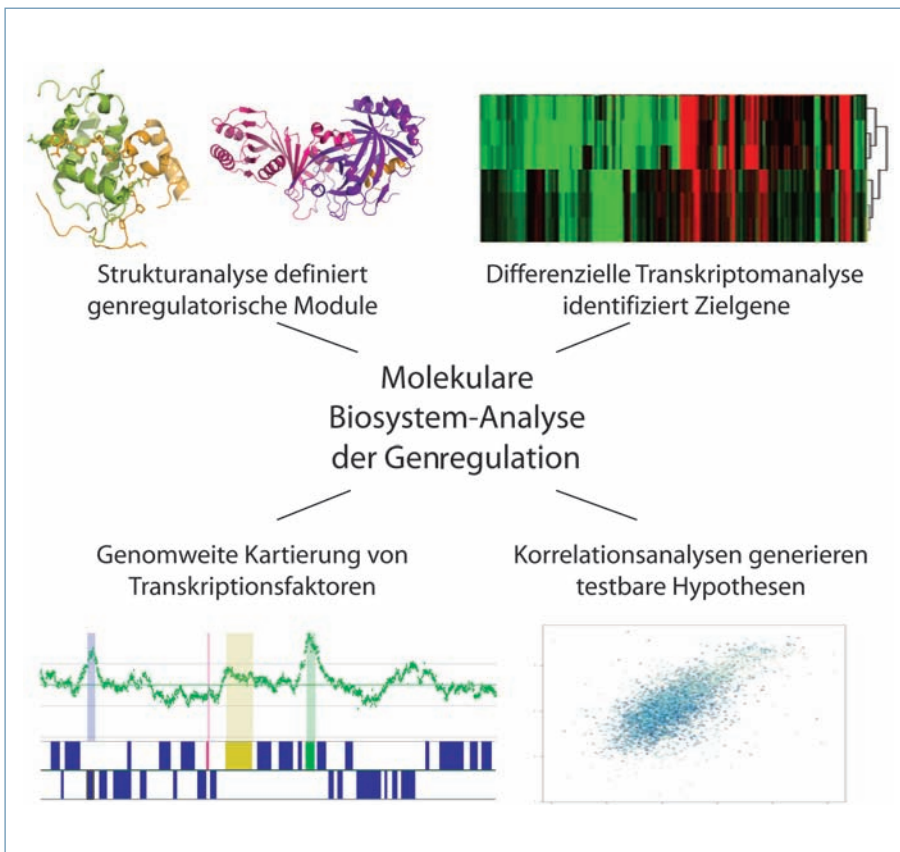
▲ **Abb. 2:** Modell der Transkriptions-Initiation. Bei der Initiation bindet TBP an die TATA-Box auf der DNA, TFIIIB bindet mit seiner C-terminalen Domäne an TBP und an die DNA und mit seiner N-terminalen Domäne an Pol II. Die doppelsträngige DNA bindet über der zentralen Spalte der Pol II, die das aktive Zentrum umfasst. Damit ist der minimale geschlossene Promotorkomplex vollständig [15]. Die doppelsträngige DNA wird dann mithilfe des TFIIIB-Linkers aufgeschmolzen, und der einzelsträngige DNA-Vorlagenstrang wird ins aktive Zentrum gefädelt, wo der TFIIIB-Reader dabei hilft, die Startstelle des Gens zu erkennen (offener Promotorkomplex) [15]. Danach beginnt die Synthese der mRNA. Die wachsende RNA verdrängt TFIIIB, der Promotorkomplex zerfällt und Pol II geht in die Elongationsphase über (Elongationskomplex) [16].

geführt haben ([8], www.LMB.uni-muenchen.de/cramer/PR-material/index.htm).

Während der Elongation durchläuft Pol II wiederholt den Zyklus der Nukleotid-Addition. Dieser beginnt mit der Bindung eines Nukleosidtriphosphats (NTP). Die darauffolgende katalytische Addition eines Nukleotids an das wachsende 3'-Ende der mRNA führt zur Abspaltung eines Pyrophosphats, die sich anschließende Translokation der DNA setzt die NTP-Bindungsstelle wieder frei, und der Zyklus beginnt von Neuem. Eine wichtige Rolle spielt die sogenannte Auslöserschleife, die sich über dem NTP-Substrat schließt, um die katalytische Konformation einzustellen (**Abb. 1**, [5]). Das NTP bindet zwei katalytische Magnesiumionen, Ion A ist ständig im aktiven Zentrum gebunden, wohingegen Ion B an den Triphosphatteil des NTPs bindet und gemeinsam mit diesem eintritt. Der Translokationsmechanismus bleibt schlecht verstanden. Unerwartete Einsichten wurden allerdings mit Aufklärung der Transkriptionshemmung durch α -Amanitin, dem Gift des grünen Knollenblätterpilzes, gewonnen (**Abb. 1**). Das zyklische Peptid α -Amanitin stabilisiert Pol II in einem Zustand, der als Translokations-Intermediat beschrieben wurde [6]. Dabei befindet sich die DNA-Base im Vorlagenstrang beim Eintreten in das aktive Zentrum in einem Zwischenzustand oberhalb der Brückenhelix. Die Brückenhelix selbst ist leicht verschoben und blo-

ckiert die eigentliche Position der DNA-Base, die als Vorlage für den nächsten Einbauschritt dient. Eine ebenso neue keilartige Konformation der Auslöserschleife an der verschobenen Brückenhelix trägt zum Mechanismus der sogenannten „Brown'schen Ratsche“ bei. Dieser Mechanismus beinhaltet ein Schwingen der Auslöserschleife zwischen einer offenen, verkeilten und geschlossenen Position.

In den letzten Jahren wurden so auch die Mechanismen der RNA-Fehlerkorrektur und der Erkennung von DNA-Schäden untersucht. Pol II zeigt eine RNA-Spaltungsaktivität, die durch den Faktor TFIIIS stimuliert wird, der durch eine Pore bis zum aktiven Zentrum vordringt (**Abb. 1**, [4, 9]). Im Falle falsch eingebauter Nukleotide wird die Fehlerkorrektur aktiviert. Die Polymerase stockt zunächst, bevor sie um ein Nukleotid zurückläuft und ein Dinukleotid vom wachsenden Ende der RNA entfernt und so den Fehler beseitigt. Die hohe Zuverlässigkeit der Pol II für den Einbau von korrekten Nukleotiden ist entscheidend, um die Produktion von fehlerhaften mRNA-Ketten zu vermeiden. Dabei benutzt Pol II unterschiedliche Strategien für die Erkennung und Entfernung von fehlgepaarten Nukleotiden [10]. Eine Ursache für inkorrekten Nukleotideinbau in der wachsenden mRNA-Kette können auch im DNA-Vorlagenstrang vorhandene Schäden sein. DNA wird durch ultraviolettes Licht, chemische Agenzien oder oxidati-



▲ **Abb. 3:** Molekulare Biosystemanalyse der Transkriptionsregulation. Genregulatorische Submodule von Proteinkomplexen wie dem Mediator-Komplex werden strukturbiochemisch identifiziert und aufgeklärt (oben links) [17, 18, 20]. Zielgene der regulatorischen Submodule werden durch eine differenzielle Transkriptomanalyse identifiziert (oben rechts) [17, 20]. Transkriptionskomplexe werden genomweit kartiert mittels Chromatin-Immuno-Präzipitation und Tiling-Mikroarrays (unten links) [19]. Korrelationsanalysen erlauben die Erstellung testbarer Hypothesen über regulatorische Netzwerke (unten rechts, unpublizierte Daten).

ven Stress geschädigt. Die dabei entstehenden Läsionen können die Transkription blockieren und einen Transkriptions-gekoppelten Reparaturmechanismus auslösen. Struktur-Funktionsanalysen von Elongationskomplexen mit großen Läsionen konnten den ersten Schritt des Reparaturmechanismus aufklären (**Abb. 1**, [11, 12]). Kleine Schäden wie etwa 8-Oxoguanin (8oxoG) können hingegen transkriptionelle Mutagenese auslösen (**Abb. 1**, [13]). Dabei entkommt ein fälschlich eingebautes Adenin der Fehlerkorrektur.

Ein Modell der Initiation

Bevor ein Gen in mRNA transkribiert werden kann, muss es erkannt werden. Diese Rolle übernimmt der Promotor am Beginn eines Gens sowie Initiationsfaktoren, die Pol II an den Promotor rekrutieren und an der Startstelle der Transkription positionieren. Die Initiation ist in eukaryotischen Zellen ein stark regulierter und sehr komplexer Vorgang. Das nachfolgend beschriebene Modell des

Initiationsprozesses beruht auf einer gerade erst publizierten Schlüsselstruktur [14, 15]. Viele Promotoren enthalten ungefähr 30 Basenpaare vor der Transkriptions-Startstelle eine DNA-Sequenz: die TATA-Box. An diese bindet der Transkriptionsfaktor TFIID, der das TATA-Box-bindende Protein (TBP) und weitere Faktoren enthält. TBP knickt die DNA an dieser Stelle rechtwinklig ab. Der Initiationsfaktor TFIIB bindet mit seiner C-terminalen Domäne an TBP und zusätzlich an DNA-Sequenzen, die die TATA-Box flankieren. Mit seiner N-terminalen Domäne bindet TFIIB an Pol II und positioniert damit die Polymerase in der Nähe der Transkriptions-Startstelle. Dieser Komplex zwischen DNA, TBP, TFIIB und Pol II wird minimaler Initiationskomplex genannt, für den nun ein dreidimensionales Modell vorliegt (**Abb. 2**). In Archaea genügt dieser minimale Komplex für die Transkriptions-Initiation, in Eukaryoten hingegen werden weitere Initiationsfaktoren, TFIIE, -F und -H, benötigt.

Pol II ist alleine nicht in der Lage, die Transkriptions-Startstelle zu erkennen, sondern ist auf weitere Hilfe durch TFIIB angewiesen. Zunächst muss die doppelsträngige DNA in Einzelstränge aufgeschmolzen werden. Dabei hilft TFIIB mit dem B-Linker, der aus einer α -Helix und einem kurzen β -Strang besteht, bei der Öffnung der DNA vor der Transkriptions-Startstelle [15]. Die einzelsträngige Templat-DNA kann nun in das aktive Zentrum der Pol II eingefädelt werden. Nun muss noch die Transkriptions-Startstelle erkannt werden. Hierbei hilft wieder TFIIB, die Erkennungsbasen des sogenannten Initiator-Elements an der Transkriptions-Startstelle auszulesen und damit den DNA-Strang im aktiven Zentrum zu platzieren. TFIIB verwendet hierzu den B-Reader, der aus einer Helix und einer flexiblen Schleife besteht. Nun werden erste kurze RNA-Stücke synthetisiert, die aber häufig wieder von der DNA-Vorlage abfallen. Der wachsende RNA-Strang stößt aber immer wieder an den noch gebundenen Faktor TFIIB und trägt dazu bei, dass TFIIB seine Bindung an Pol II verliert und vom Promotorkomplex abfällt. Damit kann Pol II dem Promotor entkommen und von der Initiations- in die Elongationsphase übertreten [15].

Eine Vision der Regulation

Es ist genau dieser komplizierte Initiationsprozess, der vielfältig reguliert wird, um die zelluläre Genexpression zu steuern. Dabei sind Proteinkomplexe wie der Mediator-Komplex wichtig, welche Signale von regulatorischen Faktoren integrieren. Der Mediator überträgt dann die Signale auf den Pol II-Initiationskomplex, um die Transkription reguliert zu starten. Wie dies genau geschieht, ist unbekannt. Der Mediator besteht aus 25 Proteinen und hat ein Molekulargewicht von über einem Mega-Dalton. Eine Kombination von Strukturanalyse und differenzieller Genexpressionsanalyse konnte zeigen, dass der Mediator Submodule enthält, die an der Transkriptionsregulation bestimmter Gruppen von Genen beteiligt sind (**Abb. 3**). So erzeugen Deletionen bestimmter Mediator-Submodule charakteristisch veränderte Genexpressionsprofile [17, 18]. Manche der Zielgene unterschiedlicher Mediator-Submodule sind gleich, aber es existieren auch viele Submodul-spezifische Zielgene.

Ziel der Transkriptionsforschung ist es, die Regulationsprinzipien nicht nur molekular und mechanistisch, sondern auch genomweit und quantitativ in der Zelle aufzuklären. Es bedarf zum einen hochauflösender Analysen

auf Ebene der mRNA-Quantifizierung. Dabei reicht es nicht aus, die Mengen aller mRNAs zu messen, da diese sowohl von der Transkriptionsrate als auch von der mRNA-Abbaurate bestimmt werden. Mit neuen Methoden muss ermittelt werden, wie hoch die mRNA-Syntheseraten tatsächlich sind. Des Weiteren müssen die zugrunde liegenden DNA-Protein-Wechselwirkungen genomweit kartiert werden. Mittels Chromatin-Immunopräzipitation und hochauflösender Analyse auf Tiling-Mikroarrays, deren DNA-Sonden das ganze Genom abdecken, werden genomweite Bindprofile von Pol II und den assoziierten Faktoren erhalten [19]. Kombiniert mit kinetischen Daten unter verschiedenen Wachstumsbedingungen (z. B. Stressbedingungen) könnten so primäre und sekundäre Zielgene regulatorischer Transkriptionsfaktoren unterschieden werden. Die Auswertung solcher genomweiter Daten erfolgt durch Korrelationsanalysen, mit deren Hilfe experimentell testbare Hypothesen generiert werden.

In der Zukunft werden Struktur-Funktionsstudien *in vitro* die molekularen Wechselwirkungen aufklären, die der Genregulation zugrunde liegen. Strukturbasierte Mutationen müssen dann genutzt werden, um die Relevanz regulatorischer Mechanismen *in vivo* zu etablieren. Diese Analysen werden von einzelnen Genen auf das zelluläre System erweitert. Mithilfe der funktionalen Genomik, insbesondere der transkriptomweiten Messung von mRNA-Syntheseraten und der genomweiten Kartierung von Faktoren wird es möglich, regulatorische Netzwerke zu eta-

blieren. Durch vergleichende Analysen verschiedener Mutanten könnte so eines Tages das transkriptionelle Netzwerk einer Zelle aufgeklärt und molekular-mechanistisch beschrieben werden. Dies wird allerdings nur mit kinetischer Information gelingen. Störungen des Systems, etwa durch definierte Stressbedingungen, bewirken Änderungen in genregulatorischen Netzwerken, die in Zeitreihen verfolgt werden müssen. Die für solche Studien nötige Kombination von Strukturbiochemie mit funktionaler Genomik und computerbasierter Biologie wird molekulare Biosystemforschung genannt und erfordert interaktive institutionelle Strukturen. ■

Literatur

- [1] Cramer P, Bushnell DA, Fu J et al. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288:640–649
- [2] Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292:1863–1876
- [3] Gnatt AL, Cramer P, Fu J et al. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292:1876–1882
- [4] Kettenberger H, Armache K-J, Cramer P (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol Cell* 16:955–965
- [5] Wang D, Bushnell DA, Westover KD et al. (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127:941–954
- [6] Brueckner F, Cramer P (2008) Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat Struct Mol Biol* 15:811–818
- [7] Vassylyev DG, Vassylyeva MN, Zhang J et al. (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448:163–168
- [8] Brueckner F, Ortiz J, Cramer P (2009) A movie of the RNA polymerase nucleotide addition cycle. *Curr Opin Struct Biol* 19:294–299
- [9] Kettenberger H, Armache K-J, Cramer P (2003) Architecture of the RNA polymerase II-TFIIIS complex and implications for mRNA cleavage. *Cell* 114:347–357
- [10] Sydow JF, Brueckner F, Cheung AC et al. (2009) Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol Cell* 34:710–721
- [11] Brueckner F, Hennecke U, Carell T et al. (2007) CPD damage recognition by transcribing RNA polymerase II. *Science* 315:859–862
- [12] Damsma GE, Alt A, Brueckner F et al. (2007) Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* 14:1127–1133
- [13] Damsma GE, Cramer P (2009) Molecular basis of transcriptional mutagenesis at 8-oxoguanine. *J Biol Chem* 284:31658–31663
- [14] Nikolov DB, Chen H, Hoffman A et al. (1995) Crystal structure of a TFIIIB-TBP-TATA-element ternary complex. *Nature* 377:119–128
- [15] Kostrewa D, Zeller ME, Armache KJ et al. (2009) RNA polymerase II-TFIIIB structure and mechanism of transcription initiation. *Nature* 462:323–330
- [16] Andrecka J, Treutlein B, Arcusa MA et al. (2009) Nano positioning system reveals the course of upstream and non-template DNA within the RNA polymerase II elongation complex. *Nucleic Acids Res* 37:5803–5809
- [17] Koschubs T, Seizl M, Larivière L et al. (2009) Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J* 28:69–80
- [18] Larivière L, Geiger S, Hoepfner S et al. (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat Struct Mol Biol* 13:895–901
- [19] Jasiak AJ, Hartmann H, Karakasili E et al. (2008) Genome-associated RNA polymerase II includes the dissociable Rpb4/7 subcomplex. *J Biol Chem* 283:26423–26427
- [20] Larivière L, Seizl M, van Wageningen S et al. (2008) Structure-system correlation identifies a gene regulatory Mediator submodule. *Genes Dev* 22:872–877

Korrespondenzadresse:

Prof. Dr. Patrick Cramer
 Genzentrum der Ludwig-Maximilians-Universität
 München
 Feodor-Lynen-Straße 25
 D-81377 München
 Tel.: 089-2180-76965
 Fax: 089-2180-76999
 cramer@LMB.uni-muenchen.de
 www.LMB.uni-muenchen.de/cramer

AUTOREN



Jasmin F. Sydow

Jahrgang 1978. 1998–2004 Biologiestudium an der LMU München. 2005–2007 Forschungsprojekt bei Prof. Dr. Coll am IBMB in Barcelona. 2005–2009 Promotion bei Prof. Dr. Cramer am Genzentrum der LMU München.



Dirk Kostrewa

Jahrgang 1961. 1980–1987 Biochemiestudium an der FU Berlin. 1987–1991 Promotion bei Prof. Dr. Saenger an der FU Berlin. 1991–1994 Postdoktorand bei Prof. Dr. Winkler an der Hoffmann-La Roche AG Basel. 1994–1999 Senior Scientist an der Hoffmann-La Roche AG Basel. 2000–2007 Staff Scientist am Paul Scherrer Institut in Villigen. Seit 2007 Staff Scientist am Genzentrum der LMU München.



Dietmar Martin

Jahrgang 1970. 1991–1996 Biologiestudium an den Universitäten Marburg und Oxford. 1996–1999 Promotion bei Prof. Dr. Reinhold Hurek am Max-Planck-Institut für terrestrische Mikrobiologie, Marburg. 2000–2001 Postdoktorand an der Universität Bremen. 2001–2008 Postdoktorand bei Prof. Dr. Hall, Universität Basel. Seit 2008 Lecturer am Genzentrum der LMU München.



Patrick Cramer

Jahrgang 1969. 1989–1994 Chemiestudium an den Universitäten Stuttgart, Heidelberg, Bristol, Cambridge. 1995–1998 Promotion bei Prof. Dr. Müller am EMBL in Grenoble, Frankreich. 1999–2001 Postdoktorand bei Prof. Dr. Kornberg an der Stanford University, USA. Seit 2001 Professor für Biochemie an der LMU München. Seit 2004 Leiter des Genzentrums der LMU München.