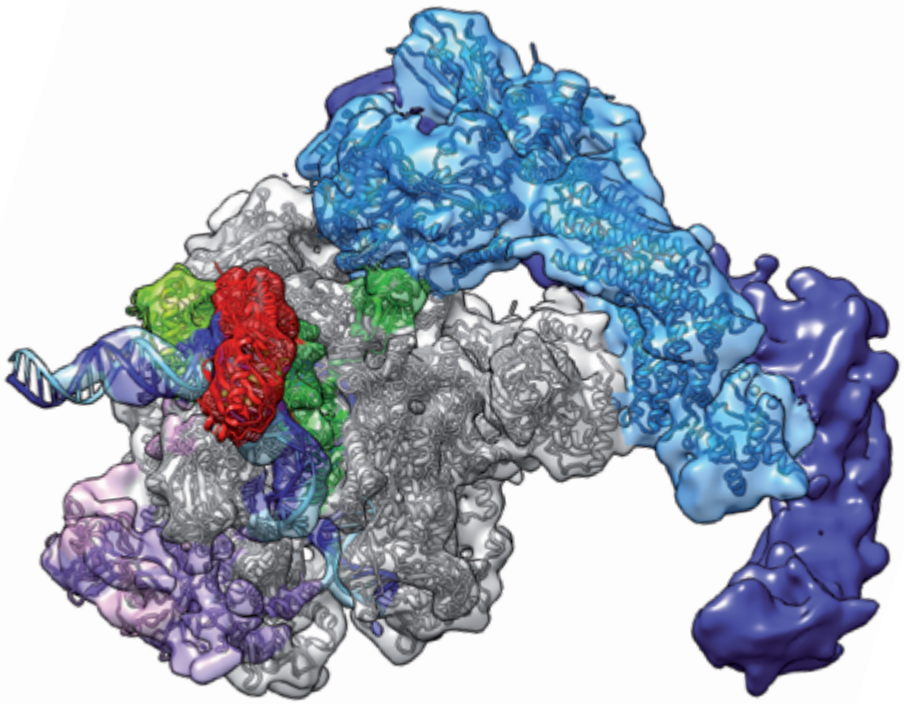


Fig.1. Architecture of the promoter assembly for regulated transcription initiation as revealed by cryo-electron microscopy. The electron density is transparent, revealing the polymerase (silver), general transcription factors (red, green, purple), and the core Mediator head module (blue). The remainder of the core Mediator corresponds to the middle module and is shown in dark blue. Promoter DNA is depicted as a ribbon model in blue and cyan, and the initial RNA transcript is in red.
(Taken from: Plaschka C, Larivière L, Wenzeck L, Seizl M, Hemann M, Tegunov D, Petrotchenko EV, Borchers CH, Baumeister W, Herzog F, Villa E, Cramer P: Architecture of the RNA polymerase II–Mediator core initiation complex. Nature 518, 376–380 (2015))

Starting to get the message

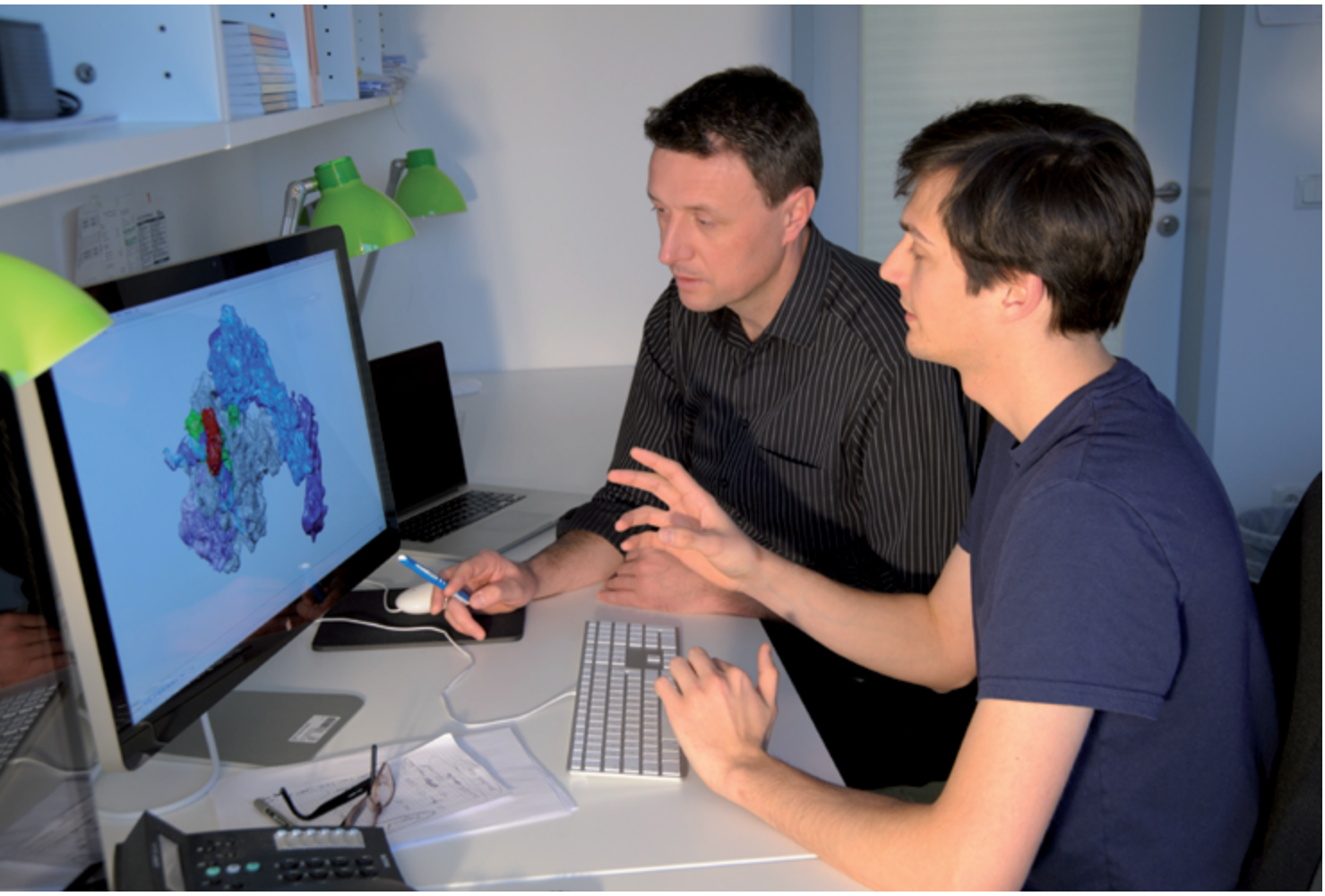
Structural analysis of a 31-protein assembly on DNA elucidates transcription initiation and regulation

Clemens Plaschka and Patrick Cramer
Department of Molecular Biology

Gene transcription is a fundamental process in living cells. It is the first step in the expression of the genome, and its regulation underlies cell differentiation and organism development. During transcription, the enzyme RNA polymerase II copies genes to produce messenger-RNA, which serves as the template for protein synthesis. Transcription is mainly regulated at the initiation stage, when the polymerase assembles on promoter DNA with the general transcription factors and the Mediator complex. Mediator receives signals from regulatory factors and assists in the assembly of the transcription initiation machinery at gene promoters. Because of the large size and complexity of the promoter assembly, its structural organization remained poorly understood.

In an interdisciplinary team effort, we have now obtained a detailed view of the promoter assembly from yeast and could outline how regulated transcription initiation occurs [1]. We studied a 31-subunit assembly that contained a minimal transcription initiation complex and the essential core of Mediator. Using cryo-electron microscopy and protein crosslinking analysis, we determined its architecture at 9.7 Å resolution (Fig. 1). This breakthrough was possible only after we had established how to prepare a stable and functionally active core Mediator.

Mediator poses a formidable challenge to the structural biologist. It contains 25 subunits and has a molecular weight of 1.4 megadaltons. Furthermore, it is modular, flexible, and post-translationally modified. After a decade of work, we learned that the essential core of Mediator could be prepared by co-expression of 15 selected subunits in bacteria. In a bottom-up approach, Sonja Baumli, Tobias Koschubs, Martin Seizl, and Laurent Lariviére had added more and more subunits. Along the way, we revealed the structure of stable



Patrick Cramer (left) and Clemens Plaschka

Mediator subassemblies by X-ray crystallography [2], including a large subassembly, called the head module [3].

When initial preparations of core Mediator became available, we elucidated its subunit architecture by protein crosslinking and mass spectrometry together with the laboratory of Christoph Borchers (University of Victoria, Canada). To facilitate structural studies, we further optimized the core Mediator preparation protocol. With assistance from Larissa Wenzeck (Ludwig-Maximilians-Universität München) we achieved preparation of core Mediator in large amounts and with very high purity. Recombinant core Mediator was functionally active, demonstrating its integrity and relevance for understanding regulated transcription initiation.

The core Mediator bound to polymerase, yet the resulting complex was not stable enough to be analyzed by electron microscopy. Fortunately, Sarah Sainsbury, Wolfgang Mühlbacher, and Merle Hantsche had just succeeded to reconstitute a minimal Pol II initiation complex that contained the polymerase and the general transcription factors TBP, TFIIB, and TFIIF [4]. This minimal initiation complex bound more tightly to core Mediator and enabled reconstitution of a 31-subunit assembly with a mass of 1.2 megadaltons.

This finally opened the door to structural studies. We pursued cryo-electron microscopy in collaboration with Elizabeth Villa in the laboratry of Wolfgang Baumeister (Max Planck Institute of Biochemistry, Martinsried) and, in parallel, we worked with Franz Herzog (Ludwig-Maximilians-Universität München) to carry out crosslinking analysis of the entire assembly. Soon the pieces came together. Like a three-dimensional puzzle, we were able to arrange available crystal structures of subunits and subassemblies. Eventually, we integrated all available information into one architectural model of the promoter assembly.

The results showed how the general transcription factors cooperate with the polymerase to bind promoter DNA and to position the start site of transcription. They also suggested how Mediator functions to switch transcription on: The Mediator head module contacts regions of the polymerase and TFIIB and stabilizes the promoter assembly. Other regions of Mediator may control polymerase conformation and can bridge to regulatory factors bound to upstream DNA, allowing for different transcriptional outputs.

The work also enabled us to imagine how even larger promoter assemblies look like (Fig.2). This became possible by comparing our structure with electron microscopic results obtained with a related complex by the laboratory of Eva Nogales (University of California, USA). Our colleagues had recently reported the architecture of a human promoter assembly that

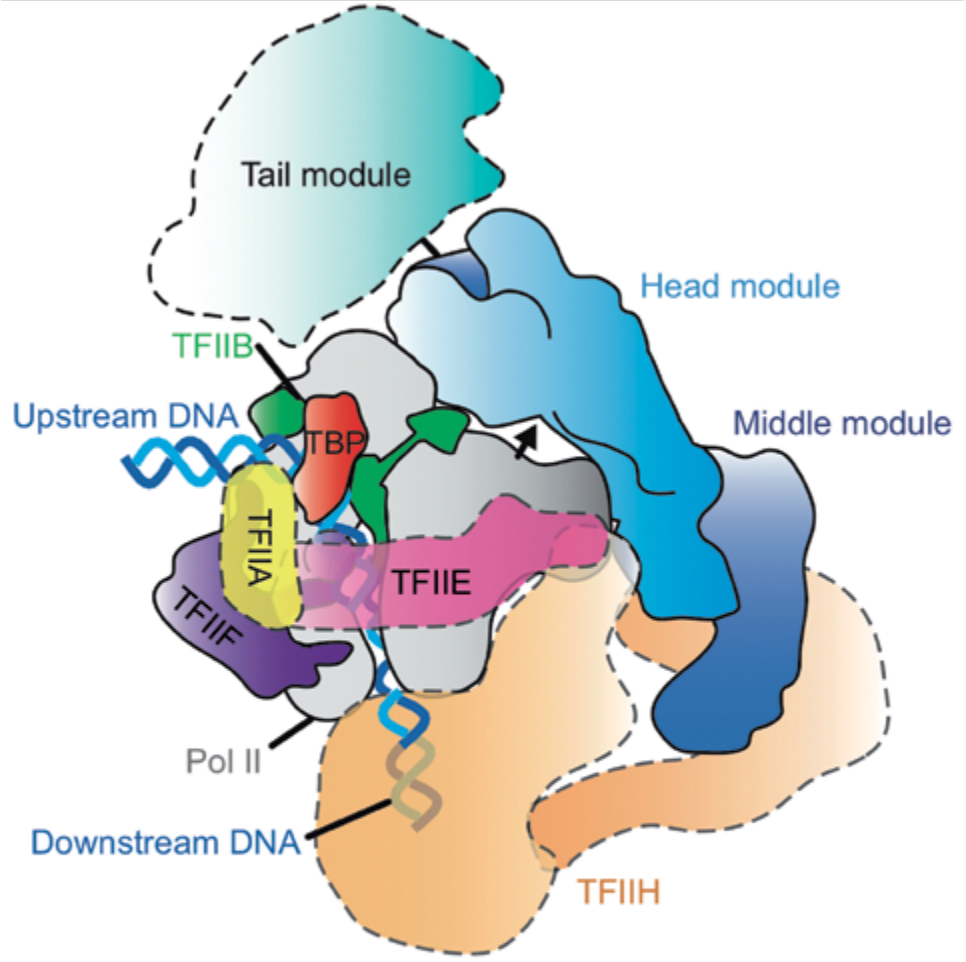


Fig. 2. Model of the complete and conserved promoter assembly. The locations of additional general transcription factors (yellow, magenta, orange) and the Mediator tail module (cyan) were inferred from work of the Nogales and Asturias groups. The resulting model suggests a complete and conserved promoter assembly and identifies important regulatory domains of Mediator.

lacked Mediator but instead contained the general factor TFIIH, which was absent in our assembly. A comparison with their work suggested how Mediator and TFIIH cooperate for successful transcription initiation and how this may result in release of the polymerase into the gene body for elongation of the messenger-RNA chain. It also revealed that the transcription initiation complex is highly conserved throughout eukaryotes, showing that studies of yeast complexes are directly relevant for understanding human gene transcription and its deregulation in disease.

Many open questions on transcription initiation remain. Can we obtain higher resolution and reveal more mechanistic details? What is the structure and location of the remaining parts of Mediator? How do these influence transcriptional output? What is the structure of the 10 subunit TFIIH complex? How does TFIIH cooperate with the polymerase and the general factor TFIIE to open promoter DNA and form a transcription bubble? What is the structure of TFIID, and how does this additional general factor contribute to promoter recognition? Finally, how does the initiation complex interact with the first nucleosome it encounters near the transcription start site? Answers to these questions will require strategies to prepare transient multiprotein complexes, the integration of available structural biology techniques, and complementary functional investigations *in vitro* and *in vivo* using system-wide approaches.

References

1. Plaschka C, Larivière L, Wenzeck L, Seizl M, Hemann M, Tegunov D, Petrotchenko EV, Borchers CH, Baumeister W, Herzog F, Villa E, Cramer P: Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature,* doi:10.1038/nature14229 (2015).

2. Larivière L, Seizl M, Cramer P: A structural perspective on Mediator function. *Curr Opin Cell Biol* **3**, 305-313 (2012).

3. Larivière L*, Plaschka C*, Seizl M, Wenzeck L, Kurth F, Cramer P: Structure of the Mediator head module. *Nature* **492**, 448-451 (2012). *(*These authors contributed equally)*

4. Mühlbacher W, Sainsbury S, Hemann M, Hantsche M, Neyer S, Herzog F, Cramer P: Conserved architecture of the core RNA polymerase II initiation complex. *Nature Commun* **5**, 4310 (2014).