Building transcriptional complexes

Alan C. M. Cheung[1,2]

1. University College London, Institute of Structural and Molecular Biology, Gower Street, London WC1E 6BT
2. Birkbeck College, Institute of Structural and Molecular Biology, Malet St, London, WC1E 7HX

The intertwined structure of the TFIID subcomplex Taf5-Taf6-Taf9 is dependent on chaperonin CCT for its assembly and subsequent integration into the TFIID general transcription factor.

Large protein complexes are ubiquitous in eukaryotic transcription, functioning in all phases of the transcription cycle, but little is known about their assembly pathways. Protein complexes require ordered pathways for their assembly[1], as misassembly can result in cellular stress and cytotoxicity[2]. In this issue, Antonova et al. reveal a key step in the biogenesis of the TFIID complex that requires the CCT chaperonin complex (Figure 1).

The TFIID complex is a general transcription factor required for recognition of the core promoter and subsequent nucleation of the RNA Polymerase II containing Pre-Initiation Complex. It is related to the SAGA complex which stimulates transcription by catalyzing histone modifications, and also by recruitment of the TATA-binding protein (TBP)[3]. Their relationship is due to their sharing of subunits, which is a recurrent theme amongst transcription-related complexes[4]. In yeast, TFIID consists of 14 subunits (TBP plus TAFs 1-13) and shares five of these TAFs with the 19-subunit SAGA complex (specifically, TAFs 5, 6, 9, 10 and 12) whereas mammalian TFIID and SAGA only share three TAFs due to duplication of the TAF5 and TAF6 genes, resulting in paralogs TAF5L and TAF6L that specifically associate with SAGA[5,6], whereas TAF5 and TAF6 are incorporated into TFIID.

In this issue, the Timmers and Berger groups determined the crystal structure of a human truncated TAF5-TAF6-TAF9 subcomplex of TFIID (hereafter referred to as TAF569). This subcomplex adopts a compact triangular architecture where the histone-folds of TAF6 and TAF9 heterodimerize, and the resultant dimer is then sandwiched between the N-terminal and WD40 domains of TAF5. A C-terminal extension of TAF9 forms an extended loop that wraps around the TAF5 WD40 domain, anchoring the complex together. The visualization of this striking interface in the crystal structure allowed the design of targeted mutations to disrupt it, and the resultant effects on TFIID and SAGA assembly monitored by co-immunoprecipitation followed by quantitative mass spectrometry (coIP-qMS) in HeLa cells. These experiments showed that this TAF5-TAF9 interface was critical for TFIID assembly, as its disruption by mutating either TAF5 or TAF9 consistently abrogated TFIID formation. As TAF9 is shared with SAGA (and presumably forms a similar interface with the WD40 domain from the SAGA-specific paralog TAF5L), it would be reasonable to expect that the same mutations in TAF9 would also affect SAGA, but limited effects were observed on SAGA assembly. Thus the human TFIID-specific TAF569 subcomplex may exhibit important differences in conformation when compared to its SAGA counterpart of TAF5L-TAF6L-TAF9, and particularly in the interface between TAF5L and TAF9. Alternatively, SAGA may make additional interactions that compensate for a loss of contact between TAF5L and TAF9.

Strikingly, one of the TAF5 mutants was found to be distributed across the cytoplasm, in contrast to the predominantly nuclear localisation of WT TAF5. coIP-qMS analysis of the same mutant showed that it significantly enriched all eight subunits of the chaperonin CCT complex, a cylindrical complex similar to bacterial GroEL that binds and sequesters unfolded proteins in its central cavity and stimulates their folding in an ATP dependent manner[7]. CCT assists folding of about 10% of the proteome, and includes many WD40-repeat containing proteins amongst its target substrates[8], making TAF5 a likely target of its activity. Antonova et al. probed the interactions between CCT and TFIID and found that WD40 domain of TAF5 (and not its N-terminal domain) was indeed the key target of CCT within TFIID. Similarly, the WD40 domain of SAGA-specific TAF5L also associated with CCT. Pulse-chase experiments revealed that newly synthesized WT TAF5 would progressively but only transiently bind CCT, representing successful folding and release of TAF5, whereas mutant TAF5 would remain sequestered by CCT and could not be incorporated into TFIID. Transient coexpression experiments then showed that the release of TAF5 from CCT is dependent on the presence of TAF6-TAF9. Collectively, these results lead the authors to postulate a model, where CCT acts as an essential step in TFIID assembly, by enabling folding of the TAF5 WD40 domain, and sequestering it until it can associate with the TAF6-TAF9 heterodimer, forming the TAF569 subcomplex for further assembly into holo-TFIID (Figure 1).

How does the TAF6-TAF9 dimer enable TAF5 release from CCT? As the folding cavity of CCT is not completely sealed, it leaves a pore that can extrude extended stretches of polypeptide, allowing CCT to fold individual domains within a multi-domain substrate[9]. This may allow CCT to fold the WD40 domain of TAF5, whilst leaving its N-terminal domain freely accessible for TAF6-TAF9 binding. The authors suggest that as the folded WD40 domain emerges from CCT, its interaction with TAF6-TAF9 may prevent its reassociation to CCT, and enable release of the complete subcomplex for holo-TFIID assembly.

Given their sharing of TAFs, a key question raised by this work is: do the principles of TFIID biogenesis also apply to SAGA? The study by Antonova et al. points toward both similarities and differences in their respective assembly pathways. Although mutations of TAF9 selectively disrupted TFIID over SAGA, the CCT chaperonin was found to interact with the SAGA-specific paralog TAF5L, so some overlap in their biogenesis mechanisms seem likely. Although TAF569 is unique to TFIID, its paralogous counterpart in SAGA is likely to have a similar conformation. This is supported by the recent Cryo-EM reconstruction of TFIID bound to promoter DNA[10] which showed that the shared TAFs between TFIID and SAGA probably have the same overall conformation in both complexes. However, as shown by the TAF9 mutations, important differences may lie in the fine molecular details. Nevertheless, the discovery of CCT in TFIID biogenesis may open a new frontier in understanding how these large macromolecular machines are built.

1. Marsh, J. A. *et al.* Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell* **153,** 461–70 (2013).
2. Ellis, R. J. Protein misassembly: macromolecular crowding and molecular chaperones. *Adv. Exp. Med. Biol.* **594,** 1–13 (2007).
3. Hahn, S. & Young, E. T. Transcriptional Regulation in Saccharomyces cerevisiae: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. *Genetics* **189,** 705–736 (2011).
4. Helmlinger, D. & Tora, L. Sharing the SAGA. *Trends Biochem. Sci.* **42,** 850–861 (2017).
5. Taatjes, D. J., Marr, M. T. & Tjian, R. Regulatory diversity among metazoan co-activator complexes. *Nat. Rev. Mol. Cell Biol.* **5,** 403–410 (2004).
6. Spedale, G., Timmers, H. T. M. & Pijnappel, W. W. M. P. ATAC-king the complexity of SAGA during evolution. *Genes Dev.* **26,** 527–541 (2012).
7. Bueno-Carrasco, M. T. & Cuéllar, J. Mechanism and Function of the Eukaryotic Chaperonin CCT. *eLS* 1–9 (2018). doi:10.1002/9780470015902.a0028208
8. Valpuesta, J. M., Martín-Benito, J., Gómez-Puertas, P., Carrascosa, J. L. & Willison, K. R. Structure and function of a protein folding machine: The eukaryotic cytosolic chaperonin CCT. *FEBS Lett.* **529,** 11–16 (2002).
9. Rüßmann, F. *et al.* Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 21208–15 (2012).
10. Patel, A. B. *et al.* Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science (80-.).* **8872,** eaau8872 (2018).

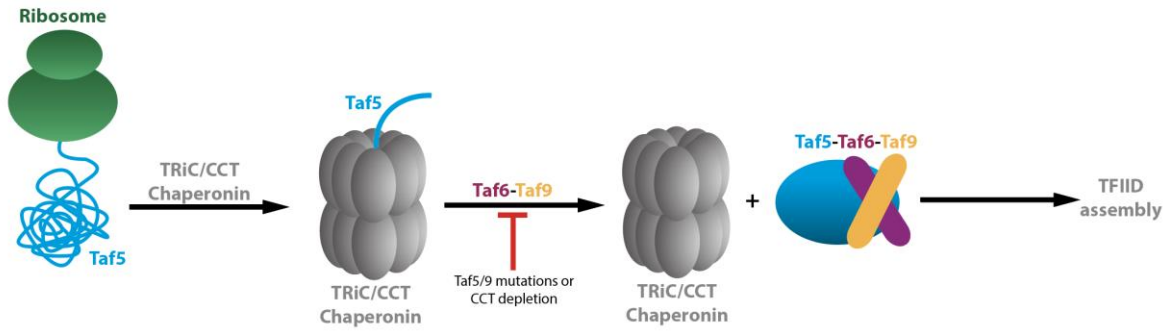

Figure 1. A model of Chaperonin CCT-assisted assembly of TFIID. Newly synthesized Taf5 is captured by the CCT chaperonin complex, enabling folding of its WD40 domain. Release of Taf5 from CCT is dependent on the recruitment of Taf6-Taf9, presumably by binding to the exposed, non-WD40 containing N-terminal region of Taf5. The release of Taf5 and the concomitant formation of the Taf5-Taf6-Taf9 subcomplex can then bind additional TAFs and TBP to form the complete TFIID complex.