

## Architecture du facteur de transcription/réparation TFIIH

Chez l'homme, l'étude de plusieurs maladies génétiques autosomiques rares, telles que le xeroderma pigmentosum (XP), le syndrome de Cockayne ou la trichothiodystrophie a mis en exergue les interconnexions entre les mécanismes de réparation et de transcription de l'ADN, deux processus cellulaires fondamentaux. Ainsi, chez certains groupes de patients atteints de xeroderma pigmentosum – une maladie caractérisée par une sensibilité accrue aux rayons ultra-violet et par des risques élevés de développer des cancers de la peau –, il a été montré que les gènes dénommés *XPB* et *XPD* sont mutés [1,2]. Ces gènes codent pour deux sous-unités du facteur de transcription/réparation TFIIH qui en comporte en tout neuf. Ce complexe protéique abrite trois activités enzymatiques distinctes, ce qui constitue une propriété unique parmi les facteurs de transcription, outre les deux activités hélicase catalysées par les protéines *XPB* et *XPD*, il possède une activité cycline kinase portée par *cdk7* [3]. Les travaux de recherche, entrepris depuis plusieurs années, ont contribué à une meilleure compréhension de la fonction de TFIIH et ont permis d'entrevoir une relation génotype-phénotype.

Il a ainsi été montré que les deux hélicases permettaient l'ouverture de l'ADN de part et d'autre de la lésion, lors de la réparation par excision-resynthèse d'un ADN endommagé (*nucleotide excision repair*, pour revue voir [4]). Cette ouverture favorise l'accès d'un complexe protéique destiné à réparer la lésion. Dans le cadre

de la réaction de transcription, le facteur TFIIH agit de concert avec cinq autres facteurs de transcription généraux (TFIIA, TFIIB, TFIID, TFIIE et TFIIIF) au sein d'un complexe de préinitiation destiné à positionner l'ARN polymérase II, elle-même responsable de la synthèse de l'ARN messager. Dans ce contexte, il a été montré que l'hélicase *XPB* de TFIIH déroule la double hélice d'ADN au niveau du promoteur pour permettre la lecture du brin codant [5]. Quant à la kinase dont on avait identifié certains substrats tels que l'ARN polymérase II ou divers activateurs de la transcription et certains récepteurs nucléaires, il semble que son action soit nécessaire pour permettre le passage de l'ARN polymérase II d'un mode d'initiation, où elle est non phosphorylée, à un mode d'élongation où elle est phosphorylée (pour revue, voir [6]).

Pour comprendre les bases moléculaires de ces divers mécanismes, et en ce qui nous concerne des maladies causées par des mutations dans les hélicases *XPB* et *XPD*, il est indispensable de connaître la structure spatiale du facteur TFIIH ainsi que son intégration dans des architectures plus vastes tels que les complexes d'initiation de la transcription ou le complexe de réparation. Ces dernières années, les structures 3D de différents éléments de ces complexes ont été résolues à l'échelle atomique. Ainsi, l'organisation spatiale de trois facteurs généraux, TBP, TFIIB et TFIIA, liés à la séquence promotrice TATA, a été déterminée [7]. De même, la structure de l'ARN polymérase II de levure a été résolue récem-

ment [8]. Des avancées considérables ont également été apportées par les structures cristallines des récepteurs nucléaires qui permettent de mieux appréhender leur mode de régulation par les ligands les concernant, mais également d'imaginer les domaines impliqués dans l'activation de la transcription [9,10]. A l'heure actuelle, les facteurs TFIID et TFIIH, probablement à cause de leur complexité, se sont montrés réticents à la cristallisation et seules des structures partielles de certaines de leurs sous-unités sont connues. Dans ce contexte, les groupes de l'IGBMC (Strasbourg, France) ont tenté d'établir une relation structure-fonction des facteurs TFIID [11,12] et TFIIH [13-15] en utilisant une approche intégrant des résultats de microscopie électronique à haute résolution, de RMN et de cristallographie par les rayons X.

Récemment, nous avons pu déterminer la structure moléculaire du facteur TFIIH par microscopie électronique à une résolution de 3,8 nm [15]. Les difficultés de production de protéines humaines ainsi que les problèmes techniques liés à la détermination de structure d'un assemblage macromoléculaire d'environ 0,5 MDa rendaient la tâche ardue. Pour mener à bien ce projet, nous avons revu le schéma de purification de TFIIH extrait de cellules HeLa en culture, mis au point des systèmes de production de TFIIH recombinant, produit des anticorps monoclonaux et utilisé les derniers développements de la microscopie électronique permettant la détermination d'une enveloppe moléculaire à partir de l'obser-

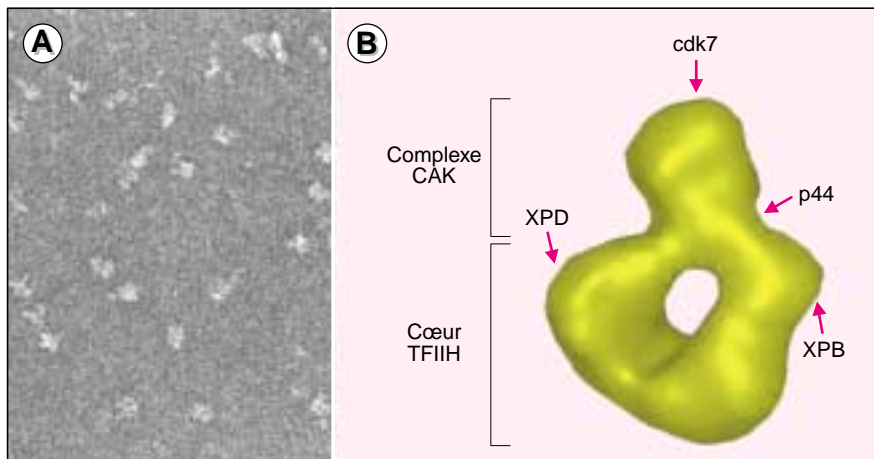
vation directe des molécules non cristallisées [16].

Le facteur TFIIH humain hautement purifié est adsorbé sur une grille de carbone, coloré à l'acétate d'uranyle et observé avec un grandissement de 45 000 fois à l'aide d'un microscope électronique. Les images moléculaires « brutes » sont fortement bruitées mais on distingue des particules légèrement allongées, réparties de façon homogène sur la grille (figure 1A). Un protocole itératif d'analyse d'image permet d'améliorer le rapport signal sur bruit par des opérations de « moyennation », d'identifier différentes vues de la molécule puis de reconstruire un modèle 3D de l'enveloppe de la molécule. Le complexe TFIIH apparaît alors comme une particule légèrement allongée et plate de dimensions 16x12,5x7,5 nm, avec une partie en forme d'anneau dont le diamètre interne d'environ 3 nm permettrait d'accommoder une molécule d'ADN double brin. Sur cet anneau est greffée une protubérance quasiment sphérique d'approximativement 5 nm de diamètre (figure 1B). Pour intéressante qu'elle soit, la connaissance de la forme de l'assem-

blage moléculaire ne nous renseigne pas sur son mode de fonctionnement. Nous avons par conséquent entrepris de localiser un certain nombre de sites fonctionnels sur cette architecture. Nos investigations ont été guidées par un ensemble de données biochimiques qui montrent que les 9 sous-unités du complexe TFIIH sont réparties en deux sous-complexes fonctionnels: le cœur-TFIIH (composé des protéines p34, p44, p52, p62 et de l'hélicase XPB) et le complexe CAK (comprenant la kinase dépendante de la cycline, cdk7, la cycline H et MAT1), la dernière sous-unité, l'hélicase XPD, pouvant être associée soit au cœur, soit au CAK selon le mode de purification employé [17]. Le cœur-TFIIH a été produit en utilisant le système d'expression de baculovirus dans des cellules d'insectes. Son étude par microscopie électronique a certes mis en évidence un polymorphisme, suggérant que la stabilité des complexes recombinants n'est pas aussi importante que celle des complexes natifs, mais a également révélé que 34 % des complexes présentaient une architecture circulaire similaire en forme et en dimensions à celles de

l'anneau observé dans le facteur entier. La comparaison du cœur-TFIIH avec le facteur entier permet de déterminer la position relative des deux entités fonctionnelles de TFIIH: les protéines du cœur-TFIIH occupent l'anneau alors que les trois sous-unités du CAK correspondent à la protubérance.

Une approche alternative et complémentaire consiste à localiser certaines sous-unités, et en particulier les sous-unités catalytiques, par des expériences d'immunomarquage utilisant des anticorps monoclonaux. Le marquage de la kinase cdk7 a tout d'abord confirmé que cette sous-unité, et par conséquent le sous-complexe CAK, faisait partie de la protubérance. Les anticorps dirigés contre les sous-unités p44, XPB et XPD se lient tous au niveau de l'anneau en accord avec la comparaison du cœur-TFIIH au facteur entier. L'épitope reconnu par l'anticorps dirigé contre p44 est situé à la base de la protubérance, à mi-chemin entre les extrémités carboxy-terminales des hélicases XPB et XPD contre lesquels sont dirigés les anticorps monoclonaux. Ces observations structurales sont conformes à de nombreuses données biochimiques et donnent un éclairage nouveau sur les études visant à élucider la relation entre les mutations et les phénotypes des patients atteints de xeroderma pigmentosum (patients XP). Il avait ainsi été montré que XPD interagissait avec p44 et que cette interaction était fonctionnellement importante dans la mesure où de nombreux patients XP portent des mutations dans la région carboxy-terminale de l'hélicase. Chez les patients, cette interaction est fragilisée et l'activité de XPD dans la réparation est réduite [18]. La proximité des épitopes de p44 et XPD observée dans nos expériences de marquage est en accord avec ces données biochimiques et suggère qu'une interaction faible entre ces deux protéines pourrait rompre la structure en anneau. Par ailleurs, XPD interagit avec MAT1, une des trois sous-unités du CAK, et certaines mutations fragilisent l'interaction CAK-XPD, affectant de ce fait la composition de TFIIH [5]. Ces données sont étayées par la



**Figure 1. (A) Molécules isolées de TFIIH observées en microscopie électronique par coloration négative. (B) Modèle 3D du facteur de transcription/réparation TFIIH humain à une résolution de 3,8 nm déterminé après analyse numérique des images moléculaires. Le sous-complexe cœur-TFIIH est localisé dans la partie annulaire alors que le sous-complexe CAK correspond à la protubérance supérieure. Les expériences d'immunomarquage moléculaire ont permis de positionner les trois sous-unités catalytiques: la kinase cdk7 ainsi que les hélicases XPB et XPD (flèches). La sous-unité p44, qui semble jouer un rôle essentiel dans la stabilité du complexe, a été localisée à la base de la protubérance.**

localisation du CAK et de l'hélicase XPD; cette dernière est positionnée sur l'anneau, à proximité de la protubérance, pouvant ainsi stabiliser l'interface entre le CAK et le cœur de TFIIH.

La détermination de la structure 3D de TFIIH constitue une étape importante pour la compréhension à l'échelle moléculaire des mécanismes de la réparation, de la transcription et des maladies associées. Les efforts futurs viseront à améliorer la résolution grâce à l'utilisation de la cryomicroscopie qui permet de mieux préserver la structure hydratée de l'échantillon dans le vide du microscope. Un enjeu essentiel consiste à étudier le facteur dans le contexte fonctionnel de la réparation ou de la transcription en présence de tous les composants caractérisés. Il faut noter l'importance des efforts de recherche consentis actuellement pour positionner l'ensemble des facteurs de transcription dans le complexe d'initiation. A cet égard, des travaux de photo-pontage mettent en évidence le rôle central du facteur TFIIH qui permettra aux deux brins d'ADN de s'écarter en partie [19,20]. La poursuite de nos travaux devrait permettre de visualiser comment TFIIH interagit avec l'ADN lors de l'ouverture du promoteur et donner une explication moléculaire à l'activité transcriptionnelle réduite observée chez certains patients XPB ■

#### Remerciements

Ce travail a été réalisé grâce à l'aide de l'Inserm, du Cnrs, des Hôpitaux Universitaires de Strasbourg, de l'Association pour la recherche contre le cancer et le programme *Human science frontiers* (Grants RG-193/97 et RG-336/99). SF a bénéficié d'une bourse du MRES et de la Ligue contre le Cancer.

#### RÉFÉRENCES

1. Sarasin A. Les gènes humains de la réparation de l'ADN. *Med Sci* 1994; 10: 43-54.
2. Citterio E, Vermeulen W, Hoeijmakers JH. Transcriptional healing. *Cell* 2000; 101: 447-50.
3. Svestrup JQ, Vichi P, Egly JM. The multiple roles of transcription/repair factor TFIIH. *Trends Biochem Sci* 1996; 21: 346-50.
4. de Laat WL, Jaspers NG, Hoeijmakers JH. Molecular mechanism of nucleotide excision repair. *Genes Dev* 1999; 13: 768-85.
5. Coin F, Bergmann E, Tremeau-Bravard A, Egly JM. Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. *EMBO J* 1999; 18: 1357-66.
6. Rield T, Egly JM. Phosphorylation intratranscription: the CTD and more. *Protein Expression* 2000 (sous presse).
7. Tan S, Richmond TJ. Eukaryotic transcription factors. *Curr Opin Struct Biol* 1998; 8: 41-8.
8. Cramer P, Bushnell DA, Fu J, et al. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 2000; 288: 640-9.
9. Moras D, Gronemeyer H. The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol* 1998; 10: 384-91.
10. Weatherman RV, Fletterick RJ, Scanlan TS. Nuclear-receptor ligands and ligand-binding domains. *Annu Rev Biochem* 1999; 68: 559-81.
11. Birck C, Poch O, Romier C, et al. Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* 1998; 94: 239-49.
12. Brand M, Laurent C, Mallouh V, Tora L, Schultz P. Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science* 1999; 286: 2151-3.
13. Andersen G, Busso D, Poterszman A, et al. The structure of cyclin H: common mode of kinase activation and specific features. *EMBO J* 1997; 16: 958-67.
14. Fribourg S, Kellenbergers E, Rogniaux H, et al. Structural characterization of the cysteine-rich domain of TFIIH p44 subunit. *J Biol Chem* 2000 (sous presse).
15. Schultz P, Fribourg S, Poterszman A, et al. Molecular structure of human TFIIH. *Cell* 2000; 102: 599-607.
16. Matadeen R, Patwardhan A, Gowen B, et al. The *Escherichia coli* large ribosomal subunit at 7.5 Å resolution. *Struct Fold Des* 1999; 7: 1575-83.
17. Drapkin R, Le Roy G, Cho H, Akoulitchev S, Reinberg D. Human cyclin-dependent kinase-activating kinase exists in three distinct complexes. *Proc Natl Acad Sci USA* 1996; 93: 6488-93.
18. Coin F, Marinoni JC, Rodolfo C, Fribourg S, Pedrini AM, Egly JM. Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet* 1998; 20: 184-8.
19. Kim TK, Ebricht RH, Reinberg D. Mechanism of ATP-dependent promoter melting by transcription factor IIIH. *Science* 2000; 288: 1418-22.
20. Douziech M, Coin F, Chipoulet JM, et al. Mechanism of promoter melting by the p89/XPB DNA helicase of transcription factor TFIIH revealed by protein-DNA photocross-linking. *Mol Cell Biol* 2000 (sous presse).

**Arnaud Poterszman**  
**Patrick Schultz**  
**Sébastien Fribourg**  
**Dino Moras**  
**Jean Marc Egly**

*Institut de Génétique et de Biologie Moléculaire et Cellulaire, Cnrs/Inserm/ULP, 1, rue Laurent Fries, BP 163, 67404 Illkirch Cedex, France.*

#### TIRÉS À PART

A. Poterszman.