ISSN 2427-4577

BULLETIN N° 236
ACADÉMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES
INTERDISCIPLINARY EUROPEAN ACADEMY OF SCIENCES

Lundi 3 juin 2019 :
à 16h30
à l'Institut Henri Poincaré salle 314
11, rue Pierre et Marie Curie 75005 PARIS/Métro : RER Luxembourg

Conférence :
" Enigmes posées par la mémoire et l'apprentissage"
par Vincent HAKIM
Directeur de Recherche CNRS, Équipe biophysique et neuroscience théoriques
Département de Physique –École Normale Supérieure
24, rue Lhomond 75231 Paris Cedex 05

Notre Prochaine séance aura lieu le lundi 9 septembre 2019 à 16h30
à l'Institut Curie, salle Joliot,
11-13 rue Pierre et Marie Curie 75005 PARIS/Métro: RER Luxembourg

Elle aura pour thème

Conférence:
" Il y a plus de marge de manœuvre en bas de l'échelle:
vers un détecteur universel des interactions moléculaires "
Par Terence STRICK
Directeur de recherche au CNRS,
Professeur et chef d'équipe Nanomanipulation de biomolécules
Institut Jacques Monod Université Paris Diderot et Institut de Biologie de l'ENS (IBENS)

ACADÉMIE EUROPÉENNE INTERDISCIPLINAIRE DES SCIENCES
INTERDISCIPLINARY EUROPEAN ACADEMY OF SCIENCES

juin 2019

N°236

TABLE DES MATIERES

p. 03 Séance du 3 juin 2019 :
p. 05 Documents

Prochaine séance : lundi 9 septembre 2019

Conférence:

" Il y a plus de marge de manœuvre en bas de l'échelle:
vers un détecteur universel des interactions moléculaires "

Par Terence STRICK

Directeur de recherche au CNRS, , Professeur et chef d'équipe Nanomanipulation de biomolécules

Institut Jacques Monod Université Paris Diderot et Institut de Biologie de l'ENS (IBENS)

ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES
Fondation de la Maison des Sciences de l'Homme, Paris.

Séance du Lundi 3 juin 2019/IHP 16h30

La séance est ouverte à 16h **sous la Présidence de Victor MASTRANGELO** et en la présence de nos Collègues Gilbert BELAUBRE, Jean BERBINAU, Gilles COHEN-TANNOUDJI, Françoise DUTHEIL, Claude ELBAZ, Jean -Pierre FRANCOISE, Michel GONDRAN, Irène HERPE-LITWIN, Claude MAURY, Marie-Françoise PASSINI, Edith PERRIER, Jacques PRINTZ, Denise PUMAIN, René PUMAIN, Jean SCHMETS, Jean-Pierre TREUIL .

Etaient excusés :François BEGON, Jean BERBINAU, Jean-Pierre BESSIS, Jean-Louis BOBIN, Bruno BLONDEL, Michel CABANAC, Alain CARDON, Juan-Carlos CHACHQUES, Eric CHENIN, Alain CORDIER , Daniel COURGEAU, Sylvie DERENNE, Ernesto DI MAURO, Jean-Félix DURASTANTI, Vincent FLEURY, Robert FRANCK, Dominique LAMBERT, Pierre MARCHAIS, Anastassios METAXAS, Jacques NIO, Pierre PESQUIES, Michel SPIRO, Alain STAHL, Mohand TAZEROUT, Jean-Paul TEYSSANDIER, Jean VERDETTI.

I. Conférence de Vincent HAKIM

A. Présentation du conférencier par notre Président Victor MASTRANGELO

Vincent Hakim est Directeur de Recherche à l'Ecole Normale Supérieure (ENS) de Paris. C'est un physicien théoricien spécialisé en physique statistique et en dynamique non linéaire. Durant les vingt dernières années , il a effectué des travaux en neurosciences théoriques , tout particulièrement dans le domaine de la dynamique du réseau et des oscillations. Il a contribué à éclaircir les mécanismes des très fréquentes "oscillations partiellement synchronisées" . Il a également travaillé étroitement avec plusieurs équipes de recherche expérimentale , tout particulièrement à l'ENS. Ceci l'a conduit à examiner divers problèmes relatifs à la physiologie synaptique tels que l'observation des synapses silencieuses du cervelet , les mécanismes moléculaires de formation et de maintenance et récemment le problème d'affectation (***"credit-assignment "***) (*entre groupes d'agent et de tâches*) des synapses lors de l'apprentissage.

II. Conférence " *Enigmes posées par la mémoire et l'apprentissage* "

Résumé de la conférence:

Énigmes concernant la mémoire à long-terme et l'apprentissage

Les synapses sont des structures biologiques importantes servant à la transmission de l'information entre les neurones et on pense qu'elles sont les sites de l'apprentissage et de la mémorisation. Néanmoins subsiste l'énigme de savoir comment la mémoire persiste pendant des années alors que le cycle des composants synaptiques se déroule à l'échelle horaire. De même, lors de l'apprentissage d'une tâche complexe, en général, la réaction semble le plus souvent globale et pauvre en information. On a alors du mal à expliquer comment les forces de nombreuses synapses différentes peuvent s'ajuster correctement , ce qu'on appelle en anglais le "*credit-assignment problem*" (" problème d'attribution de la compétence") . Après un rappel de quelques éléments de biophysique synaptique et de découvertes expérimentales significatives , je vais discuter de nos travaux récents en collaboration avec les équipes d'A. TRILLER et B. BARBOUR à l'IBENS en vue de la résolution de ces deux énigmes.

Un compte-rendu rédigé par un membre de l'AEIS qui est neurobiologiste, sera prochainement disponible sur le site de l'AEIS http://www.science-inter.com.

REMERCIEMENTS

Nous tenons à remercier vivement Mme Sylvie BENZONI Directrice de l'Institut Henri POINCARÉ et Mmes Florence LAJOINIE et Chantal AMOROSO ainsi que les personnels de l'IHP pour la qualité de leur accueil.

Documents

- p. 06 : Notre collègue René PUMAIN nous a confié un compte-rendu de la conférence du Pr Antoine TRILLER du 6 mai 2019 intitulée "***Biologie quantitative de la communication entre neurones : instabilité moléculaire et mémoire, du normal au pathologique***".

Pour préparer la conférence du Pr Terence STRICK , nous vous proposons :

- p. 10: un résumé de sa conférence ***" Il y a plus de marge de manœuvre en bas de l'échelle: vers un détecteur universel des interactions moléculaires "***

- p.11 : ***Dissection of DNA double-strand break repair using novel single-molecule forceps.*** *Nature Struct. Mol. Biol.*, **25**: 482—487 (2018) J. Wang, C. Duboc, Q. Wu, T. Ochi, S. Liang, S.E. Tsutakawa, S.P. Lees-Miller, M. Nadal, J.A. Tainer, T.L Blundell, & T.R. Strick accessible sur le site https://www.researchgate.net/publication/329558051_Dissection_of_DNA_double-strand-break_repair_using_novel_single-molecule_forceps

- p.34 : ***Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching.*** *Science* **314**: 1139—1143 (2006). A. Revyakin, C.-Y. Liu, R.H. Ebright and T.R. Strick accessible sur le site : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754787/pdf/nihms-142605.pdf

Biologie quantitative de la communication entre neurones : instabilité moléculaire et mémoire, du normal au pathologique

Compte rendu de la Conférence d'Antoine Triller du 6 mai 2019

rédigé par notre Collègue René PUMAIN

La conférence a été présentée en trois parties :

- Une première partie pour rappeler les bases fondamentales fonctionnement du système nerveux central, en particulier au niveau synaptique.
- Une deuxième partie pour présenter les travaux réalisés pendant ces dernières années se poursuivant par nombre de projets.
- La troisième partie pour montrer comment certains résultats de la recherche fondamentale peuvent permettre de mieux comprendre les mécanismes de certaines pathologies, en particulier certaines maladies dégénératives du système nerveux telle que la maladie d'Alzheimer et donc de proposer de nouvelles approches thérapeutiques.

Mécanismes neurophysiologiques de base de la transmission synaptique

Le système nerveux central comprend chez les mammifères essentiellement les hémisphères cérébraux, le cervelet et la moelle épinière, représentant chez l'homme une masse d'environ 1,5 kg. Cette structure est formée principalement de deux types de cellules, les neurones et les cellules gliales en proportions à peu près équivalentes. Les neurones dont on estime le nombre dans le cerveau humain de 80 à 100 milliards peuvent former des réseaux de connexions grâce à leurs extensions que l'on appelle des dendrites et des axones. La structure fine ces connexions a été décrite au début du XXe siècle par un histologiste espagnol du nom de Santiago Ramon y Cajal en utilisant une technique imprégnation argentique développée par Camillo Golgi. Ces deux auteurs ont obtenu le prix Nobel ensemble en 1906. Ramon y Cajal a publié l'essentiel de son œuvre dans un livre écrit en français, Histologie du système nerveux de l'homme et des vertébrés.

Les neurones sont capables de générer une activité rapide sous forme de variations de potentiel membranaire correspondant à des courants ioniques transmembranaires que l'on appelle des potentiels d'action. Ces activités peuvent être transmises à d'autres neurones proches par l'intermédiaire de structures spécifiques que l'on appelle des synapses. Le prolongement axonal (généralement un par neurone) se projette sur les autres prolongements de certains neurones voisins que l'on appelle les dendrites (qui peuvent être relativement nombreux) au niveau de ces synapses. Les connexions synaptiques peuvent avoir sur les neurones cible soit des influences excitatrices pouvant déclencher des potentiels d'action dans les neurones cible soit des actions inhibitrices permettant de prévenir le déclenchement de potentiels d'action dans ces neurones cible. La propagation des activités électriques excitatrices dépend donc du nombre et de l'efficacité des synapses excitatrices présentes sur le neurone cible et de la présence éventuelle d'activités inhibitrices plus ou moins synchrones avec des activités excitatrices.

Les potentiels d'action axonaux transmettent généralement leur activité à leur extrémité, au niveau présynaptique, par la libération au niveau de la membrane du neurone cible de substances chimiques que l'on appelle des neurotransmetteurs. On a mis en évidence un grand nombre de neurotransmetteurs qui sont

généralement des molécules de petite taille. Dans le système nerveux central le neurotransmetteur excitateur le plus important est un acide aminé, le glutamate, et le neurotransmetteur inhibiteur le plus efficace est l'acide gamma aminobutyrique, aussi appelé GABA. Ces substances se retrouvent localisées et concentrées dans de petites vésicules situées dans la terminaison présynaptique. La survenue d'un potentiel d'action au niveau de ces terminaisons provoque leur libération dans la fente synaptique par des mécanismes complexes impliquant l'entrée d'ions calcium dans cette terminaison. Ces substances diffusent très rapidement dans la fente synaptique (car la distance entre la terminaison présynaptique et la membrane post synaptique est très petite) et agissent sur la membrane présynaptique par l'intermédiaire de protéines membranaires dont la structure est modifiée par leur liaison avec le neurotransmetteur. En créant des pores dans la membrane post synaptique ces récepteurs-canaux provoquent l'apparition de courants ioniques dans cette membrane qui modifient le potentiel membranaire dans le neurone cible et déclenchent éventuellement des potentiels d'action dans ce neurone (la dépolarisation post synaptique se produit en environ 1 milliseconde dans les synapses excitatrices). Ces protéines que l'on appelle des récepteurs post synaptiques ont donc deux caractéristiques, un site de liaison au neurotransmetteur et la possibilité de former un canal membranaire permettant le transfert de différentes espèces ioniques à travers la membrane post synaptique.

Un neurone peut être la cible de milliers de contacts synaptiques, soit excitateurs soit inhibiteurs. L'efficacité de ces connexions donne naissance à des circuits neuronaux privilégiés qui sont impliqués dans toutes les fonctions du système nerveux, en particulier dans les phénomènes d'apprentissage à court terme et de mémorisation à long terme, par renforcement ou délitement de l'efficacité de ces circuits neuronaux. L'efficacité de la transmission synaptique dépend donc à la fois de la stabilité des synapses et de leur possibilité d'être modulées par les modifications de l'environnement, par l'intermédiaire des afférences sensorielles. Ces mécanismes très importants dépendent de ce qu'on appelle la plasticité synaptique.

Recherches récentes

En ce qui concerne les mécanismes moléculaires impliqués dans la stabilité des synapses, Antoine Triller et ses collaborateurs ont développé en relation avec des physiciens des techniques (quantum-dots ou points quantiques) permettant de marquer et de visualiser des molécules isolées dans des neurones vivants, essentiellement en culture. Ils ont en particulier étudié les récepteurs post synaptiques pour différents neurotransmetteurs. Cette avancée a permis de passer d'une observation statique à un suivi dynamique à haute résolution. Ils ont par la suite utilisé des techniques d’imagerie en super-résolution (PALM, STORM) ce qui a permis de préciser l'organisation des structures post synaptiques.

Il avait été montré préalablement que les protéines attachées aux membranes lipidiques des cellules pouvaient se déplacer en permanence, comme un phénomène de diffusion du fait de l’agitation thermique brownienne. Comment donc les récepteurs pop synaptiques peuvent rester stabilisés en des endroits précis pour que la transmission synaptique puisse se produire ?

Cette stabilisation transitoire des récepteurs au niveau synaptique se réalise par des interactions avec des protéines intracellulaires voisines de la membrane post synaptique que l'on appelle les protéines d'ancrage ou protéines d'échafaudage. Ces interactions stabilisent à la fois les récepteurs post synaptiques et les protéines d'ancrage mais seulement pendant un «certain temps». En réalité, il y a des récepteurs-canaux post synaptiques en dehors de la zone proprement synaptique qui peuvent venir s'agréger dans cette zone, alors que des récepteurs initialement dans la zone post synaptique peuvent s'en détacher (diffusion–rétention). Il s'agit donc d’un mécanisme essentiel de la régulation du nombre des récepteurs post synaptiques, mais en quelque sorte de façon probabiliste. Le nombre de ces récepteurs au niveau d’une zone post synaptique peut

varier même si la valeur moyenne reste à peu près constante. Ce mécaniste se traduit par un «bruit» du courant post synaptique.

Ce mécanisme est régulé par des ions, en particulier les ions calcium et par une série d'enzymes de phosphorylations et de déphosphorylations de protéines. On peut donc considérer qu'il y a un processus de régulation biologique qui s'exerce sur une activité de diffusion qui est un processus physique. Il est lui-même dépendant de l'activité des neurones et peut être modifié par des substances pharmacologiques.

On savait depuis longtemps que les protéines synthétisées dans les cellules sont métabolisées relativement rapidement, ce qui pose un problème pour le maintien de la mémoire à long terme. Ce processus statistique permet de comprendre le maintien des fonctions sur des périodes de temps assez longues, de nouvelles protéines étant synthétisées et amenées à la membrane en ayant une certaine probabilité de remplacer les protéines ayant été métabolisées (système ergodique).

Le déplacement d'une molécule dépend du potentiel de capture et peut-être modulée par la viscosité du milieu et de la diffusivité. L'énergie nécessaire pour la capture est relativement faible et cette capture est donc labile. Les neurones dépensent de l'énergie pour contrôler ces mouvements et maintenir les molécules à des sites particuliers par des mécanismes qui peuvent diminuer la fluidité membranaire.
Si on modifie les protéines récepteur par des mutations on peut induire une variation de l'énergie des interactions de capture entre protéines.

Modélisations

Des modèles ont été développés au sein de l'équipe pour répondre à différentes questions, par exemple pourquoi les agrégats de récepteurs restent relativement restreints et n'envahissent pas toute la surface membranaire ? Dans ces modèles contraints, les agrégats atteignent une certaine dimension d'équilibre. Il a été possible également de mettre en évidence des transitions de phase montrant que la stabilisation des récepteurs post synaptiques et des protéines d'échafaudage est réciproque. Il est possible de modifier la taille des agrégats de récepteurs post synaptiques par différents moyens et donc de modifier l'efficacité de la transmission synaptique pour rendre compte de sa plasticité.

Maladies neurodégénératives

Parmi les maladies neurodégénératives, la maladie d'Alzheimer est une maladie qui évolue et qui devient progressivement très handicapante. Elle concerne un grand nombre de personnes. Il n'existe pour l'instant aucune thérapie curative ou préventive. Elle apparaît dans certaines zones du lobe temporal, essentiellement niveau de l'hippocampe et des structures péri- hippocampiques et progresse par la suite vers des structures voisines en ne se manifestant que très peu dans les zones corticales primaires.Les signes cliniques se manifestent relativement tardivement et de nombreuses recherches sont réalisées pour la détecter plus précocement. Elle a été décrite initialement par des caractères spécifiques histopathologiques, essentiellement par le dépôt dans l'espace extracellulaire de protéines bêta amyloïde sur certains neurones et des modifications de protéines intracellulaires (protéines tau) qui servent à maintenir la structure des neurones. On observe également une diminution du nombre des synapses.

Les recherches sur cette maladie ont pour l'instant concerné essentiellement les agrégats de protéines bêta amyloïde: sont-elles toxiques et induisent-telles la dégénérescence neuronale? Cette hypothèse n'a pas été confirmée et lors du vieillissement des individus on peut en mettre en évidence dans le tissu nerveux central

de personnes ne présentant aucune manifestation clinique. Néanmoins, les plaques peuvent libérer dans l'espace extracellulaire des peptides qui peuvent être toxiques pour les neurones. Apparemment ces protéines sont mal repliées ou malformées et peuvent induire chez des neurones sains la formation de protéine malformées en quelque sorte par contamination. **Ce mécanisme rappelle celui qui avait été décrit pour les maladies à prions**. Ces protéines malformées contaminantes doivent s'accrocher à la membrane des neurones, probablement au niveau de certains récepteurs au glutamate. Elles peuvent alors induire par une série de mécanismes complexes un dysfonctionnement de la protéine membranaire qui maintient les concentrations intracellulaires des ions sodium et potassium et ainsi provoquer la dégénérescence neuronale. Le projet est de retarder au maximum l'accolement des protéines malformées à la membrane des neurones, soit au niveau des protéines malformées soit au niveau des récepteurs membranaires sur lesquels elles se fixent. Par les mêmes méthodes utilisées pour étudier les récepteurs post synaptiques, l'efficacité de différents agents pharmacologiques destinés à empêcher la fixation des protéines malformées sur la membrane des autres neurones est étudiée.

Résumé Article Terence STRICK:

Il y a plus de marge de manœuvre en bas de l'échelle:
vers un détecteur universel des interactions moléculaires

Depuis maintenant plusieurs décennies des outils biophysiques ont permis aux chercheurs de manipuler et observer en temps réel les réactions biologiques ayant lieu à l'échelle du réactif moléculaire individuel. Avec un pouvoir de résolution qui atteint vraiment l'échelle allant de l'Angstrom au nanomètre qui est celle des liaisons et des molécules individuelles, le nouveau domaine appelé " biophysique des molécules individuelles" a depuis lors fourni aux biologistes et aux physiciens une vision sans précédent sur la manière dont les moteurs et les machines biologiques moléculaires assemblent, fonctionnent et désassemblent pour permettre à nos cellules de vivre et de fonctionner correctement. Dans cette présentation nous fournirons une vue d'ensemble historique du domaine et nous montrerons comment il a modifié notre compréhension de la structure , de la fonction et de l'organisation moléculaire et cellulaire, par exemple dans des processus clés tels que l'expression et la réparation des gènes. En retour, ces idées nous ont permis de voir sous un nouveau jour les biomatériaux et de commencer à les assembler dans de nouvelles voies permettant de générer de nouvelles et utiles fonctionnalités et mesures . Nous expliquons ceci grâce au développement d'une nouvelle catégorie de détecteurs moléculaires avec des applications au dépistage, à la caractérisation et au perfectionnement des médicaments.

Dissection of DNA double-strand-break repair using novel single-molecule forceps

Article *in* Nature Structural & Molecular Biology · June 2018

CITATIONS
10

READS
83

11 authors, including:

Qian Wu
University of Cambridge
17 PUBLICATIONS **383** CITATIONS
SEE PROFILE

Marc Nadal
Paris Diderot University
40 PUBLICATIONS **1,071** CITATIONS
SEE PROFILE

John Tainer
University of Texas MD Anderson Cancer Center
607 PUBLICATIONS **39,081** CITATIONS
SEE PROFILE

Tom L Blundell
University of Cambridge
872 PUBLICATIONS **52,384** CITATIONS
SEE PROFILE

Article:
Wang, JL, Duboc, C, Wu, Q orcid.org/0000-0002-6948-7043 et al. (8 more authors) (2018) Dissection of DNA double-strand-break repair using novel single-molecule forceps. Nature Structural & Molecular Biology, 25 (6). pp. 482-487. ISSN 1545-9993

https://doi.org/10.1038/s41594-018-0065-1

Published in final edited form as:
Nat Struct Mol Biol. 2018 June ; 25(6): 482–487. doi:10.1038/s41594-018-0065-1.

Dissection of DNA double-strand break repair using novel single-molecule forceps

Jing L. Wang[#1], **Camille Duboc**[#1], **Qian Wu**[#2], **Takashi Ochi**[2,&], **Shikang Liang**[2], **Susan E. Tsutakawa**[3], **Susan P. Lees-Miller**[4], **Marc Nadal**[1], **John A. Tainer**[3,5], **Tom L. Blundell**[2], and **Terence R. Strick**[1,6,7,†]

[1]Institut Jacques Monod, CNRS, UMR7592, University Paris Diderot, Sorbonne Paris Cité 75013 Paris, France

[2]Department of Biochemistry, Cambridge University, Tennis Court Road, Cambridge, CB2 1GA,UK

[3]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[4]Department of Biochemistry and Molecular Biology, Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, Alberta T2N 4N1, Canada

[5]Department of Molecular and Cellular Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA

[6]Ecole normale supérieure, Institut de Biologie de l' Ecole normale supérieure (IBENS), CNRS, INSERM, PSL Research University, 75005 Paris, France

[7]Programme Equipe Labellisées, Ligue Contre le Cancer, 75013 Paris, France

[#] These authors contributed equally to this work.

Abstract

Repairing DNA double-strand breaks (DSBs) by non-homologous end-joining (NHEJ) requires multiple proteins to recognize and bind DNA ends, process them for compatibility, and ligate them together. We constructed novel DNA substrates for single-molecule nano-manipulation allowing us to mechanically detect, probe, and rupture in real-time DSB synapsis by specific human NHEJ components. DNA-PKcs and Ku allow DNA end synapsis on the 100 ms timescale, and addition of PAXX extends this lifetime to ˜2s. Further addition of XRCC4, XLF and Ligase IV resulted in minute-scale synapsis and led to robust repair of both strands of the nanomanipulated DNA. The

[†]To whom correspondence should be adressed: strick@ens.fr.
[&]Present address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK

Author Contributions

J.L.W, C.D., Q.W., T.O., S.L., S.E.T., S.P.L.-M., M.N., J.A.T., T.L.B., and T.R.S designed experiments; J.L.W, C.D., Q.W., T.O., S.L., and T.R.S. prepared reagents; J.L.W, C.D., Q.W., and T.R.S carried out experiments; J.L.W, C.D. and T.R.S analyzed data; J.L.W., C.D., Q.W., T.O., S.L., S.E.T., S.P.L-M., M.N., J.A.T., T.L.B. and T.R.S. wrote the paper.

Competing Financial Interests The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

energetic contribution of the different components to synaptic stability is typically on the scale of a few kCal/mol. Our results define assembly rules for NHEJ machinery and unveil the importance of weak interactions, rapidly ruptured even at sub-picoNewton forces, in regulating this multicomponent chemomechanical system for genome integrity.

Introduction

NHEJ is a complex multicomponent process, which is conceptually divided into three steps: DNA end synapsis, processing, and ligation. Ku70–Ku80 heterodimer (Ku) recognizes DNA ends [1] and recruits DNA-PKcs to form the DNA-PK holoenzyme, which interacts with another equivalent holoenzyme to form a synaptic complex [2]. Through autophosphorylation, the holoenzyme then provides a platform for assembly and regulation of downstream components [3, 4, 5, 6, 7, 8]. For instance, DNA-PKcs recruits Artemis to open the DNA hairpin structure during V(D)J recombination [9, 10]. Although PAXX has recently been identified as important for NHEJ, its role remains unclear [11, 12, 13, 14]. XRCC4 and XLF interact with each other and form filament structures [15, 16, 17, 18, 19]. The XLF-XRCC4 complexes interact with DNA-PK assembled at DNA ends and are believed to bridge spatially two DNA ends so as to facilitate ligation of the break by Ligase IV [20, 21]. Although the overall process is understood with the key components identified and characterized through various approaches, critical information concerning the forces that hold together DNA ends, the kinetics of assembly and disassembly of NHEJ complexes, as well as the robustness of the repair pathway to fluctuations in complex composition remain unknown [22, 23]. Here we used single-molecule experimentation along with a novel DNA substrate to enable mechanical and energetic analysis of molecular interactions across a DNA break and address these questions.

Results

The DNA molecular forceps sustains a functional NHEJ reaction

To measure objectively and quantitatively the properties of complex molecular synapsis and interactions, we developed a DNA-scaffold based single-molecule assay that allows one to observe repeated cycles of synapsis and rupture. Microscopic properties such as kinetics of formation and lifetime of synapsis can be derived from such assays (Figure 1A). Briefly, the DNA system consists of two linear double-strand DNA segments ~1510 bp in length, connected to each other by a third double-strand segment, termed a “bridge,” of ~690 bp (see Online Methods and Supplementary Figure 1). The bridge is anchored 58 bp away from the ends of each DNA segment it tethers, allowing the ends to fluctuate freely as they face each other and providing ample space for loading of NHEJ components [6, 24, 25]. We attach this construct at one end to a digoxigenin-modified glass surface and at the other end to a streptavidin-coated magnetic bead via 1-kb digoxigenin- or biotin-labelled DNA fragments attached to either end of the ~3.6 kbp construct. With a magnetic trap and computer-aided video-microscopy we applied a vertical extending force, F, to the bead and hence the construct, all the while imaging the position of the magnetic bead above the surface [26]. This system allows real-time monitoring of DNA end synapsis by monitoring the molecule’ s overall extension, *l*. If there is no synapsis between the DNA ends, then the extension of the

~3.6 kbp construct under a 1.4 picoNewton (pN) force (1 pN = 10^{-12} N) is predicted to be 1085 nm based on the worm-like chain (WLC) model which describes the mechanical properties of DNA [27]. If there is synapsis between the DNA ends, then the bridge no longer contributes to the extension of the system, which is now predicted to be only 913 nm, or 172 nm shorter than in the absence of synapsis.

We first show that this system can successfully recapitulate complete NHEJ of blunt-ended DSBs (Fig. 1B). We digested the DNA with SmaI restriction enzyme, assembled constructs in the magnetic trap (~25-50 constructs/field-of-view), and introduced into the reaction chamber the full complement of human NHEJ components (10 nM Ku, 100 pM DNA-PKcs, 20 nM PAXX, 20 nM XLF and 20 nM XRCC4–Ligase IV, see Supp. Fig. 1C,D). We repeatedly cycled the force applied to the DNA between a low and high value with 450 second period (F_{low} = 0.04 pN for ~225 seconds, F_{high} = 1.4 pN for ~225 seconds); in the low-force state the DNA ends are free to encounter each other whereas in the high-force state we can reliably interrogate the DNA as to the status of its ends. After one or two traction cycles we frequently observed a DNA state with reduced extension (Fig. 1B) and which never recovered initial extension, indicative of successful DNA ligation. Successful ligation of the blunt DNA ends of the construct generates a SmaI site. Thus, after disrupting residual NHEJ components with 0.2% SDS and then washing out the detergent (see Materials and Methods), we introduced SmaI into the experiment and the initial DNA state was restored via an abrupt increase in DNA extension $<\Delta>$ = 161 +/- 2 nm (SEM, n = 28, Fig. 1C). This repair-specific cleavage reaction provides us with a length-change calibration that we can then use as a robust signature of bonafide, specific interactions between the ends of the DNA construct. Most DNA molecules (36 out of 50, or 72%) were repaired within the first few traction cycles (Fig. 1D), underscoring the robust nature of DSB repair in this system when all components are present. Ligation was still observed, although at lower efficiency, if any one of PAXX, XLF or XRCC4 C-terminal domain were omitted from the reaction (Table 1). No ligation was observed if DNA-PKcs or Ligase IV was omitted (Table 1) or if DNA ends were dephosphorylated (see below).

The ability of the ligated DNA to supercoil demonstrated that there was a double ligation with both strands of the DNA ligated to their respective counterpart in the other DNA end (Supp. Fig. 2A). We also show (Supp. Fig. 2B-D) that T4 DNA ligase, although capable of efficiently ligating the construct prepared with overhanging ends (via XmaI digestion), was unable to efficiently ligate the construct with blunt ends, underscoring the requirement *in vivo* for robust molecular pathways to repair non-complementary DSBs. We conclude that this DNA system functionally recapitulates human NHEJ at single-molecule resolution, providing us with a basis by which to evaluate the contribution of each reaction component to synaptic stability.

Stepwise Assembly of the human NHEJ machinery

We thus proceeded by progressively assembling the NHEJ machinery and testing for DNA end synapsis by the force-modulation method described above. We find that the minimal combination of the DNA-PK holoenzyme (DNA-PKcs plus Ku) and PAXX is necessary to observe frequent, second-scale interactions between the two DNA ends (Fig. 2A,B and

Supp. Table 1). In this assay an interaction was observed as a transient plateau in the DNA extension signal obtained upon increasing the force from F_{low} to F_{high}; duration of the plateau reflects the lifetime of the end interaction. The histogram depicting the distribution of observed Δl values (Fig. 2C) again displays a Gaussian peak located close to the expected amplitude (166 ± 1 nm SEM, n =129). Comparing the change in DNA extension detected at the end of the plateau to our calibrated length-change obtained via specific recleavage of the repaired DNA ends confirms we are detecting bonafide interactions between the ends of the DNA construct; we specifically designate as "synaptic" those interaction events within three standard deviations of the likeliest value. In Fig. 2D we show that the lifetime distribution for synaptic events follows single-exponential behavior with an average, $<t_{synapsis}>$, of 2.2 ± 0.3 s (SEM, n = 98). Synapsis frequency, defined as the fraction of force-modulation cycles in which we observe a synaptic event, is roughly 6% (102 synaptic events out of 1611 force-modulation cycles collected on ˜90 DNA substrates, see Supp. Table 1).

These synaptic events did not depend significantly on the phosphorylation state of the DNA ends (Fig. 3A). Notably, synaptic events were abolished when we replaced wild-type PAXX with a mutant (V199A and F201A) deficient in its ability to interact with the Ku subunit of the DNA-PK holoenzyme (Fig. 3B). XLF was unable to substitute for PAXX (Fig. 3C) despite its ability to interact with Ku [28, 29]. Neither DNA-PK holoenzyme nor PAXX on their own generates synapsis on the 2-second scale (Fig. 3D, E). We conclude that the "upstream" components Ku, DNA-PKcs and PAXX sustain formation of DNA end synapsis which can resist up to ˜1 pN traction forces (0.15 kCal/mol·nm) for second time-scales, and that PAXX plays a central role in stable bridging of the gap between the two DNA ends.

To assess the role of the "downstream" components XRCC4, XLF and Ligase IV we added them in combination with the "upstream" components used earlier, to obtain a "complete" reaction (Fig. 4). Unlike in the prior assays with all components however, here we dephosphorylated the DNA ends so as to prevent ligation, allowing us to repeatedly interrogate the construct. We now observed long-lived synaptic interactions: $<t_{synapsis}>$ = 66.4 ± 7.6s (SEM, n=175). Omitting any one of Ligase IV, XLF and the N-terminal portion of XRCC4 from the above reaction (Supp. Fig. 3A-C) resulted in the same short synaptic lifetime as was observed in the presence of Ku, DNA-PKcs and PAXX using dephosphorylated DNA substrate (Fig. 3A), i.e. of the order of 2-3 seconds. This supports and extends observations that XRCC4, XLF and Ligase IV are involved in a series of joint interactions that could lead them to act as a modular, functional unit of NHEJ; Ligase IV interacts with the coiled-coil (cc) region of XRCC4 (ccXRCC4, corresponding to residues 138-213), the N-terminal domain of which associates with XLF [21, 30]. Removing any one of these three components thus abolishes the stabilizing effects of this functional unit. The same results were obtained when these experiments were repeated using a phosphorylated DNA substrate (compare Supp. Fig. 3D-F and Fig. 2).

Replacing Ligase IV by a catalytically-inactive mutant (K273A, see Supp. Figs. 1D and 3G) resulted in a synaptic lifetime of 34.5±6.9 s (SEM, n = 82) reflecting a mild impact on synapsis lifetime and indicating this component could still sustain the web of interactions necessary for downstream components to stabilize synapsis by DNA-PK and PAXX. Surprisingly, stable interactions were observed when wild-type XLF was replaced by a

mutant (XLF_{1-233}) deficient in its interactions with Ku (44±12.6s SEM, n=84, see Supp. Fig. 3H). However, removing DNA-PKcs from the complete reaction essentially abolished synapsis (Supp. Fig. 3I). Synapsis was also abolished when both XLF and PAXX were omitted from the complete reaction (Supp. Fig. 3J), supporting and extending recently-described overlapping function between XLF and PAXX and resulting synthetic embryonic lethality in double-knockout mouse models [31, 32]. We note that the ternary combination of XLF, XRCC4, and Ligase IV does not support synapsis formation under the physiological salt conditions employed in this work (Supp. Fig. 4).

To investigate further the role of PAXX, we next removed it from the complete reaction, resulting in a seven-fold reduction in synaptic lifetime compared to when PAXX was present (9 ± 1.8 s SEM, n = 77, see Supp. Fig. 5A-C). We used these conditions to probe the effect of the traction force on synaptic lifetime. We thus observed that varying F_{high} from 0.1 pN to 4.7 pN had no significant effect on the lifetime of synapsis (Supp. Fig 5D). This indicates that the second-to-minute lifetimes we measure are relevant to the kinetics of NHEJ *in vivo* where the forces acting on broken DNA ends are essentially diffusive. We also note that substituting the PAXX mutant for wild-type PAXX resulted in roughly the same reduction in synaptic lifetime as removing PAXX entirely from the reaction (here to 5.4 ± 1.3s, see Supp. Fig. 5E-G).

Discussion

We thus find that synaptic junctions of varied composition can resist picoNewton-scale forces for times ranging from seconds to minutes. Presumably this is a reflection of actual commitment to repair. Based on these results we propose a simple model for NHEJ synapsis (Fig. 5). Taking the lifetime of the complete synaptic junction as a baseline (66s), we find that removing PAXX is responsible for a roughly seven-fold destabilization of the complete synaptic junction (9s synapsis without PAXX). We likewise determine that disabling the XRCC4–XLF–Ligase IV system is responsible for a roughly 30-fold destabilization of the complete synaptic junction (2s synapsis for just DNA-PK and PAXX, without XRCC4–XLF–Ligase IV). Because the lifetime of a complex is related to its free energy of activation by Boltzmann' s law, we estimate that PAXX contributes k_BT ln (66s/9s) ˜ 2 k_BT or 1.2 kcal/mol to the stability of the complete synaptic junction, whereas XRCC4–XLF–Ligase IV contributes k_BT ln (66s/2s) ˜ 3.5 kBT or 2.1 kcal/mol to the stability of the complete synaptic junction. Together these two sets of components would stabilize the synaptic state by roughly 5.5 k_BT, increasing its lifetime roughly two-hundred fold. These results suggest that the primary synapsis formed by the DNA-PK holoenzyme alone is very short-lived, in the range of hundreds of miliseconds. Finally therefore, to attempt to detect this short-lived interaction, we increased by a factor of ten the rate at which the force is switched between low and high values. The large number of traction cycles thus generated allowed us to detect rare and very short-lived synaptic events of expected synaptic amplitude (Supp. Fig. 6). The lifetime of this synapsis is, as predicted, in the 100 milisecond range. Control experiments show this short-lived and specific synapsis is only observed in the presence of DNA-PKcs and Ku (Supp. Fig. 6 and Supp. Table 2).

The additivity of these short-lived and weak interactions, revealed by these single-molecule experiments, indicates a multivalent system where multiple protein interfaces stabilize DNA end-synapsis. Use of multiple weak interactions to obtain high affinity balanced by regulation is a critical emerging common theme in DNA repair systems, such as RPA whose sub-nanomolar affinity is built on multiple weak interactions [33]. Longer synapsis effectively increase ligation probability.

Our work supports and broadens the observation that PAXX stabilizes the core NHEJ proteins at damaged chromatin [11, 31, 32], and suggests PAXX as a target for cancer drug development. Indeed removing PAXX from the complete reaction reduces the probability of repair from 72% to 28% but does not completely abolish repair – as is the case when XLF, or XRCC4 are absent from the reaction (Table 1). Whether or not PAXX can incorporate into NHEJ complexes only at the initial stage of assembly, or also at later stages of assembly for instance once XRCC4–XLF–Ligase IV have assembled, remains to be determined. With this validated single-molecule system, it will be feasible and intriguing to examine the impacts of other components that act in NHEJ, such as Artemis and APLF [34] proteins as well as noncoding RNA [35], and relate their energetic and mechanistic roles to cancer predispositions resulting from mutational defects and aberrant regulation.

Currently, our results show that novel and functional DNA scaffolds provide unique and unexpected kinetic insights into the human protein machinery responsible for carrying out NHEJ. Our studies thus provide a solid foundation upon which to further explore the order of assembly and stoichiometry of functional repair complexes using fluorescently-labeled protein variants and correlative single-molecule nanomanipulation and fluorescence (NanoCOSM [36]). Importantly, the measurements of synapsis lifetime and k_BT-scale energetics uncover the nature of multi-component DNA break repair complexes. This knowledge has broad scientific and biomedical implications including strategies for targeting the DNA-damage response by small molecules. Specifically, we propose a strategy of chemical inhibitors to compete with the functionally-important yet individually-weak interactions in this major system for break repair in humans.

Materials and Methods

DNA Construct

The construct is assembled from two precursor dsDNA molecules. For the first, we assemble two ssDNA oligos, a short and a long, into a covalently-linked, asymmetric branched structure. In a PCR reaction using a pair of such branched oligos, the longer oligos prime PCR while the shorter one remains unused. The branchpoints end up 58 bp from the end of dsDNA. This “Break” PCR product is 3 kb in length, and by hybridization one can convert the remaining ssDNA branches into dsDNA branches with 4-bp overhangs. This fragment can then be “circularized” by hybridization and ligation to a ~600 bp “Bridge” DNA fragment, which is simply a standard dsDNA fragment digested so as to have overhangs compatible with those of the short dsDNA branches. The circularized DNA can then be reopened with a pair of restriction enzymes and gel purified so as to obtain the desired product, namely two 1.5 kbp DNA segments connected to each other via an internal DNA “bridge” anchored 58 bp from the closest DNA ends. Finally, one of the 1.5 kbp fragments is

ligated to a biotin-labeled dsDNA fragment (at the extremity distal to the anchor) while for the other 1.5 kbp fragment is ligated to a digoixgenin-labeled dsDNA fragment (at the extremity distal to the anchor). The construct, assembled to a micron-size magnetic bead at the biotin-labeled end and a treated glass coverslip at the dig-labeled end can then be handled and observed in a magnetic trap (˜30-50 constructs per field-of-view). The ends can be processed into blunt or overhang ends by restriction digest either after ligation to the biotin- and dig-labeled fragments, or directly under the magnetic trap.

Oligonucleotides DBCO-Hyb1 5' **Y**CCATGGGCATACTGATCGGTAGGG and DBCO-Hyb2 5' **Y**GAGCCAAGACGCCTCCATCCATGCA, where Y indicates DBCO-dT (courtesy of John Randolph, Glen Research, USA), and oligos Az-Charomid-1456-SmaI (5' GAGAGACCCGGGCACCGTCTCCTTCGAACTTATTCG CAATGGAGTGTCATTCATCAAGGACGCCGC**Z**ATCGCAAATGGTGCTATCC) and Az-Charomid-3778-SmaI (5' GAGAGACCCGGGCACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA GCAGAGCGAGG**Z**ATGTAGGCGGTGCTACAGAG), where Z represents Azido-dT (Trilink, USA) and the underlined sequence is the SmaI/XmaI restriction site, were resuspended to 10 µg/µl in formamide at 37°C. 50 µg each of oligos DBCO-Hyb1 and Az-Charomid-1456 were combined. Similarly, 50 µg each of DBCO-Hyb2 and Az-Charomid-3778 were combined. Combined oligos were incubated 8h at 37°C, and products corresponding to covalently-coupled oligos were separated on a 8M Urea, 8% acrylamide:bisacrylamide (29:1) gel running at ˜10V/cm in 1xTris-borate-EDTA buffer. Gel slices corresponding to product were excised from the gel, using a sacrificial imaging lane to carry out UV shadowing to identify the relevant bands. DNA was purified by first eluting DNA from the crushed gel fragment overnight in 10 mM phosphate buffer pH 8.0 with agitation at 4°C. A SepPak C18 cartridge (Waters) was conditioned with first acetonitrile and then water; the eluate was loaded onto the cartridge which was then washed with water, and the oligo released from the cartridge by eluting with acetonitrile. Acetonitrile was evaporated under vacuum centrifugation, and oligo resuspended to ˜10 µM. The coupled oligos were then used as primers in a PCR reaction using Charomid 9-5 ΔSbfI as template (i.e. a Charomid 9-5 plasmid derivative from which the native SbfI site has been removed by SbfI digestion, fill-in and religation), and ˜3 kbp product DNA was purified using agarose gel electrophoresis and extraction (Macherey-Nagel).

Next, the covalently-coupled oligos which do not participate in the PCR reaction were annealed, respectively, to ssDNA oligos O1-Comp (5' Pho-CGCGCCCTACCGATCAGTATGCCCATGGA, complementary to DBCO-Hyb1 but allowing formation of an AscI overhang compatible with ligation to an MluI overhang as well as an AscI overhang) and O2-Comp (5' Pho-TGGATGGAGGCGTCTTGGCTCA, complementary to DBCO-Hyb2 but allowing formation of an NsiI overhang compatible with ligation to an SbfI overhang as well as an NsiI site) in equimolar ratio at a final concentration of ˜300 nM each for one hour at room temperature in 1x SureCut buffer (New England Biolabs).

Precursor to the "bridge" DNA was obtained by PCR amplifcation of Charomid C9-5 ΔSbfI template with oligos Charo-3600-MluI (5'

GAGAGAACGCGTTACCTGTCCGCCTTTCTCCCTTCGGG) and Charo-4230-SbfI (5' GAGAGACCTGCAGGC CTCACTGATTAAGCATTGGTAACTGTCAGACC), digesting with MluI and SbfI, and agarose gel separation and extraction (Macherey-Nagel).

"Break" DNA and precursor to "bridge" DNA prepared as above were combined at ˜130 nM final concentration each in 1x SureCut buffer (New England Biolabs), along with 0.5 µl each of SbfI, MluI-HF, NsiI and AscI, HC-T4 DNA ligase (New England Biolabs), and 1 mM ATP and 1 mM DTT, and left overnight at room temperature. After 20' heat inactivation at 65°C, the DNA was digested with XbaI and SacI. This resulted in a linear ˜3.6 kbp DNA product consisting of the ˜690 bp bridge segment in the center and two linear 1.5 kbp DNA segments at each end anchored to the bridge internally ˜58 bp from a SmaI/XmaI restriction site. This 3.6 kbp fragment was purified by agarose gel separation and extraction (Macherey-Nagel). The DNA was finally ligated to biotin-labelled ˜1 kbp DNA bearing an XbaI site and digoxigenin-labelled ˜1 kbp DNA bearing a SacI site for attachment in the magnetic trap to, respectively, streptavidin-coated magnetic bead and antidigoxigenin-modified glass surface. In this 10 µl ligation reaction conducted in the manufacturer' s recommended conditions (New England Biolabs T4 DNA Ligase buffer 1x), the DNA was at a concentration of 3 nM, the biotin- and dig-labeled fragments were each at 10 nM concentration, and 200 units (0.5 µl) of T4 DNA ligase were used to ligate the DNA for 3h at room temperature before thermally inactivating the ligase.

To complete preparation of the DNA ends, we then carried out overnight digestion of 1 µl of this DNA preparation in a final reaction volume of 10 µl using either 10 units of XmaI (New England Biolabs) or 20 units of SmaI (New England Biolabs) restriction enzyme. Restriction enzyme was heat inactivated as per manufacturer recommendation, and construct was diluted 30-fold to a nominal DNA concentration of 100 pM and stored at -20°C. For experiments conducted using dephosphorylated DNA as a substrate for NHEJ, we dephosphorylated the construct by combining into a 10 µl reaction volume 2 units of antarctic phosphatase (New England Biolabs) and 0.6 fmoles of the ligation product obtained in the previous paragraph. The reaction was allowed to proceed for 1h at 37°C before being thermally inactivated as per manufacturer' s recommendation.

For assembly onto the microscope, the ligation reaction was first diluted sixty-fold to 50 pM nominal concentration of DNA in Tris buffer (10 mM TrisCl pH 8). Then, 0.5 µl of this dilution was mixed with 10 µl of Dynal MyOne C1 magnetic beads. The beads were prepared by taking 10 µl of stock solution, washing them with 100 µl of RB and concentrating the beads, discarding the supernatant, and resuspending the beads in 10 µl of RB. The bead+DNA mixture was diluted after 5-10 seconds with an additional 15 µl of RB and injected onto functionalized surfaces.

Functionalized surfaces consist in two glass coverslips (#1, ˜180 µm thick) separated by two thicknesses of parafilm into which a channel (1mm X 50 mm) has been cut. Both coverslips are functionalized with antidigoxigenin and passivated as described recently [39], but one of the coverslips has 2 mm diameter holes located above each end of the channel and which are used for filling and draining the reagent-filled channel. The surface is mounted on a custom-designed holder and placed atop the oil-immersion objective of a magnetic trap, in which a

pair of high-grade permanent magnets located above the sample can be translated (to vary the force) or rotated (to rotate the bead and supercoil the DNA) to modify the mechanical constraints applied to the DNA. Real-time particle tracking software (PicoJai SARL) allows for videomicroscopy based tracking of the 1-micron magnetic beads with nanometer resolution and in real-time (˜30 Hz).

A typical sample consists in a microscope field-of-view containing ˜30-50 individual DNA tethers which can be monitored simultaneously. These tethers are systematically verified to display an appropriate change in DNA extension when the applied force is varied from F_{low}=0.04 pN to F_{high}=1.4 pN, and when necessary are further verified to not change extension when supercoiling is changed.

Proteins

Constructs—Ku70 and Ku80 were both cloned into pFastBac Dual vector (Thermo Fisher) and co-expressed in insect Sf9 cells. Ku70 contains an N-terminal hexahistidine tag followed by a TEV cleavage site. XLF and XRCC4 were individually cloned into pHAT5 vector [40] and transformed into Rosetta™2 (DE3) cells (Invitrogen). Both XLF and XRCC4 contain a C-terminal hexahistidine tag. Cell cultures were grown in LB medium until OD_{600} was approximately 0.6. IPTG was added to a final concentration of 1 mM. Proteins were expressed at 16°C overnight for XLF and at 37°C for 4 hours for XRCC4. The PAXX and full-length LigaseIV-XRCC4 co-expression constructs are as described previously [11, 41]. XLF_{1-233} (C-terminal truncation mutant, which can not bind to Ku70/80) was cloned into pETG-41A (Gateway™ Destination vector, EMBL) and expressed as descripted before [42]. The catalytically-dead LigaseIV(K273A)-XRCC4 protein complex was generated from the LigaseIV-XRCC4 co-expression plasmid (A gift from Prof. Ming-Daw Tsai) by the QuikChange method (Agilent Technologies). Full-length Ligase IV fused with a hexahistidine tag at the C-terminus and ccXRCC4 (residue 138-213) were amplified from the Ligase IV-XRCC4 co-expression plasmid and an XRCC4 plasmid, of which cysteines were mutated to alanines, respectively and cloned into pRSFDuet1 vector (Novagen).

Protein purifications—Ku70/80

Insect Sf9 cells containing Ku70/80 were lysed by sonication in lysis buffer (50 mM Tris pH 8.0, 5% glycerol, 150 mM NaCl, 2 mM β-mercaptoethanol, 20 mM imidazole, protease inhibitor (Roche) and 40 μg/ml Deoxyribonuclease I (Sigma)). After 30 minutes incubation in 4°C after sonication, the salt concentration of the lysate was adjusted to 500 mM NaCl. After centrifugation, Ku70/80 supernatant was mixed with Ni-NTA affinity resin (Qiagen) pre-equilibrated with binding buffer (50 mM Tris pH 8.0, 5% glycerol, 500 mM NaCl, 2 mM β-mercaptoethanol, and 20 mM imidazole). After washing beads with binding buffer for 10x column volume, protein was eluted using elution buffer (50 mM Tris pH8.0, 5% glycerol, 500 mM NaCl, 2 mMβ-mercaptoethanol, and 100 mM imidazole). The N-terminal hexahistidine tag of Ku70 was cleaved by TEV protease. Uncleaved Ku70/80 and TEV (also hexahistidine-tagged) were removed by running through a gravity column containing Ni-NTA resin. The flow-through containing Ku70/80 was dialysed with Q column (GE

healthcare) buffer A (20 mM Tris pH8, 50 mM NaCl, 5% glycerol, and 5 mM DTT) and loaded onto the column. Protein was eluted in a gradient against buffer B (20 mM Tris pH8, 1M NaCl, 5% glycerol, and 5 mM DTT). Eluted sample was further purified by running through Superdex 200 10/300 (GE Healthcare) equilibrated in Buffer GF (20 mM Tris pH8, 150 mM NaCl, 5% glycerol, and 5 mM DTT). Protein samples were analysed on 4-12% NuPAGE® Bis-Tris gels, concentrated and stored in -80°C.

XLF and XRCC4

XLF cells were purified by a similar process as for Ku70/80. The lysis buffer for XLF is 20mM HEPES pH8.0, 2M NaCl, 10mM imidazole, 2mM β-Mercaptoethanol and protease inhibitor). No cleavage was carried out for the elution sample from after Ni affinity purification. The Q column for XLF is buffer A (20 mM HEPES pH 8.0, 10 mM NaCl, and 5 mM DTT) and buffer B (20mM HEPES pH8.0, 1M NaCl, and 5 mM DTT). The final gel filtration buffer is 20 mM HEPES pH 8.0, 150 mM NaCl, and 5 mM DTT. Purification of XRCC4 was performed as described previously [43]. Purification of XLF_{1-233} was identical to that of wild-type XLF, except that the mutant contains an N-terminal 6-His-MPB tag followed by aTEV cleavage site. After Ni affinity purification, the tag was cleaved before proceeding to the Q column.

Ligase IV-XRCC4, PAXX and DNA-PKcs

Purifications of all Ligase IV-XRCC4 constructs, PAXX and DNA-PKcs were performed as described previously [5, 11, 41].

Bulk Ligation of DNA with overhang ends by XRCC4-LigIV System

pHAT4 was double-digested with XhoI and NcoI to generate four base-pair overhangs. 200 ng of the digested plasmid was incubated with 25 nM of the indicated proteins in 20 ml of reaction buffer containing 25 mM Tris-HCl pH 7.5, 150 mM KCl, 1 mM MgCl2, 1 mM DTT, 10 μM ATP, 10% (w/v) PEG10,000, and 10 μg/ml BSA. Mixtures were incubated at 37 °C for 5 min before initiating ligation by XRCC4/LIG4 at 37 °C for 30 min. The mixtures were incubated at 50 °C for another 30 min after adding 2 μl of a reaction-stop solution (100 mM EDTA, 0.1% (w/v) SDS) and 0.2 μl of 20 mg/ml Proteinase K. Reaction mixtures were separated by electrophoresis on a 0.8% agarose gel in TBE buffer. The gel was stained with SYBR Gold, visualized using a UV imager and quantified using GeneTools (Synegene). We note that the presence of four base-pair overhangs makes this assay more permissive than the blunt ends used in the scaffold assay.

Experimental Conditions

All assays were conducted at 34°C in reaction buffer RB (20 mM K·Hepes pH 7.8, 100 mM KCl, 5 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 0.05% Tween-20, and 0.5 mg/ml BSA), unless noted otherwise. When present, NHEJ components were used at a concentration of 10 nM Ku70/80 (dimer), 100 pM DNA-PKcs, 20 nM PAXX, 20 nM XRCC4, 20 nM XLF, and 20 nM Ligase IV. When mutated versions of proteins were used, they were at the same concentration as wild-type proteins.

To digest DNA molecules which have undergone NHEJ and repair under the magnetic trap, we first disrupt any remaining components of the repair machinery by rincing the capillary with Wash Buffer supplemented with 0.2% SDS (Wash Buffer: 20 mM Tris·Cl pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 0.05% Tween-20, and 0.5 mg/ml BSA), then rinsing the capillary with Wash Buffer, then infusing the capillary with 100 μl of RB containing either 20 units of XmaI or 10 units of SmaI (New England Biolabs). Data presented in histograms of changes in DNA extension upon NHEJ repair or T4-based ligation are obtained by comparing DNA extension before and after cleavage.

Data Collection and Analysis

Single-molecule magnetic trapping data giving DNA extension as a function of time and under force-modulation cycles was obtained using the Picotwist software suite (Picotwist S.A.R.L). Magnetic trapping is carried out as has been described in the literature. In the magnetic trap used the magnetic bead z-position was tracked at 31 Hz to ˜5 nm resolution for F = 1.4 pN.

Unless noted otherwise, force was repeatedly alternated between low (F_{low} = 0.04 pN) and high (F_{high}=1.4 pN) with a total cycle time of typically 450 seconds (˜225 seconds at F_{low} followed by ˜225 seconds at high force). For experiments presented in Supp. Fig. 6 and Supp. 2 the total cycle time was about 50 seconds (25 seconds at low force followed by 25 seconds at high force).

Time-traces of DNA extension vs. time were analyzed to determine the lifetime ($t_{synapsis}$) and extension change (Δl) upon rupture of individual interaction events observed on the scaffold after the extending force had been increased from low to high force.

Finally, to determine the frequency of synaptic events, we included all events with expected amplitude-change signature for specific end-end interactions (i.e. within three standard deviations of the mean, or Δl between 130 and 200 nm). In this calculation we also accounted for apparently irreversible events characteristic of ligation, as well as very long-lived events which end up reversing after apparently multiple traction cycles.

Events for which rupture did not take place within the same traction cycle in which the event was first observed, but instead ruptured in later traction cycles, could not be reliably assessed for continuity or scored for $t_{synapsis}$ and were therefore excluded from ($t_{synapsis}$, Δl) analysis. Such long-lived events lasting thousands of seconds were observed in the presence of downstream components (XLF, XRCC4 and Ligase IV), and in particular in the “complete” reaction for which they represent approximately half of all events. When any one of the downstream components was removed, these long-lived events spanning at least two traction cycles were no more than 10% of all observed events. No such long-lived events were observed when only upstream components (Ku, DNA-PKcs, and PAXX) were present. We note that the extended lifetime of such events precludes their detection by fluorescence-based methods which are essentially unable to access such timescales. We further note that the amplitude of such long-lived events was systematically within the expected range for bonafide synaptic events, but the very long-lived nature of these events suggests an alternative conformation or state of one of the NHEJ components employed.

Statistics and Reproducibility

Data were collected from at least 50-100 monitored DNA molecules observed over the course of at least two independent experimental replicates, and events were distributed according to a Poisson law among DNA molecules, with ~1-3 events per DNA molecule generating interactions (with typically at least ~30% of the 50-100 DNA molecules monitored reporting interactions).

Data Availability Statement

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

J.L.W is supported by a PhD scholarship from the China Scholarship Council; C.D. is supported by a PhD scholarship from the University of Paris V and the Frontieres du Vivant doctoral program. For funding and support we thank the Ligue Nationale Contre le Cancer “Equipe Labellisée” program and PSL University (NanoRep grant) as well as the University of Paris VII, the CNRS, and the Ecole normale supérieure (to T.R.S); the National Institute of Health (PO1CA92584 to S.P.L.-M. and J.A.T and R01GM110387 to S.E.T); and the Wellcome Trust for support for T.O. and Q.W. (093167MA and 200814/Z/16/Z to T.L.B). J.A.T. is partly supported by a Robert A. Welch Chemistry Chair, the Cancer Prevention and Research Institute of Texas, and the University of Texas System Science and Technology Acquisition and Retention.

References

[1]. Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature. 2001; 412:607–614. [PubMed: 11493912]

[2]. Smith GC, Jackson SP. The DNA-dependent protein kinase. Genes Dev. 1999; 13:916–934. [PubMed: 10215620]

[3]. Dvir A, Stein LY, Calore BL, Dynan WS. Purification and characterization of a template associated protein kinase that phosphorylates RNA polymerase II. J Biol Chem. 1993; 268:10440–10447. [PubMed: 8486698]

[4]. Carter T, Vancurova I, Sun I, Lou W, DeLeon SA. DNA-activated protein kinase from HeLa cell nuclei. Mol Cell Biol. 1990; 10:6460–6471. [PubMed: 2247066]

[5]. Sibanda BL, Chirgadze DY, Blundell TL. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. Nature. 2010; 463:118–121. [PubMed: 20023628]

[6]. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S, Abolfath RM, et al. Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. J Biol Chem. 2010; 285:1414–1423. [PubMed: 19893054]

[7]. Jiang W, Crowe JL, Liu X, Nakajima S, Wang Y, Li C, Lee BJ, Dubois RL, Liu C, Yu X, Lan L, et al. Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. Molecular Cell. 2015; 58:172–185. [PubMed: 25818648]

[8]. Sibanda BL, Chirgadze DY, Ascher DB, Blundell TL. DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. Science (New York, N.Y.). 2017; 355:520–524.

[9]. Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezcan I, Sanal O, Bertrand Y, Philippe N, Fischer A, et al. Artemis, a novel DNA double-strand break

repair/V(D)J recombination protein, is mutated in human severe combined. Cell. 2001; 105:177–186. [PubMed: 11336668]

[10]. Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell. 2002; 108:781–794. [PubMed: 11955432]

[11]. Ochi T, Blackford AN, Coates J, Jhujh S, Mehmood S, Tamura N, Travers J, Wu Q, Draviam VM, Robinson CV, Blundell TL, et al. DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. Science. 2015; 347:185–188. [PubMed: 25574025]

[12]. Xing M, Yang M, Huo W, Feng F, Wei L, Jiang W, Ning S, Yan Z, Li W, Wang Q, Hou M, et al. Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway. Nat Commun. 2015; 6:6233. [PubMed: 25670504]

[13]. Craxton A, Somers J, Munnur D, Jukes-Jones R, Cain K, Malewicz M. XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair. Cell Death Differ. 2015; 22:890–897. [PubMed: 25941166]

[14]. Tadi SK, Tellier-Lebegue C, Nemoz C, Drevet P, Audebert S, Roy S, Meek K, Charbonnier JB, Modesti M. PAXX is an accessory c-NHEJ factor that associates with Ku70 and has overlapping functions with XLF. Cell Rep. 2016; 17:541–555. [PubMed: 27705800]

[15]. Ropars V, Drevet P, Legrand P, Baconnais S, Amram J, Faure G, Marquez JA, Pietrement O, Guerois R, Callebaut I, Le Cam E, et al. Structural characterization of filaments formed by human XRCC4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. Proc Natl Acad Sci USA. 2011; 108:12663–12668. [PubMed: 21768349]

[16]. Hammel M, Rey M, Yu Y, Mani RS, Classen S, Liu M, Pique ME, Fang S, Mahaney BL, Weinfeld M, Schriemer DC, et al. XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair. J Biol Chem. 2011; 286:32638–32650. [PubMed: 21775435]

[17]. Wu Q, Ochi T, Matak-Vinkovic D, Robinson CV, Chirgadze DY, Blundell TL. Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions. Biochem Soc Trans. 2011; 39:1387–92. [PubMed: 21936820]

[18]. Andres SN, Vergnes A, Ristic D, Wyman C, Modesti M, Junop M. A human XRCC4-XLF complex bridges DNA. Nucleic Acids Res. 2012; 40:1868–1878. [PubMed: 22287571]

[19]. Brouwer I, Sitters G, Candelli A, Heerema SJ, Heller I, de Melo AJ, Zhang H, Normanno D, Modesti M, Peterman EJG, Wuite GJL. Sliding sleeves of XRCC4-XLF bridge DNA and connect fragments of broken DNA. Nature. 2016; 535:566–569. [PubMed: 27437582]

[20]. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell. 2006; 124:301–313. [PubMed: 16439205]

[21]. Mahaney BL, Hammel M, Meek K, Tainer JA, Lees-Miller SP. XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair 1. Biochem Cell Biol. 2013; 91:31–41. [PubMed: 23442139]

[22]. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010; 79:181–211. [PubMed: 20192759]

[23]. Liang S, Esswein SR, Ochi T, Wu Q, Ascher DB, Chirgadze D, Sibanda BL, Blundell TL. Achieving selectivity in space and time with DNA double-strand-break response and repair: molecular stages and scaffolds come with strings attached. Struct Chem. 2017; 28:161–171.

[24]. Yaneva M, Kowalewski T, Lieber MR. Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. The EMBO journal. 1997; 16:5098–5112. [PubMed: 9305651]

[25]. West RB, Yaneva M, Lieber MR. Productive and nonproductive complexes of Ku and DNA-dependent protein kinase at DNA termini. Mol Cell Biol. 1998; 18:5908–5920. [PubMed: 9742108]

[26]. Strick T, Allemand JF, Bensimon D, Bensimon A, Croquette V. The elasticity of a single supercoiled DNA molecule. Science. 1996; 271:1835–1837. [PubMed: 8596951]

[27]. Bouchiat C, Wang MD, Block SM, Allemand J-F, Strick TR, Croquette V. Estimating the persistence length of a Worm-Like Chain Molecule from Force-extension Measurements. Biophys J. 1999; 76:409–413. [PubMed: 9876152]

[28]. Yano K-I, Morotomi-Yano K, Wang S-Y, Uematsu N, Lee K-J, Asaithamby A, Weterings E, Chen DJ. Ku recruits XLF to DNA double-strand breaks. EMBO Rep. 2008; 9:91–96. [PubMed: 18064046]

[29]. Grundy GJ, Rulten SL, Arribas-Bosacoma R, Davidson K, Kozik Z, Oliver AW, Pearl LH, Caldecott KW. The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins. Nat Commun. 2016; 7:11242. [PubMed: 27063109]

[30]. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA. XLF regulates filament architecture of the XRCC4-Ligase IV complex. Structure (London, England : 1993). 2010; 18:1431–1442.

[31]. Balmus G, Barrons AC, Wijnhoven PWG, Lescale C, Hasse HL, Boroviak K, le Sage C, Doe B, Speak AO, Galli A, Jacobsen M, et al. Synthetic lethality between PAXX and XLF in mammalian development. Genes Dev. 2016; 30:2152–2157. [PubMed: 27798842]

[32]. Liu X, Shao Z, Jiang W, Lee BJ, Zha S. PAXX promotes ku accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice. Nat Commun. 2017; 8:13816. [PubMed: 28051062]

[33]. Walther AP, Gomes XV, Lao Y, Lee CG, Wold MS. Replication protein A interactions with DNA. 1. functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. Biochemistry. 1999; 38:3963–3973. [PubMed: 10194308]

[34]. Hammel M, Yu Y, Radhakrishnan SK, Chokshi C, Tsai M-S, Matsumoto Y, Kuzdovich M, Remesh SG, Fang S, Tomkinson AE, Lees-Miller SP, et al. An intrinsically disordered APLF links Ku, DNA-PKcs and XRCC4-DNA ligase iv in an extended flexible non-homologous end joining complex. J Biol Chem. 2016

[35]. Lees-Miller SP, Beattie TL, Tainer JA. Noncoding RNA joins Ku and DNA-PKcs for DNA-break resistance in breast cancer. Nat Struct Mol Biol. 2016; 23:509–510. [PubMed: 27273637]

[36]. Graves ET, Duboc C, Fan J, Stransky F, Leroux-Coyau M, Strick TR. A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. Nat Struct Mol Biol. 2015; 22:452–457. [PubMed: 25961799]

[37]. Bouchiat C, Mezard M. Elasticity model of a supercoiled DNA molecule. Phys Rev Lett. 1998; 80:1556–1559.

[38]. Hammarsten O, Chu G. DNA-dependent protein kinase: DNA binding and activation in the absence of Ku. Proc Natl Acad Sci USA. 1998; 95:525–530. [PubMed: 9435225]

[39]. Duboc C, Fan J, Graves ET, Strick TR. Preparation of DNA substrates and functionalized glass surfaces for correlative nanomanipulation and colocalization (NanoCOSM) of single molecules. Methods Enzymol. 2017; 582:275–296. [PubMed: 28062038]

[40]. Peranen J, Rikkonen M, Hyvonen M, Kaariainen L. T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. Anal Biochem. 1996; 236:371–373. [PubMed: 8660525]

[41]. Ochi T, Wu Q, Chirgadze DY, Grossmann JG, Bolanos-Garcia V-M, Blundell TL. Structural insights into the role of domain flexibility in human DNA ligase IV. Structure (London, England : 1993). 2012 J;20:1212–1222.

[42]. Yi, Li, Chirgadze, DY., Bolanos-Garcia, VM., Sibanda, BL., Davies, OR., Ahnesorg, P., Jackson, SP., Blundell, TL. Crystal structure of human XLF/Cernunnos reveals unexpected differences from XRCC4 with implications for NHEJ. EMBO J. 2008 Jan.27:290–300. [PubMed: 18046455]

[43]. Murray JE, van der Burg M, Ijspeert H, Carroll P, Wu Q, Ochi T, Leitch A, Miller ES, Kysela B, Jawad A, Bottani A, et al. Mutations in the NHEJ component XRCC4 cause primordial dwarfism. Am J Hum Genet. 2015; 96:412–424. [PubMed: 25728776]

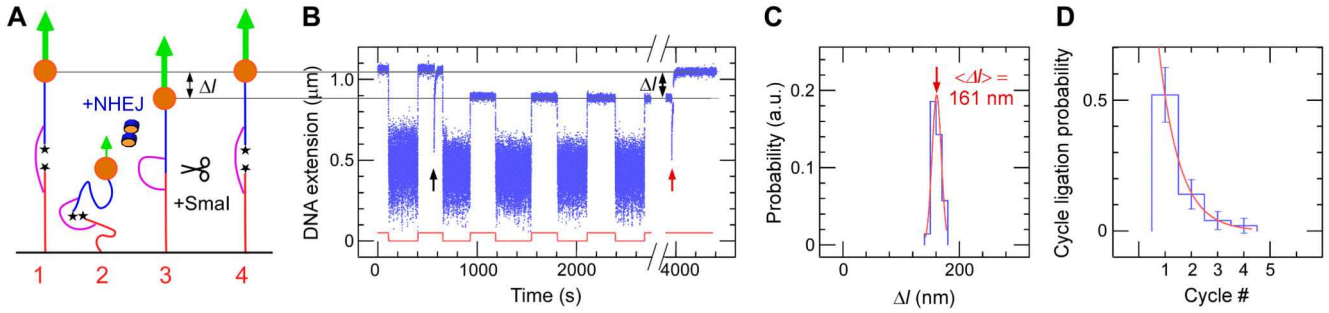


Figure 1. Double-strand break repair by NHEJ proteins at single-molecule resolution

(A) A 600 bp dsDNA segment (magenta) joins two 1.5 kbp dsDNA segments (blue, red), forming a construct in which two blunt ends face each other. Stars represent phosphate groups. The construct is tethered to a treated glass surface and a 1-micron magnetic bead. Magnets located above the sample generate a controlled extending force on the DNA (green arrow), and the DNA end-to-end extension is determined in real-time. Ligation is observed via a series of four steps: (1) at high force a high-extension state is initially observed, (2) the force is lowered allowing the DNA ends to interact, (3) the force is returned to its initial value but if end ligation has occurred the construct cannot recover its initial extension, (4) the initial extension is recovered upon specific cleavage of repaired DNA. **(B)** Time-trace obtained upon force-modulation (red) in the presence of Ku, DNA-PKcs, PAXX, XLF, XRCC4 and Ligase IV. Initially, DNA extension (blue points) is shown to alternate between a low and a high value upon force modulation. After addition of NHEJ components (black up arrow) the maximum extension displayed by the construct is reduced. After washing the sample with 0.2% SDS (break in time-trace), addition of SmaI (red up arrow) results in an increase, Δl in DNA extension. **(C)** Histogram of Δl values. Red line is a fit to a Gaussian distribution, with a maximum at 161 ± 8 nm (SD, n=28 cleavage events). **(D)** DNA ligation probability per traction cycle. We monitored 50 DNA molecules; of 36 molecules repaired 28 were monitored throughout the cleavage reaction. Red line: single-exponential fitting yields a time constant of ˜0.8 ± 0.2 cycles (SEM, n=36) or ˜175s.

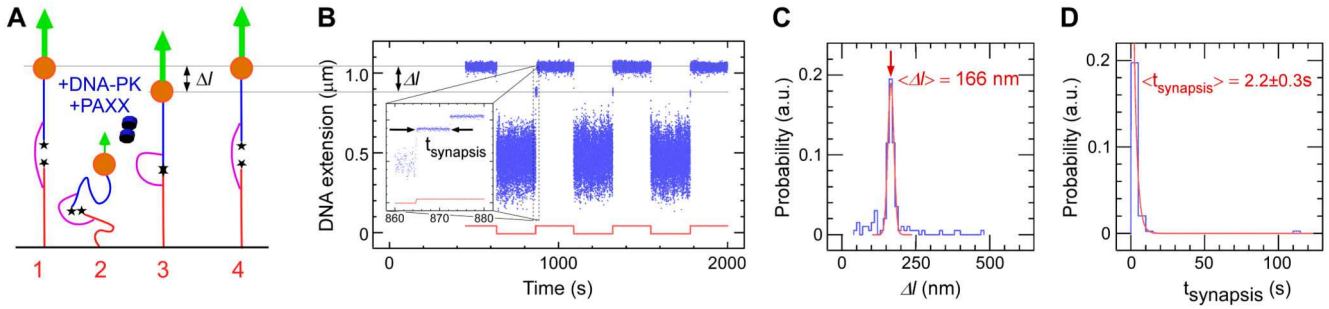


Figure 2. Ku, DNA-PKcs and PAXX are necessary to stabilize DNA-end synapsis.

(A) Experimental design. DNA is prepared with blunt ends using SmaI digest, stars represent phosphate groups. **(B)** Representative time-trace obtained upon application of the force-modulation pattern (red). DNA is prepared with blunt ends using SmaI digest. Inset shows an expanded view of an end-interaction rupture event, which can be characterized by both the change in DNA extension upon rupture, Δl, and the duration of the synaptic event prior to rupture, $t_{synapsis}$. **(C)** Histogram of DNA extension change, Δl, upon rupture event. Red line is a fit to a Gaussian distribution, with a maximum (red arrow) at 166 ± 10 nm (SD). The entire histogram contains n=129 events, of which 98 are within three standard deviations from the peak. **(D)** Lifetime distribution of the synaptic state is fit to a single-exponential distribution (red line), giving a lifetime of 2.2 ± 0.3 s (SEM, n=98).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

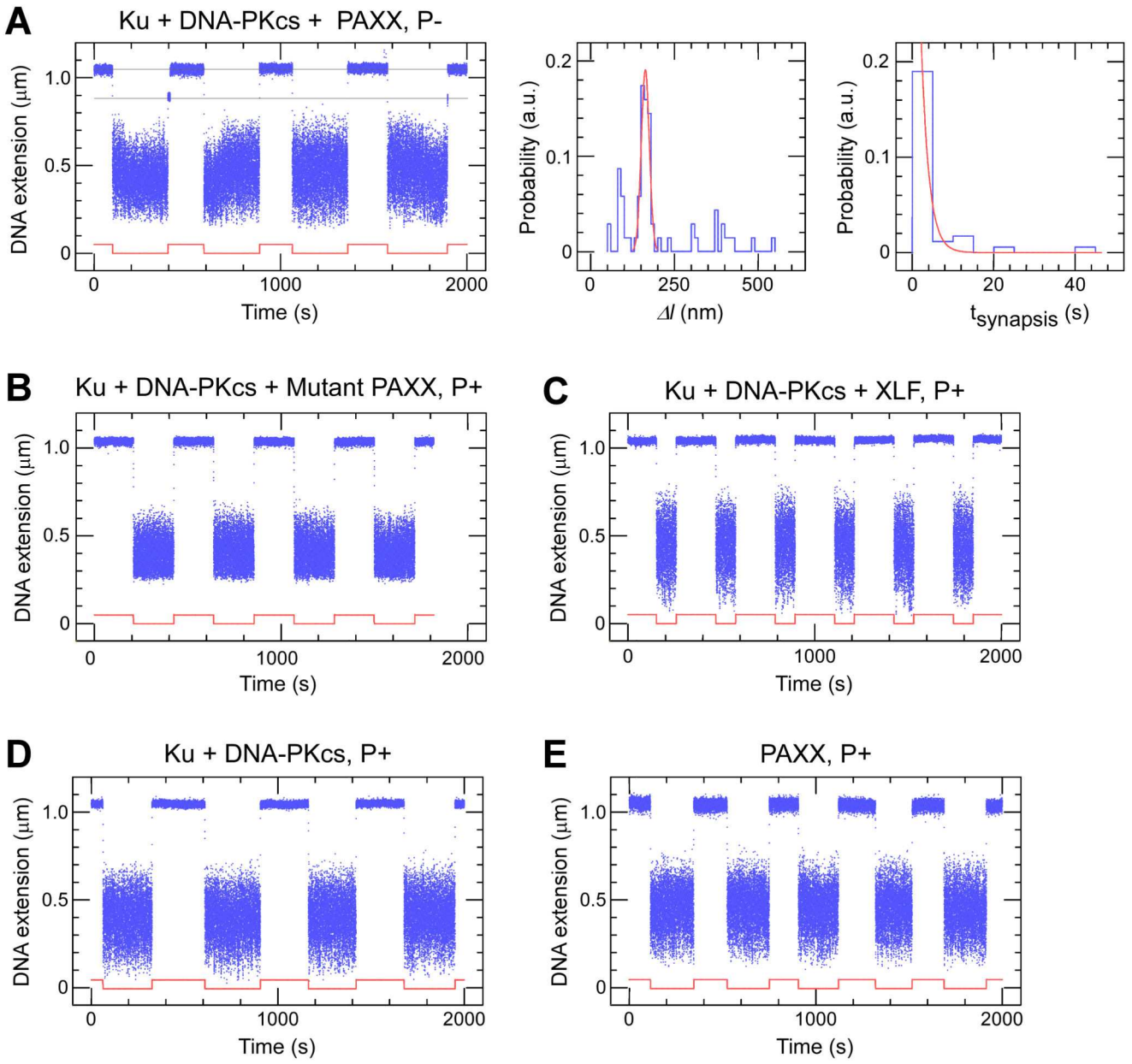


Figure 3. DNA synapsis on the 2-second timescale in the presence of Ku, DNA-PKcs, and PAXX: control experiments.

(A) Synapsis by Ku, DNA-PKcs and PAXX does not require phosphorylated DNA ends. **(Left)** Time-trace showing rupture events. **(Middle)** Amplitude distribution of rupture events (n=69). Distribution peak is fit to a Gaussian (red line) with a maximum at 163 ± 2 nm (SEM; σ=12 nm; n=40). The distribution displays additional density for *ΔI* values of 75 nm and 400 nm. Values of 400 nm are consistent with the extension of a 1500 bp DNA segment at the force employed, and thus likely correspond to "loop-back" interactions between the tip of a 1500 bp DNA segment and the surface to which that segment is anchored. The smaller peak is consistent with local bending or wrapping deformations of DNA with ˜50 nm persistence length [37]. **(Right)** Lifetime distribution of synaptic events follows a single-

exponential distribution (red line) with a mean of 1.9±0.5s (SEM, n=40). **(B-E)** Representative time-traces show that the combinations of **(B)** Ku, DNA-PKcs, and the PAXX mutant; **(C)** Ku, DNA-PKcs, and XLF; **(D)** Ku and DNA-PKcs; and **(E)** PAXX alone do not lead to 2-second synapsis. The well-known contamination of DNA-PKcs preparations by Ku [25, 38] means it is not formally possible to test DNA-PKcs alone.

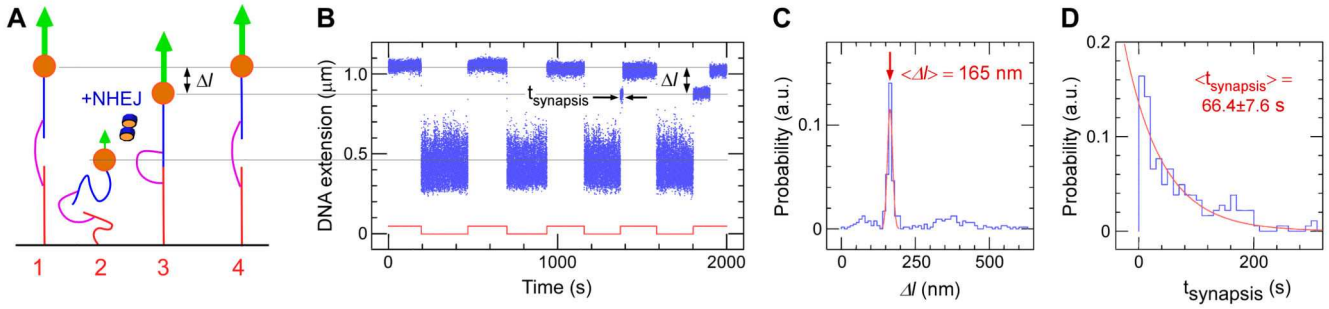


Figure 4. Maximum stabilization of DNA end synapsis by Ku, DNA-PKcs, PAXX, XLF, XRCC4 and Ligase IV.

(A) Experimental design. DNA is prepared with blunt ends using SmaI digest and then dephosphorylated (see Materials & Methods). **(B)** Representative time-trace obtained upon application of the force-modulation pattern (red) in the presence of Ku, DNA-PKcs, PAXX, XLF, XRCC4 and Ligase IV. Transient synapsis is observed as an intermediate DNA extension state, which spontaneously reverts to the maximum extension, allowing us to characterize the duration of the intermediate state and the change in extension upon reversion to the maximum extension. **(C)** Histogram of DNA extension change (Δl) upon rupture event. Red line is a fit to a Gaussian distribution, with a maximum (red arrow) at 165 ± 9 nm (SD, n=324 events in the histogram, of which 183 within three standard deviations from the peak). **(D)** Lifetime distribution of the synaptic state is fit to a single-exponential distribution (red line), giving a lifetime of 66.4 ± 7.6 s (SEM, n=183).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

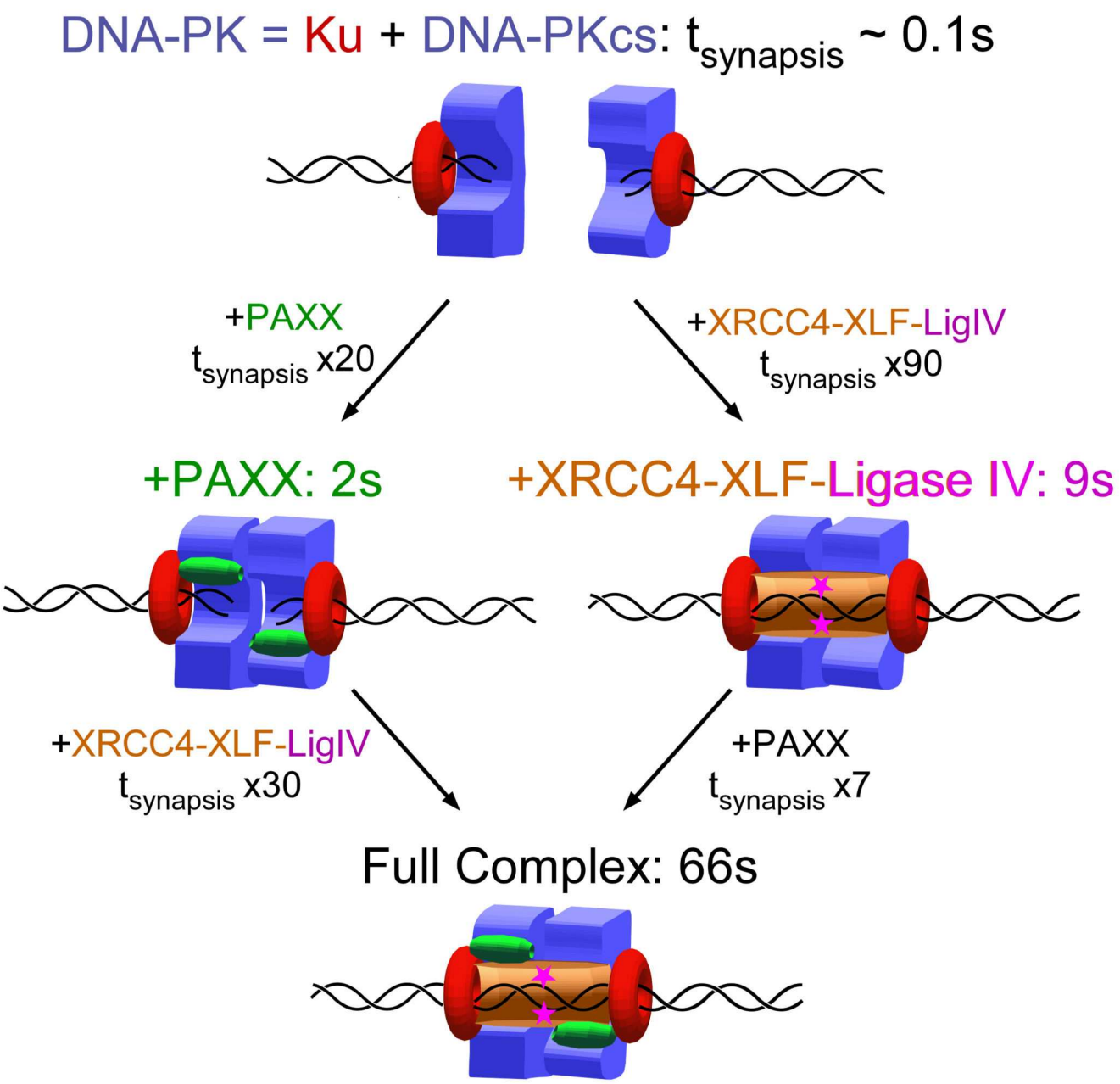


Figure 5. Model for multivalent stabilization of DNA-end synapsis by NHEJ machinery.

Schematic model for synapsis and repair involves (top) an initial synaptic complex formed by two DNA-PK holoenzymes but with a lifetime in the range of hundreds of milliseconds and which can be stabilized by incorporation of PAXX (left path) or XRCC4–XLF–Ligase IV (right path). A complete complex stabilized by both PAXX and XRCC4–XLF–Ligase IV has the longest lifetime (bottom) and leads to efficient ligation of the DSB.

Europe PMC Funders Author Manuscripts

Table 1

Efficiency of observed NHEJ ligation reactions as a function of NHEJ components present.

Ligation events were identified as an irreversible shortening of the DNA scaffold' s extension by Δl ˜ 165 nm. For the complete reaction, 28 of the 36 molecules scored for repair were further confirmed by DNA cleavage via SmaI digest (Fig. 1C); the remaining 8 molecules were lost from tracking during SDS washes which can destabilize streptavidin and antidigoxigenin links between DNA and surfaces)

Assay	# molecules	Number of molecules ligated at # pulling cycle									Total	Efficiency %	Normalized Efficiency %
		1st	2nd	3rd	4th	5th	6th	7th	8th	9th			
Complete reaction	50	26	7	2	1	N/A	N/A	N/A	N/A	N/A	36	72 ± 12	100
Complete but for DNA-PKcs	40	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A	0	0	0
Complete but for Ku	69	0	1	0	1	N/A	N/A	N/A	N/A	N/A	2	3 ± 2	4
Complete but for PAXX	57	6	5	1	1	1	0	1	0	1	16	28 ± 7	39
Complete but for XLF	184	0	1	0	1	1	3	1	2	0	9	5 ± 2	7
Complete but for XRCC4 NTD	193	1	0	1	1	0	0	0	0	0	3	2 ± 1	3
Complete but for Ligase IV	102	0	0	0	0	0	0	0	0	0	0	0	0

Nat Struct Mol Biol. Author manuscript; available in PMC 2018 November 21.

Published in final edited form as:
Science. 2006 November 17; 314(5802): 1139–1143. doi:10.1126/science.1131398.

ABORTIVE INITIATION AND PRODUCTIVE INITIATION BY RNA POLYMERASE INVOLVE DNA SCRUNCHING:

Single-molecule-nanomanipulation experiments--in which magnetic tweezers are used to monitor RNA-polymerase-induced DNA unwinding with ~1 bp resolution and ~1 s temporal resolution--show that both abortive initiation and productive initiation by RNA polymerase involve DNA "scrunching."

Andrey Revyakin[1,2,†], **Chenyu Liu**[1,2,3], **Richard H. Ebright**[1,*], and **Terence R. Strick**[2,3,*]

[1]Howard Hughes Medical Institute, Waksman Institute, and Department of Chemistry, Rutgers University, Piscataway NJ 08854, USA

[2]Cold Spring Harbor Laboratory, Cold Spring Harbor NY 11724, USA

[3]Institut Jacques Monod, Centre National de la Recherche Scientifique UMR7592, 2 Place Jussieu 75251 Paris Cedex 05, France

Abstract

Using single-molecule DNA nanomanipulation, we show that abortive initiation involves DNA "scrunching"--in which RNA polymerase (RNAP) remains stationary and unwinds and pulls into itself downstream DNA--that scrunching requires RNA synthesis, and that scrunching depends on RNA length. We show further that promoter escape involves scrunching, and that scrunching occurs in most or all instances of promoter escape. Our results support existence of an obligatory stressed intermediate, with ~1 turn of additional DNA unwinding, in escape and are consistent with the proposal that stress in this intermediate provides the driving force to break RNAP-promoter and RNAP-initiation-factor interactions in escape.

Transcription initiation involves a series of reactions (1-3). RNA polymerase (RNAP) binds to promoter DNA to yield an RNAP-promoter closed complex (RP_c). RNAP then unwinds ~1 turn of DNA surrounding the transcription start site to yield an RNAP-promoter open complex (RP_o). RNAP then enters into abortive cycles of synthesis and release of short RNA products as an RNAP-promoter initial transcribing complex (RP_{itc}) and, upon synthesis of an RNA product ~9-11 nt in length, escapes the promoter and enters into productive synthesis of RNA as an RNAP-DNA elongation complex (RD_e).

The mechanism by which the RNAP active-center translocates in abortive initiation and promoter escape has remained problematic. The problem has been posed by two seemingly contradictory observations: First, RNA products up to ~8-10 nt in length are synthesized in abortive initiation (4-6); thus the RNAP active center translocates relative to DNA in abortive initiation. Second, DNA-footprinting results indicate that there upstream boundary of the DNA

---

*Corresponding authors Phone: +33-1-4427-8175 (T.S.), 721-445-5179 (R.H.E.) Fax: +33-1-4427-5716 (T.S.); 732-445-5735 (R.H.E.) strick@ijm.jussieu.fr (T.S.); ebright@waksman.rutgers.edu (R.H.E.).
†Current address: Howard Hughes Medical Institute and Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720-3204, USA

Supporting Online Material www.sciencemag.org Materials and Methods Figs. S1-S15

segment protected by RNAP is the same in RP_o and in RP_{itc} engaged in abortive synthesis (7-10); thus RNAP appears not to translocate relative to DNA in abortive initiation. To reconcile the observation that the RNAP active center translocates in abortive initiation with the observation that RNAP appears not to translocate in abortive initiation, three models have been proposed: (i) "scrunching," which invokes contraction of DNA, (ii) "inchworming," which invokes expansion of RNAP, and (iii) "transient excursions" which invokes transient cycles of forward and reverse RNAP translocations with long intervals between cycles (Fig. 1A;4,7,9-12; see also proposals for structurally unrelated single-subunit RNAP derivatives in 13-19).

In previous work, we have developed a single-molecule-DNA-nanomanipulation approach that detects RNAP-dependent DNA unwinding with ~1 bp resolution and ~1 s temporal resolution, and we have applied this approach to detect and characterize RNAP-dependent promoter unwinding upon formation of RP_o (Figs. 1B, S1, 20-22). In this work, we have applied this approach to test the scrunching model for RNAP-active-center translocation in abortive initiation and promoter escape (Fig. 1A;4,7,11,12). The scrunching model--and only the scrunching model--postulates changes in RNAP-dependent DNA unwinding during abortive initiation and promoter escape. Specifically, the scrunching model postulates that RNAP pulls into itself downstream DNA; for each base pair that RNAP pulls into itself, a base pair must be broken and must be maintained broken, and, correspondingly there must be one base pair of additional DNA unwinding.

Our first set of experiments addressed abortive initiation occurring in complexes engaged in iterative abortive initiation [complexes prepared using subsets of nucleoside triphosphates (NTPs) insufficient to permit promoter escape and productive initiation]. Our primary experimental system was the T5 N25 promoter, a classic model system for analysis of abortive initiation and promoter escape (11; Fig. S2). The initial transcribed sequence of the N25 promoter has no C or G residues in the first 8 nt (Fig. S2); therefore, upon preparation of RP_o at the N25 promoter and addition of ATP and UTP, one obtains RP_{itc} engaged in iterative abortive synthesis of RNA products up to 8 nt in length ($RP_{itc,\leq 8}$; Fig. S2). We also analyzed the N25A5C promoter, a derivative of the N25 promoter that has an altered initial-transcribed-region sequence (Fig. S3). With the N25A5C promoter, with appropriate NTP subsets, one obtains RP_{itc} engaged in iterative abortive synthesis of products up to 4 nt in length ($RP_{itc,\leq 4}$) and RP_{itc} engaged in iterative abortive synthesis of products up to 8 nt in length ($RP_{itc,\leq 8}$) (Fig. S3).

To determine whether scrunching occurs in abortive initiation, we quantified RNAP-dependent DNA unwinding at the N25 promoter in RP_o (0 NTPs) and in $RP_{itc,\leq 8}$ (ATP+UTP) (Fig. 2A,D). Comparison of the amplitudes of transitions in RP_o and in $RP_{itc,\leq 8}$ in representative single-molecule time traces indicates that the amplitude of transitions is greater in $RP_{itc,\leq 8}$ (Fig 2A, left). Comparison of histograms shows clearly that the amplitude of transitions is greater in $RP_{itc,\leq 8}$ (Fig 2A, right). By combining data obtained with positively supercoiled DNA and data obtained with negatively supercoiled DNA, we can establish unequivocally that there is a difference in DNA unwinding in RP_o and $RP_{itc,\leq 8}$, and we can extract the extent of the difference in DNA unwinding in RP_o and $RP_{itc,\leq 8}$: namely, upon transition from RP_o to $RP_{itc,\leq 8}$, there is an increase in unwinding of 5±1 bp (Fig. 2D). Consistent with these results, we find that, upon transition from RP_o to $RP_{itc,\leq 8}$ at the N25 promoter, the downstream boundary of the potassium-permanganate-sensitive promoter region shifts from position +2 to position +8, implying an increase in DNA unwinding of 6 bp (Fig. S4). We conclude that scrunching occurs in abortive initiation.

To determine whether scrunching requires RNA synthesis, we performed a control experiment in which we provided only the initiating nucleotide, ATP ($RP_{itc,\leq 1}$; Figs. 2B,D, S2, S5). In this

case, scrunching was not observed (Fig, 2B,D). We also performed a control experiment in which we provided ATP, UTP, and rifampicin--an inhibitor that blocks synthesis of RNA products >2 nt in length ($RP_{itc,\leq 2}$; Figs. 2C,D, S2, S5). In this case also, scrunching was not observed (Fig. 2C,D). We conclude that scrunching requires RNA synthesis, and, more particularly, that scrunching requires synthesis of an RNA product >2 nt in length.

To determine whether the extent of scrunching correlates with RNA length, we quantified RNAP-dependent DNA unwinding at the N25A5C promoter in RP_o (no NTPs), $RP_{itc,\leq 4}$ (ATP +UTP), and $RP_{itc,\leq 8}$ (ATP+UTP+CTP) (Figs. 3A,B, S3, S6). We observed successive, step-wise increases in the amplitudes of transitions (Fig. 3A) and in the corresponding extents of DNA unwinding (Fig. 3B). We observed an increase in DNA unwinding of 2±1 bp upon transition from RP_o to $RP_{itc,\leq 4}$ (Fig. 3B), and we observed an increase in DNA unwinding of 5±1 bp upon transition from RP_o to $RP_{itc,\leq 8}$ (Fig. 3B). Within experimental error, the observed increases in unwinding in the preceding experiments and in this experiment agree with the quantitative predictions of the simplest version of the scrunching model, wherein increases in unwinding are predicted to equal N-2, where N is the length of the RNA in nucleotides (Fig. 3C). [In the simplest version of the scrunching model, the RNAP active center is predicted to be able to make an RNA product 2 nt in length without translocation, but to need to translocate, and to scrunch, to make longer RNA products; thus the increase in unwinding is predicted to equal N-2 (Fig. 3C).] We conclude that the extent of scrunching in abortive initiation correlates with RNA length and that it correlates quantitatively as predicted by the simplest model of scrunching.

In each of the preceding experiments, complexes engaged in abortive synthesis and release of RNA products >2 nt in length were observed to be present predominantly in the scrunched state; cycles of transitions between the scrunched state and the unscrunched state having the extent of unwinding in RP_o were not observed (Figs. 2A, S5, S6). We infer that, at the promoters studied, at the saturating NTP concentrations studied, abortive-product synthesis and scrunching are fast relative to abortive-product release and unscrunching, and also are fast relative to the second-scale temporal resolution of our method. Consistent with this inference, recently published results indicate that, at a consensus promoter, at saturating NTP concentrations, the rate-limiting step in abortive initiation is abortive-product release and RNAP-active-center reverse translocation (23).

Our results for abortive initiation receive unequivocal support from the companion paper, which analyzes abortive initiation using an independent single-molecule method: single-molecule fluorescence resonance energy transfer (in which scrunching is detected as a decrease in distance between the DNA segments upstream and downstream of the unwound region; 24).

Our second set of experiments addressed promoter escape in complexes engaged in productive initiation (complexes prepared in the presence of all four NTPs). For these experiments, we used DNA constructs having the N25 promoter, followed by a 400 bp or 100 bp transcribed region, followed by a terminator (N25-400-tR2 and N25-100-tR2; Fig. S7). These constructs allow us to monitor complete transcription cycles in the presence of all four NTPs--monitoring promoter unwinding, promoter escape, elongation, and termination, in real time. In addition, these constructs, due to the presence of a terminator, automatically recycle the DNA molecule, facilitating collection of large data sets (N>100). (After each transcription cycle, RNAP leaves the DNA molecule, rendering the DNA molecule available for the next transcription cycle.) For these experiments, we also used shorter DNA molecules (2 kb vs. 4 kb; Fig S8; see 22). This resulted in a decrease in noise and an increase in spatial and temporal resolution.

To determine whether scrunching occurs in productive initiation, we collected and analyzed single-molecule time traces in experiments with N25-400-tR2 and N25-100-tR2 in the presence of all four NTPs (Figs. 4, S9). Representative single-molecule time traces exhibit series of events, wherein each individual event corresponds to a complete transcription cycle--from promoter unwinding through termination (Figs. 4A, S9; events underscored in black). The events are remarkably uniform in duration and overall form (Figs. 4A, S9). (Struck by the uniformity of duration and form of the events, we refer to these time traces as “EKG strips of transcription.”) We parse each event into four unwinding and rewinding transitions: a first transition from the initial state to a state having the extent of unwinding in RP_o, a second transition to a state having the extent of unwinding in the scrunched RP_{itc}, a third transition to a state having an extent of unwinding comparable, but not identical, to that in the RP_o, and a fourth transition returning to the initial state (Figs. 4A, S9, S10). We assign these four transitions to the formation of RP_o, the formation of RP_{itc} (with scrunching), the formation of RD_e (with reversal of scrunching), and termination. Scrunching occurs in these events (Figs. 4A,B, S9, S10). Scrunching is manifest as the ”overshoot” in unwinding that follows formation of RP_o and that precedes formation of RD_e. The observed extent of scrunching is 9±2 bp (Fig. 4B), which agrees, to the base pair, with the predicted extent of scrunching under the N-2 rule (Fig. 3C). [Promoter escape at N25 occurs upon synthesis of an RNA product having a length, N, of 11 nt (Fig. S2B).] We conclude that scrunching occurs in promoter escape in productive initiation.

Extensive control experiments document the assignment of transitions in the preceding paragraph (Figs. S11-S14). If we perform experiments in the absence of NTPs, yielding RP_o, we observe transition 1 only (Fig. S11A). If we perform experiments with an NTP subset that yields $RP_{itc,\leq 8}$, we observe transitions 1 and 2 only (Fig. S11B). If we perform experiments with an NTP subset that yields a halted elongation complex, $RD_{e,29}$, we observe transitions 1, 2, and 3 only (Fig. S11C). If we add back the remaining NTPs to any of the preceding cases, we observe the full process recapitulated, with full cycles of transcription (Fig. S11A-C). If we omit the terminator, we observe transitions 1, 2, and 3 only (Fig. S12). If we vary the length of the transcribed region, we vary the duration of the phase between transition 3 and transition 4--and we vary it according to a relationship suggesting an elongation rate of ~10 nt/s (which equals the expected elongation rate for the NTP concentration and temperature; 25-28) (Figs. S13, S14). Again, we conclude that scrunching occurs in promoter escape in productive initiation.

To determine whether scrunching occurs in few, many, or all, productive initiation events, we compared the number of transcription cycles that exhibit detectable scrunches to the number of transcription cycles that do not (Fig. 4C). Fully 80% of transcription cycles exhibit a detectable scrunch (Fig. 4C). Thus, most transcription cycles involve scrunching. This percentage, however, represents an underestimate, since the temporal resolution of our approach is insufficient to detect fast scrunches (scrunches that have a duration <1 s). From the observed distribution of scrunch lifetimes, we estimate that 20% of scrunches have a duration <1 s (Fig. 4D). Based on the percentage of transcription cycles that exhibit a detectable scrunch (80%) and the estimated percentage of scrunches that are not detected because they have a duration <1 s (20%), it is apparent that ~100% of transcription cycles involve scrunching. We conclude that scrunching occurs in all, or very nearly all, transcription cycles. We conclude further that scrunching may be obligatory for promoter escape in productive initiation.

Our overall conclusions are as follows: Abortive initiation involves scrunching. Promoter escape involves scrunching. Promoter escape may--and we believe does--involve obligatory scrunching. Promoter escape may--and we believe does--involve an obligatory “stressed intermediate,” as originally suggested two decades ago (9; see also 11).

At a typical promoter, promoter escape occurs only after synthesis of an RNA product ~9-11 nt in length (1-11) and thus can be inferred to require scrunching of ~7-9 bp (N-2 where N = ~9-11; see Fig. 3C). Assuming an energetic cost of base-pair breakage of ~2 kcal/mol/bp (29), it can be inferred that, at a typical promoter, a total ~14-18 kcal/mol of base-pair-breakage energy is accumulated in the stressed intermediate. This free energy is high relative to the free energies for RNAP-promoter interaction (~7-9 kcal/mol for sequence-specific component of RNAP-promoter interaction; 1) and RNAP-initiation-factor interaction (~13 kcal/mol for σ^{70};30). We believe that our results demonstrate the existence of an obligatory stressed intermediate, and we believe that the energy accumulated in that obligatory stressed intermediate is the energy that drives disruption of interactions between RNAP and promoter DNA, that drives disruption of interactions between RNAP and initiation factors, and, thus, that drives the transition from initiation to elongation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Borukhov, L. Hsu, and E. Nudler for plasmids and discussion. This work was supported by Cold Spring Harbor, Centre National de la Recherche Scientifique, Fondation pour la Recherche Médicale, Fondation Fourmentin-Guilbert, European Molecular Biology Organization, City of Paris, and Universities of Paris VI and Paris VII grants to T.R.S. and by NIH grant GM41376 and a Howard Hughes Medical Investigatorship to R.H.E.

Abbreviations

RNAP, RNA polymerase; RP_o, RNAP-promoter open complex; RP_{itc}, RNAP-promoter initial transcribing complex; RD_e, RNAP-DNA elongation complex; NTPs, nucleoside triphosphates.

References and Notes

1. Record, MT.; Reznikoff, W.; Craig, M.; McQuade, K.; Schlax, P. Escherichia coli and Salmonella. Neidhart, F., editor. Vol. vol. 1. ASM Press; Washington, D.C.: 1996. p. 792-820.
2. Young B, Gruber T, Gross C. Cell 2002;109:417–420. [PubMed: 12086598]
3. Murakami K, Darst S. Curr. Opin. Structl. Biol 2003;13:31–39.
4. Carpousis AJ, Gralla JD. Biochem 1980;19:3245–3253. [PubMed: 6996702]
5. Grachev M, Zaychikov E. FEBS Lett 1980;115:23–26. [PubMed: 6156091]
6. Munson L, Reznikoff W. Biochem 1981;20:2081–2085. [PubMed: 6165380]
7. Carpousis A, Gralla J. J. Mol. Biol 1985;183:165–177. [PubMed: 2409292]
8. Spassky A. J. Mol. Biol 1986;188:99–103. [PubMed: 3519983]
9. Straney D, Crothers D. J. Mol. Biol 1987;193:267–278. [PubMed: 2439694]
10. Krummel B, Chamberlin M. Biochem 1989;28:7829–7842. [PubMed: 2482070]
11. Hsu L. Biochim. Biophys. Acta 2002;1577:191–207. [PubMed: 12213652]
12. Pal M, Ponticelli A, Luse D. Mol. Cell 2005;19:101–110. [PubMed: 15989968]
13. Cheetham G, Jeruzalmi D, Steitz T. Nature 1999;399:80–83. [PubMed: 10331394]
14. Cheetham G, Steitz T. Science 1999;286:2305–2309. [PubMed: 10600732]
15. Brieba L, Sousa R. EMBO J 2001;20:6826–6835. [PubMed: 11726518]
16. Jiang M, Rong M, Martin C, McAllister W. J. Mol. Biol 2001;310:509–522. [PubMed: 11439019]
17. Liu C, Martin C. J. Biol. Chem 2002;277:2725–2731. [PubMed: 11694519]
18. Esposito E, Martin C. J. Biol. Chem 2004;279:44270–44276. [PubMed: 15304497]
19. Gong P, Esposito E, Martin C. J. Biol. Chem 2004;279:44277–44285. [PubMed: 15337752]
20. Revyakin A, Allemand J-F, Croquette V, Ebright RH, Strick TR. Meths. Enzymol 2003;370:577–598.

21. Revyakin A, Ebright RH, Strick TR. Proc. Natl. Acad. Sci. USA 2004;101:4776–4780. [PubMed: 15037753]
22. Revyakin A, Ebright RH, Strick TR. Nature Meths 2005;2:127–138.
23. Margeat E, et al. Biophys. J 2006;90:1419–1431. [PubMed: 16299085]
24. Kapanidis A, et al. Science. (in press)
25. Wang M, et al. Science 1998;282:902–907. [PubMed: 9794753]
26. Adelman K, et al. Proc. Natl. Acad. Sci. USA 2002;99:13538–13543. [PubMed: 12370445]
27. Abbondanzieri E, Greenleaf W, Shaevitz J, Landick R, Block S. Nature 2005;438:460–465. [PubMed: 16284617]
28. Abbondanzieri E, Shaevitz J, Block S. Biophys. J 2005;89:L61–L63. [PubMed: 16239336]
29. Breslauer K, Frank R, Blocker H, Marky L. Proc. Natl. Acad. Sci. USA 1986;83:3746–3750. [PubMed: 3459152]
30. Gill S, Weitzel S, von Hippel P. J. Mol. Biol 1991;220:307–324. [PubMed: 1856861]

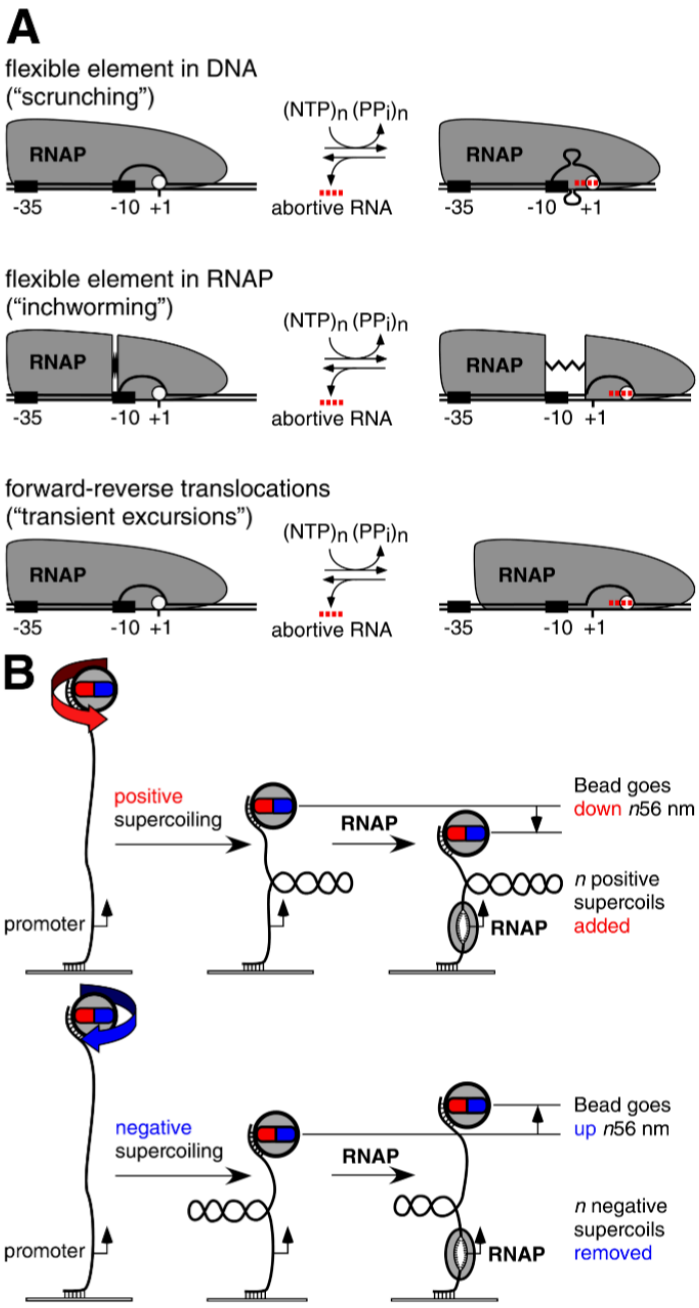


Fig. 1. Background and experimental approach

(A) Models for RNAP-active-center translocation in abortive initiation: (i) The "scrunching" model invokes a flexible element in DNA (4,7,11,12; see also 13-19). In each cycle of abortive initiation, RNAP unwinds and pulls into itself downstream DNA, accommodating the accumulated DNA as single-stranded bulges in the unwound region; upon release of the abortive RNA, RNAP extrudes the internalized DNA. (ii) The "inchworming" model invokes a flexible element in RNAP (9,10). In each cycle of abortive initiation, a module of RNAP containing the active center (white circle) detaches from the remainder of RNAP and translocates downstream; upon release of the abortive RNA, this module of RNAP reverse translocates. (iii) The "transient-excursions" model invokes abortive cycles that are transient--

so short in lifetime, and so infrequent in occurrence, that they cannot be detected in a time-averaged, population-averaged approach such as DNA footprinting (7). In each cycle of abortive initiation, RNAP translocates downstream as a unit; upon release of the abortive RNA, RNAP reverse translocates as a unit. **(B)** Experimental approach (Fig. S1;20-22). The end-to-end extension of a mechanically stretched, negatively supercoiled (top) or positively supercoiled (bottom), single DNA molecule containing a single promoter is monitored. Unwinding of *n* turns of DNA by RNAP results in the compensatory loss of *n* negative supercoils, or gain of *n* positive supercoils, and a readily detectable, nanometer-scale (*n*·56 nm), movement of the bead.

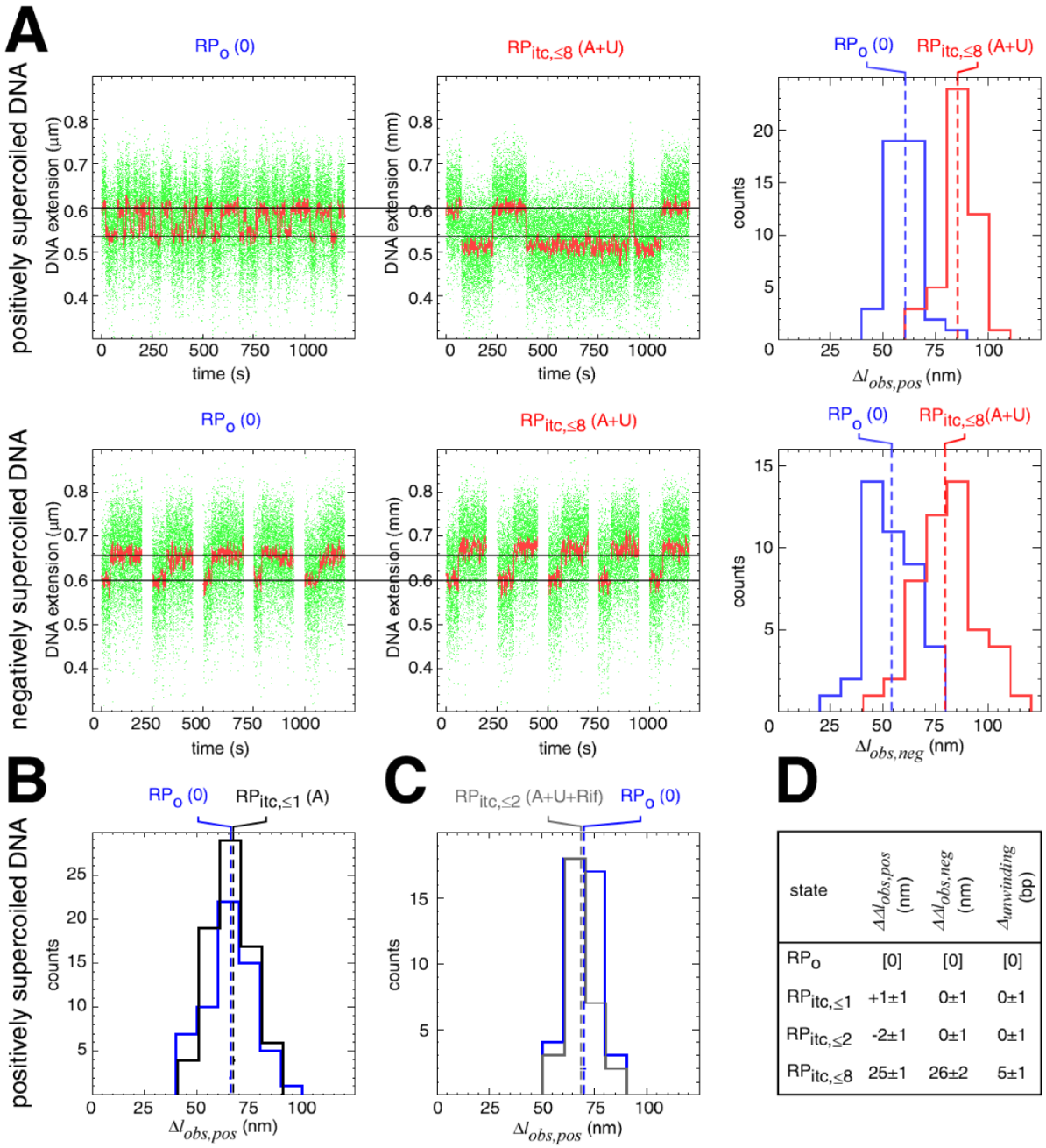
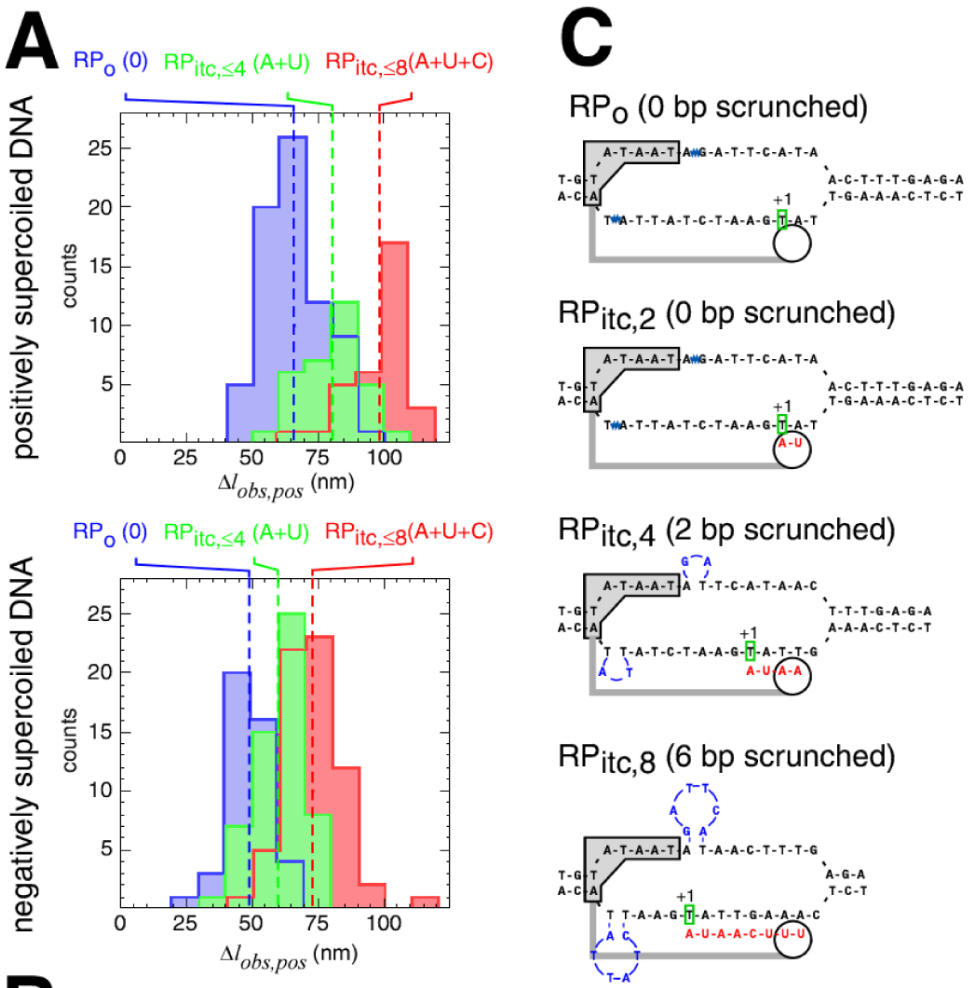


Fig. 2. Scrunching occurs in abortive initiation

(A) Single-molecule time traces and transition-amplitude histograms for RP_o (0 NTPs) and $RP_{itc,\leq 8}$ (ATP+UTP) at the N25 promoter. Data for positively supercoiled DNA are at top; data for negatively supercoiled DNA are at bottom. Green points, raw data (30 frames/s); red points, averaged data (1 s window); dashed lines in histograms, means; $\Delta l_{obs,pos}$, transition amplitude with positively supercoiled DNA; $\Delta l_{obs,neg}$, transition amplitude with negatively supercoiled DNA.

(B) Transition-amplitude histogram for $RP_{itc,\leq 1}$ (from control experiment providing only ATP).

(C) Transition-amplitude histogram for $RP_{itc,\leq 2}$ (from control experiment providing ATP, UTP, and rifampicin).
(D) Differences in $\Delta l_{obs,pos}$, $\Delta l_{obs,neg}$, and unwinding relative to values in RP_o (mean ± SE).



state	$\Delta\Delta l_{obs,pos}$ (nm)	$\Delta\Delta l_{obs,neg}$ (nm)	$\Delta_{unwinding}$ (bp)
RP_o	[0]	[0]	[0]
$RP_{itc,\leq 4}$	15±1	10±2	2±1
$RP_{itc,\leq 8}$	33±1	23±2	5±1

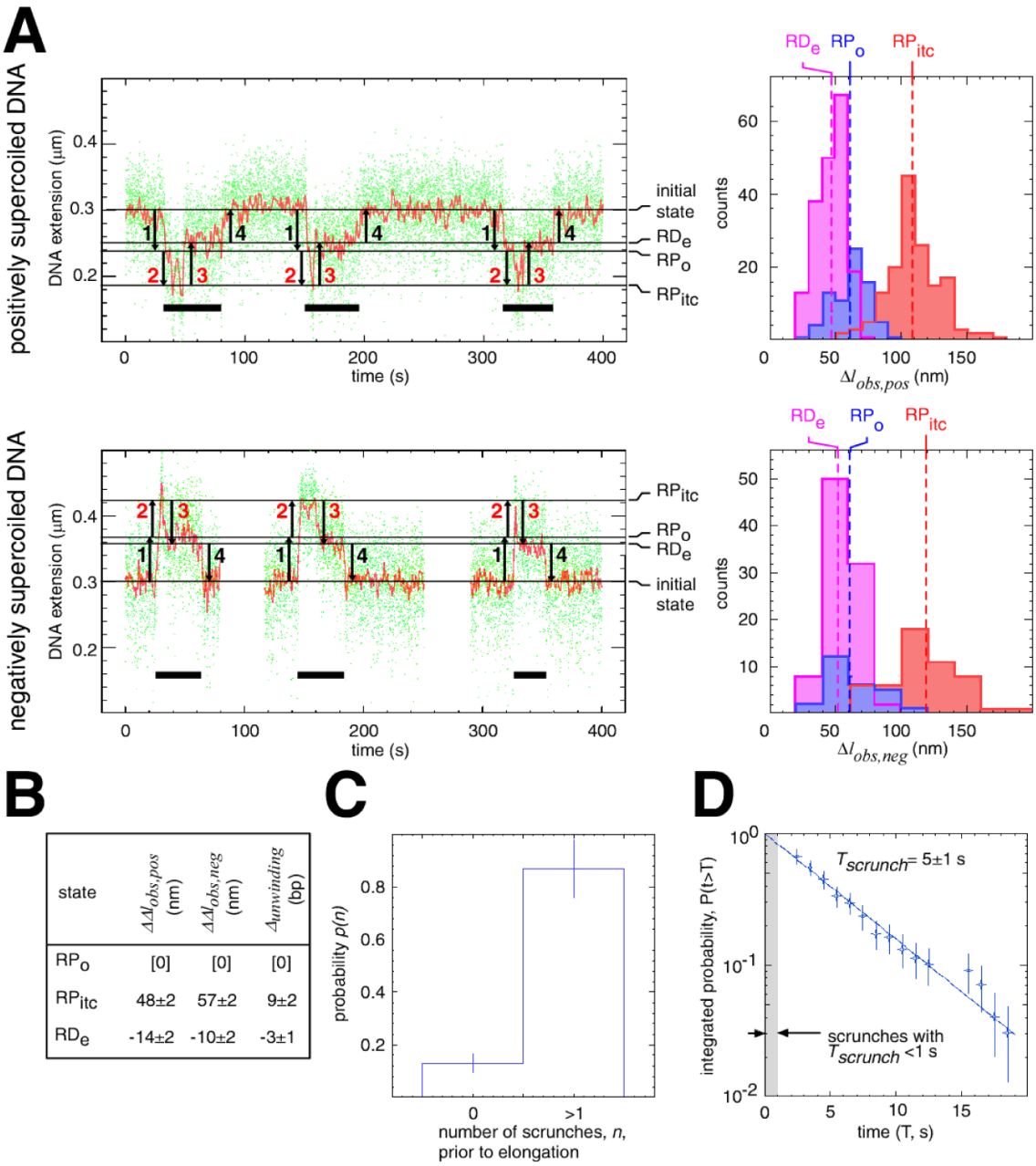
Fig. 3. The extent of scrunching correlates with the length of the RNA product

(A) Transition-amplitude histograms for RP_o (0 NTPs), $RP_{itc,\leq 4}$ (ATP+UTP), and $RP_{itc,\leq 8}$ (ATP+UTP+CTP) at the N25A5C promoter. Data for positively supercoiled DNA are at top; data for negatively supercoiled DNA are at bottom.

(B) Differences in $\Delta l_{obs,pos}$, $\Delta l_{obs,neg}$, and unwinding relative to values in RP_o (mean ± SE).

(C) Prediction of scrunching model: number of base pairs scrunched equals N-2, where N is the length of the RNA product. Sequence-specific RNAP-promoter interactions that define the upstream boundary of the unwound region are indicated by a gray box; RNAP structural elements that constrain the spacing between the upstream boundary of the unwound region and the RNAP active center are indicated by a gray bar; the RNAP active center is indicated by a

white circle; the RNA product is in red; position +1 of the template DNA strand is in green; scrunched DNA nucleotides are in blue (and are positioned as proposed in the companion paper; 24).



state	$\Delta\Delta I_{obs,pos}$ (nm)	$\Delta\Delta I_{obs,neg}$ (nm)	$\Delta_{unwinding}$ (bp)
RP_o	[0]	[0]	[0]
RP_{itc}	48±2	57±2	9±2
RD_e	-14±2	-10±2	-3±1

Fig. 4. Scrunching occurs in promoter escape in productive initiation

(A) Single-molecule time traces and transition-amplitude histograms for complete transcription cycles on N25-400-tR2 in the presence of all four NTPs. Data for positively supercoiled DNA are at top; data for negatively supercoiled DNA are at bottom. Transcription cycles are indicated by horizontal black bars; unwinding and rewinding transitions are indicated by numbered arrows (red numbered arrows for scrunching and reversal of scrunching); and states are indicated by horizontal lines and labelled at the right of the time traces.

(B) Differences in $\Delta I_{obs,pos}$, $\Delta I_{obs,neg}$, and unwinding relative to values in RP_o.

(C) Fraction of transcription cycles exhibiting at least one detectable scrunch (mean ± SEM; N = 100).

(D) Distribution of scrunch lifetimes (measured from midpoint of transition 2 to midpoint of transition 3; N = 100).