



HAL
open science

Quadruplex-DNA: structures, functions and detection

David Monchaud

► **To cite this version:**

David Monchaud. Quadruplex-DNA: structures, functions and detection. Médecine/Sciences, 2017, 33 (12), pp.1042-1045. 10.1051/medsci/20173312008 . hal-02268711

HAL Id: hal-02268711

<https://hal.science/hal-02268711>

Submitted on 21 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

enjeu majeur pour le traitement futur des maladies liées à un déséquilibre de la flore intestinale. **◇**

Impact of the evolution of mammalian intestinal microbiotas on human health

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Lozupone CA, Stombaugh JI, Gordon JI, et al. Diversity, stability and resilience of the human gut microbiota. *Nature* 2012 ; 489 : 220-30.
2. Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. *Genome Med* 2016 ; 8 : 51.
3. El Kaoutari A, Armougom F, Raoult D, et al. Le microbiote intestinal et la digestion des polysaccharides *Med Sci (Paris)* 2014 ; 30 : 259-65.
4. Bordenstein SR, Theis KR. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol* 2015 ; 13 : e1002226.
5. Moran NA, Sloan DB. The hologenome concept: helpful or hollow? *PLoS Biol* 2015 ; 13 : e1002311.
6. Groussin M, Mazel F, Sanders JG, et al. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nat Commun* 2017 ; 8 : 14319.
7. Brucker RM, Bordenstein SR. The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science* 2013 ; 341 : 667-9.
8. Ley RE, Hamady M, Lozupone C, et al. Evolution of mammals and their gut microbes. *Science* 2008 ; 320 : 1647-51.
9. Muegge BD, Kuczynski J, Knights D, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 2011 ; 332 : 970-4.
10. Papa E, Doktor M, Smillie CS, et al. J. Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. *PLoS One* 2012 ; 7 : e39242-12.
11. Normand S, Secher T, Chamaillard M, La dysbiose, une nouvelle entité en médecine ? *Med Sci (Paris)* 2013 ; 29 : 586-9.

NOUVELLE

Quadruplexes d'ADN : structures, fonctions et détection

David Monchaud

Institut de chimie moléculaire de l'université de Bourgogne (ICMUB), CNRS UMR6302, 9, avenue Alain Savary, 21078 Dijon, France.

david.monchaud@cnrs.fr

> L'acide désoxyribonucléique, ou ADN, est le support de l'information génétique ; il occupe, de fait, une place centrale dans notre compréhension des mécanismes sous-jacents aux processus vitaux comme le code génétique ou encore l'hérédité. Cependant, notre appréhension de l'ADN d'un point de vue structural et fonctionnel connaît actuellement un changement radical. Depuis la découverte de sa structure en double-hélice (ou duplexe) en 1953 [1], notre vision de l'ADN était celle d'une structure sophistiquée, sanctuarisée dans le coffre-fort de la cellule (le noyau), ainsi précieusement protégée des aléas de la vie cellulaire. Or, il apparaît aujourd'hui que l'ADN n'est pas seulement un code barre qui attend patiemment d'être scanné au gré de l'activité cellulaire, mais plutôt une structure hautement dynamique qui serait capable de gérer elle-même une partie de son activité. Notre patrimoine génétique se compose de près de 3 milliards de paires de bases azotées dont seulement 1 à

2 % constituerait de l'ADN codant, qui chez l'homme se compose d'environ 20 à 25 000 gènes fonctionnels. La partie non-codante de notre génome, improprement nommée « ADN poubelle », se caractérise par de nombreuses séquences dites répétées (qui représenteraient environ 50 % de notre génome) dont on commence à comprendre aujourd'hui l'étendue et la portée de leurs implications fonctionnelles. Parmi ces répétitions, les séquences riches en guanines sont au cœur d'une attention scientifique sans précédent : les guanines offrent en effet des possibilités d'appariements entre nucléobases bien au-delà de l'appariement dit de Watson-Crick (entre guanine et cytosine), permettant de former des triplets voire des quartets de bases (ces derniers étant également nommés G-quartet, ou tétrade de guanines) [2]. Connue depuis les années 1960, cette capacité unique des guanines à s'auto-assembler en tétrade a pris une dimension particulière à la suite de la découverte de séquences répétées riches en

guanines dans des régions clés de notre génome, notamment au niveau des extrémités terminales des chromosomes (les télomères), des régions promotrices de gènes et des origines de réplication [3, 4] (→).

La richesse en guanines et la nature répétée de ces séquences leur offrent la possibilité d'adopter des structures différentes du duplexe (Figure 1) : la formation de tétrades de guanines permet notamment la stabilisation d'édifices non plus à deux brins (comme dans le duplexe) mais à quatre brins appelés quadruplexes, qui se caractérisent par une très grande stabilité thermodynamique et une cinétique de formation très favorable (l'assemblage se faisant de façon intramoléculaire). Ces observations expliquent pourquoi les premières recherches sur les quadruplexes ont été conduites sur les régions télomériques ; une partie des télomères existe, en effet, libérée de la contrainte duplexe (c'est-à-dire

(→) Voir la Synthèse de J. Poulet-Benedetti et al., page 1063 de ce numéro

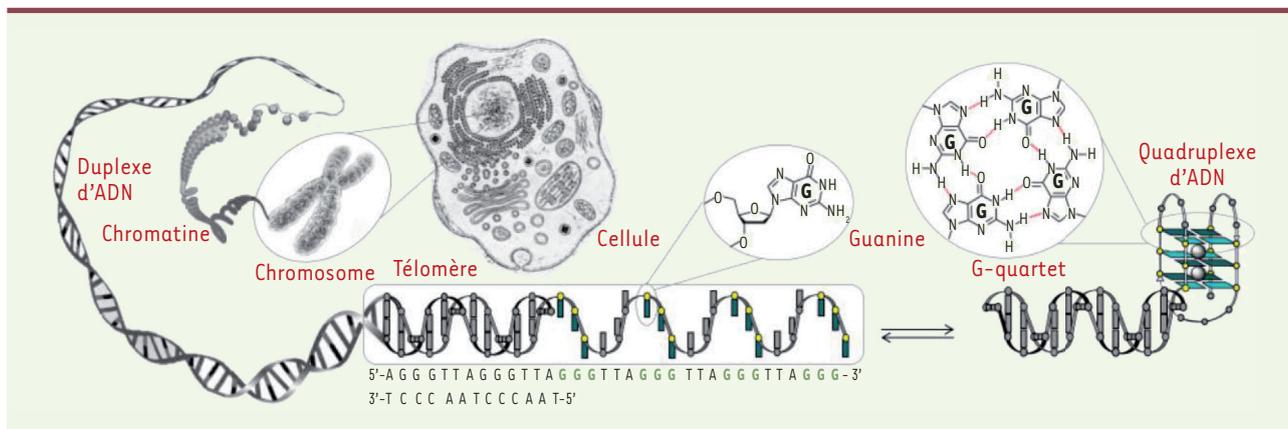


Figure 1. Représentation schématique d'une cellule et de son ADN. Emphase particulière sur l'ADN télomérique riche en guanines et sa partie terminale monocaténaire, capable de se structurer en quadruplexe via la formation de G-quartets (ou quartets de guanines).

non associée à un brin complémentaire), un fait unique dans la cellule. Ces séquences monocatéennes, constituées du motif TTAGGG répété plusieurs dizaines de fois, peuvent se replier sur elles-mêmes pour former une structure quadruplexe (Figure 1) qui a été et reste certainement la plus étudiée à ce jour, notamment dans la structuration des télomères (impactant la stabilité des chromosomes) et pour sa capacité à inhiber l'action de la télomérase (notamment dans le cadre de la lutte contre le cancer) [5].

Des études bioinformatiques puis de séquençage haut-débit ont montré que les séquences capables de former des structures quadruplexes étaient nombreuses (plus de 700 000 au sein du génome humain) [6]. Comme indiqué ci-dessus, elles sont réparties de façon non-aléatoire, étant particulièrement présentes au niveau des télomères, des régions promotrices de gènes et des origines de réplication. Elles sont en revanche absentes au niveau des gènes suppresseurs de tumeur. Ces séquences peuvent adopter une structure quadruplexe même en conditions natives, c'est-à-dire lorsqu'elles sont associées à leur brin complémentaire et engagées dans l'organisation nucléoprotéique d'ordre supérieur qu'est la chromatine [7] : les grandes manœuvres génomiques, que sont la réplication et la transcription,

offrent en effet une opportunité unique à ces séquences de se libérer de ces contraintes et de se structurer transitoirement en quadruplexe.

Les rôles qui sont à ce jour attribués aux structures quadruplexes sont divers et dépendent de leur localisation. Il est cependant difficile de comprendre toute la portée de leurs implications, car ces structures semblent jouer des rôles tour à tour activateurs ou inhibiteurs, en fonction de leur localisation génomique, de leur abondance, de leur stabilité, de leur capacité à échapper aux systèmes protéiques de régulation (comme les hélicases), etc. Par exemple, les quadruplexes peuvent aider à l'organisation structurale des télomères ou à l'initiation de la réplication de l'ADN (rôles activateurs) [4], tout comme sous-réguler l'expression de certains gènes ou générer de l'instabilité chromosomique (rôles inhibiteurs) [8].

Pour déchiffrer ces différents rôles, des outils permettant de détecter les quadruplexes au sein de notre génome ont été développés. Cette détection devait être dynamique, de façon à comprendre où, quand et comment les quadruplexes se structurent (et se déstructurent) au cours de la vie de la cellule. Du fait de leur formation transitoire et de leur faible abondance au regard de l'ADN génomique dans son entièreté, il a été particulière-

ment difficile de les localiser dans les cellules vivantes et en conditions fonctionnelles. Ce n'est que récemment qu'un effort de recherche au long cours a porté ses fruits, grâce à la conception d'outils moléculaires toujours plus précis et efficaces. Les grandes étapes qui ont conduit à la visualisation des quadruplexes *in cellulo* pourraient être résumées ainsi (Figure 2) : une première démonstration intervient en 2001, via l'utilisation d'un anticorps anti-quadruplexe (nommé Sty49)¹ qui permet la visualisation des quadruplexes télomériques présents dans les macronoyaux (contenant des millions de mini-chromosomes) du cilié *Stylonychia lemnae* [9]. En 2005, une équipe de recherche française utilise une molécule à haute affinité et sélectivité pour les quadruplexes (également appelée ligand de quadruplexe) radiomarquée au tritium (³H-360A) pour détecter par autoradiographie ces structures sur des cellules en culture et sur des étalements chromosomiques [10]. Les auteurs de cette étude notent une accumulation importante de la molécule ³H-360A au niveau des régions télomériques. En 2012, un autre ligand de quadruplexe

¹ Il s'agit d'un fragment d'anticorps simple chaîne (scFv) identifié dans une banque combinatoire d'anticorps qui a de l'affinité pour les répétitions télomériques.

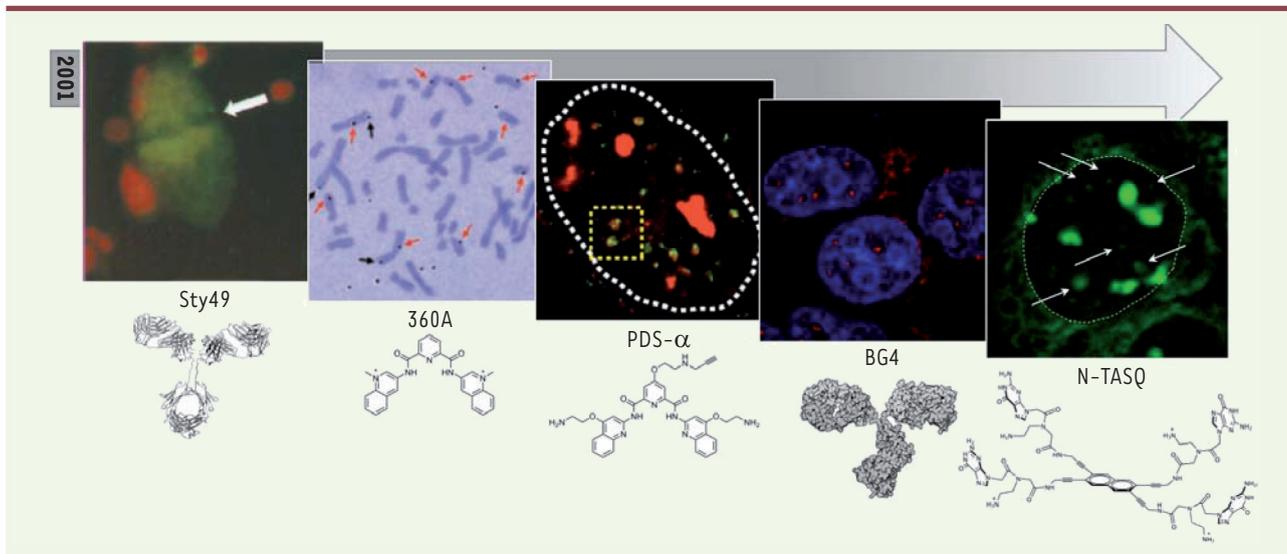


Figure 2. Évolution de l'imagerie des quadruplexes au cours du temps. Différentes approches d'imagerie cellulaire suivies depuis 2001 pour la visualisation des G-quadruplexes par l'utilisation d'anticorps (nommés Sty49 visualisé en vert [le contrôle interne négatif pointé par la flèche représente la bande de réplication] et BG4 en rouge [noyaux contre-colorés en bleu]) ou de sondes moléculaires, soit radioactives (utilisées sur un étalement de chromosome, telles que ^3H -360A [les quadruplexes télomériques sont pointés en noir et intragénomiques en rouge]), soit fluorescentes (utilisées en imagerie cellulaire, comme Alexa 594-PDS- α [les quadruplexes apparaissent en rouge, le co-marquage à l'aide d'une protéine sélective des quadruplexes, l'hélicase Pif1 conjuguée à la fluorescéine, en vert] et N-TASQ [les quadruplexes sont indiqués par les flèches]). Les images sont extraites des références [9-13].

affin et sélectif (la pyridostatine, ou PDS) est utilisé dans une approche complémentaire, non plus fondée sur de la radiodétection mais sur une analyse par fluorescence, qui présente des avantages notables en termes de sensibilité et de précision de détection. Par une approche multi-étape élégante, cette étude apporte une démonstration éclatante de l'existence des quadruplexes dans les cellules humaines [11]. Les images obtenues, alliées à une analyse pangénomique, conduisent à l'identification de quadruplexes télomériques, mais également intragénomiques, apportant ainsi des informations précieuses quant à leurs implications possibles dans la vie de la cellule. Ces études seront par la suite complétées par 1) le développement de l'anticorps anti-quadruplexe BG4, identifié dans une banque de phages (en 2013), qui a permis une visualisation par des techniques d'immunodétection classiques [12], puis 2) par celui de sondes fluorescentes dites intelligentes (ou *smart*

probes,² comme le NaphtoTASQ en 2015) pour la visualisation directe des quadruplexes dans les cellules humaines vivantes. [13]

La conception, la production et les études d'outils moléculaires visant à la détection – voire la quantification – des structures quadruplexes sont donc en plein essor, notamment *via* l'élaboration de nouveaux anticorps (comme 1H6 ou D1) et de ligands fluorescents (tels le composé *triangulenium* DAOTA-M2³ ou encore un composé dérivé du naphthalène diimide, le c-exNDI⁴) toujours plus performants. Ces outils permettront *in fine* d'obtenir des informations précieuses quant à la spatio-temporalité de la formation des quadruplexes au sein de

notre génome et ainsi, de mieux appréhender les différents processus cellulaires dans lesquels ces structures sont impliquées. La compréhension des rôles que ces structures non-usuelles de l'ADN jouent dans nos cellules nous permettra de percer encore un peu plus le mystère du fonctionnement de notre génome ; cependant, au regard de nos connaissances actuelles et de la diversité tant structurale que fonctionnelle des acides nucléiques, nul doute que l'ADN n'a pas encore livré tous ses secrets. \diamond

Quadruplex-DNA: structures, functions and detection

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Watson JD, Crick FH. Molecular structure of nucleic acids. *Nature* 1953; 171 : 737-8.
2. Sivakova S, Rowan SJ. Nucleobases as supramolecular motifs. *Chem Soc Rev* 2005; 34 : 9-21.
3. Hänsel-Hertsch R, Di Antonio M, Balasubramanian S. DNA G-quadruplexes in the human genome : detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* 2017; 18 : 279-84.

² Son interaction avec un G-quadruplexe provoque un changement conformationnel qui induit une augmentation de sa fluorescence.

³ DAOTA-M2 est un composé dont la fluorescence augmente en présence d'acides nucléiques et dont la durée de vie de la fluorescence dépend de la topologie des quadruplexes.

⁴ Ce composé s'agrège en solution, ce qui le rend non-fluorescent ; son interaction avec les G-quadruplexes provoque sa désagrégation et restaure sa fluorescence.



RÉFÉRENCES

4. Poulet-Benedetti J, Prioleau MN, Valton AL. G-quadruplexes : acteurs majeurs de la duplication du génome humain. *Med Sci (Paris)* 2017 ; 33 : 1063-70.
5. De Cian A, Lacroix L, Douarre C, et al. Targeting telomeres and telomerase. *Biochimie* 2008 ; 90 : 131-55.
6. Chambers VS, Marsico G, Boutell et al. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol* 2015 ; 33 : 877-81.
7. Hänsel-Hertsch R, Beraldi D, Lensing SV, et al. G-quadruplex structures mark human regulatory chromatin. *Nat Genet* 2016 ; 48 : 1267-72.
8. Tarsounas M, Tijsterman M. Genomes and G-quadruplexes: for better or for worse. *J Mol Biol* 2013 ; 425 : 4782-9.
9. Schaffitzel C, Berger I, Postberg J, et al. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylochyia lemnae* macronuclei. *Proc Natl Acad Sci USA* 2001 ; 98 : 8572-7.
10. Granotier G, Pennarun G, Riou L, et al. Preferential binding of a G-quadruplex ligand to human chromosomes ends. *Nucleic Acids Res* 2005 ; 33 : 4182-90.
11. Rodriguez R, Miller KM, Forment JV, et al. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat Chem Biol* 2012 ; 8 : 301-10.
12. Biffi G, Tannahill D, McCafferty J et al. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* 2013 ; 5 : 182-6.
13. Laguerre A, Wong JMY, Monchaud D. Direct visualization of both DNA and RNA quadruplexes in human cells via an uncommon spectroscopic mechanism. *Sci Rep* 2016 ; 6 : 32141.

NOUVELLE

Séquençage d'exome et de transcriptome : des approches complémentaires pour identifier des gènes impliqués dans la survenue de maladies

Agnès Rötig

UMR1163, université Paris Descartes, Sorbonne Paris Cité, Institut Imagine, 24, boulevard du Montparnasse, 75015 Paris, France.
agnes.rotig@inserm.fr

► Les techniques de séquençage haut débit ont révolutionné l'identification des gènes impliqués dans la survenue de maladies génétiques [1] (→).

L'approche la plus utilisée est le séquençage

d'exome, c'est-à-dire des exons, les séquences codantes du génome [2]. L'exome ne représente que 3 % du génome humain, mais son étude est particulièrement appropriée dans les maladies mendéliennes [3] (→).

La très grande majorité des mutations touche, en effet, les régions

codantes, que ce soit des mutations faux-sens, non-sens ou des modifications d'épissage conduisant à un transcrit anormal ou absent. La puissance de cette technique est telle que des centaines de gènes ont pu être identifiés par séquençage de l'exome d'un seul patient, ou mieux du trio père-mère-

(→) Voir la Nouvelle de D. Stoppa-Lyonnet et C. Houdayer, *m/s* n° 2, février 2012, page 123

(→) Voir la Nouvelle de A. Criqui et S. Baulande, *m/s* n° 5, mai 2010, page 452

enfant atteint d'une pathologie. Dans la moitié des cas environ, le séquençage d'exome ne permet cependant pas d'identifier la mutation responsable de la maladie pour différentes raisons : (1) certains exons sont mal couverts car ils sont riches en GC (guanine/cytosine) et difficiles à amplifier ; (2) les mutations modifient l'épissage, mais sont localisées loin dans l'intron ; ou (3) parce qu'il est souvent impossible d'identifier la mutation parmi tous les variants de signification inconnue (VUS : *variants of unknown significance*), mutation qui pourrait jouer sur la nature ou l'abondance de l'ARN messager (ARNm).

Récemment, le groupe d'Holger Prokisch, à Munich, a montré l'intérêt du séquençage du transcriptome (séquençage de l'ARN, ou *RNA-seq*) dans l'identification de gènes impliqués dans des maladies mendéliennes [4]. Le *RNA-seq* permet, en effet, de compléter l'information génétique brute donnée par le séquençage d'exome, par une donnée fonctionnelle qui révèle les conséquences de

VUS, ou met en évidence des transcrits quantitativement ou qualitativement anormaux du fait d'une mutation dans des régions non-codantes. Cette étude a été menée sur une série de patients suspectés, sur des critères cliniques, d'avoir une maladie mitochondriale et pour lesquels le séquençage d'exome avait été infructueux.

Les maladies mitochondriales représentent collectivement les maladies métaboliques les plus fréquentes, mais leur extrême hétérogénéité clinique et génétique rend leur diagnostic génétique de plus en plus compliqué. Ces maladies résultent, en effet, de mutations touchant soit l'ADN mitochondrial (ou ADNmt), soit des gènes nucléaires responsables d'un très grand nombre de fonctions mitochondriales [5]. On connaît actuellement plus de 250 gènes nucléaires impliqués dans des maladies mitochondriales et des centaines de mutations de l'ADNmt. Les corrélations entre génotype et phénotype se révèlent parfois difficiles puisqu'un