

► Une vaste plasticité des acides nucléiques s'offre à la biologie synthétique, pour les surproduire et diversifier leurs structures chimiques, et aussi pour établir *in vivo* et *in vitro* des processus informationnels qui n'ont pas été explorés au cours de l'évolution des espèces. ◀

Horizon scientifique et enjeu industriel de la biosynthèse des acides nucléiques

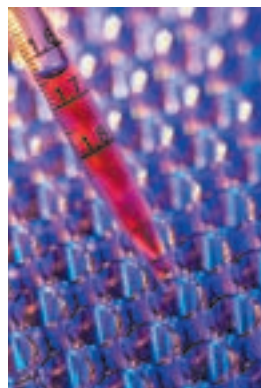
Philippe Marlière

La découverte du statut de mémoire moléculaire des acides nucléiques, ADN et ARN, en lesquels s'enregistre l'ensemble des processus biologiques et par lesquels ils se propagent, a déclenché au XX^e siècle la révolution paradigmatique connue sous le nom de « biologie moléculaire ».

Voici comment en peu de mots pourrait se résumer l'apport essentiel du siècle passé à la techno-science dans le domaine de la biologie synthétique : (1) Toute séquence de nucléobases appariée à sa séquence complémentaire présente la même géométrie et la capacité d'être reproduite par copie dédoublée, sous l'action d'enzymes de réplication par des organismes vivants ou suivant des procédés d'amplification *in vitro*. (2) Toute séquence nucléique peut être fabriquée suivant un procédé de chimie automatisée sur support solide pour être ensuite concaténée sous l'action d'enzymes de ligation ou de recombinaison à d'autres séquences nucléiques et former des réplicons plasmidiques de synthèse ou pour altérer, accroître ou réduire des réplicons génomiques naturels. (3) Toute séquence polypeptidique est susceptible d'être spécifiée dans un gène synthétique suivant la grille du code génétique, que cette séquence corresponde à une protéine native ou à un prototype virtuel.

Vers la « génématique »

La mise en œuvre systématique et à grande échelle de cette capacité technique se traduira dans le futur par la synthèse intégrale de génomes bactériens, c'est-à-dire l'assemblage d'instructions génétiques nécessaires et suffisantes à la reproduction d'une cellule vivante à partir de nutriments chimiquement définis. La construction des OGM bactériens, végétaux et animaux a consisté jusqu'à présent à effectuer des altérations génomiques limitées à quelques délétions ou insertions de gènes.



Isthmus SARL,
31, rue Saint-Amand,
75015 Paris, France.
phmarliere@isthmus-sarl.com

La réécriture globale de génomes cellulaires va faire passer la biologie, dans un premier temps la microbiologie, sous le régime de la génétique automatique (« génématique »). Toutes les étapes, allant de la conception immatérielle des biocatalyseurs à l'assemblage chimique des instructions génétiques spécifiant ceux-ci et leur composition dans des génomes, seront gouvernées par des algorithmes et l'incarnation chimique proprement dite sera accomplie dans des réseaux fluidiques, sans intervention humaine. La métaphore technique qui paraît la plus appropriée pour rendre compte de la réécriture des génomes est celle de l'imprimerie, passant des copistes à la presse, puis aux rotatives, puis à la photocomposition assistée par ordinateur.

Ces avancées, qui furent longtemps marquées d'une suspicion d'ésotérisme, sont maintenant largement admises comme souhaitables et inexorables par la communauté scientifique. La commémoration des dix ans de Genopole®, qui s'est institué à l'époque charnière où la biologie a basculé de la démarche analytique à la démarche synthétique, est le moment propice pour balayer d'un bref tour d'horizon les projets relatifs à la bioproduction et à la diversification *in vivo* des acides nucléiques.

L'ADN comme expédient métabolique

Le rôle central de l'ADN comme support chimique des génomes de toutes les cellules et de nombreux virus,



son rôle de véhicule informationnel inter-espèces dans les transferts horizontaux, la sophistication des processus de réplication, réparation et recombinaison semblent ne pas avoir une contrepartie métabolique aussi bien conçue dans la biosynthèse des précurseurs métaboliques de l'ADN, les désoxynucléosides triphosphates (dNTP). Ceux-ci sont formés à partir des précurseurs de l'ARN, les rNTP, sous l'action d'une famille d'enzymes au mécanisme radicalaire bien controuvé, la ribonucléotide réductase. Cette organisation présente tous les stigmates d'un expédient chimique qui serait survenu au cours de l'évolution du métabolisme. L'émergence de l'ARN comme type moléculaire primordial, ayant à la fois le statut de génotype et de phénotype et ayant préexisté à la fois aux protéines et à l'ADN, pourrait expliquer cet état de fait.

La réduction des rNTP en dNTP aurait en effet permis de diminuer la propension à l'hydrolyse spontanée des polynucléotides de type ARN et, ce faisant, accru la capacité de stockage de dizaines de milliers à des dizaines de millions de nucléobases par chaîne nucléique. La biosynthèse des rNTP est intriquée aux voies centrales du métabolisme, le cycle des phosphopentoses et la branche sérine/glycine. Les bases pyrimidine et purine sont assemblées et diversifiées sur le motif phosphoribose, avant que celui-ci ne soit phosphorylé puis réduit pour donner les dNTP.

C'est seulement une faible fraction du flux métabolique vers l'ARN qui est déviée vers l'ADN par la ribonucléotide réductase, suivant un mécanisme radicalaire délicat, et ce sous un contrôle allostérique plafonnant strictement la concentration intracellulaire des substrats dNTP. Ces deux caractéristiques, mécanisme radicalaire et régulation allostérique, restreignent la production massive d'ADN par voie microbiologique.

C'est donc suivant un changement de perspective que notre équipe scientifique, coordonnée par Volker Döring, a entrepris de rectifier la biosynthèse des désoxynucléotides en engendrant directement le sucre désoxyribose, qui ne survient normalement pas dans les cellules. Les dNTP peuvent en effet être dérivés par introduction de nucléobases et de phosphate sur le motif désoxyribose par l'action d'enzymes connues.

L'activité virtuelle d'une enzyme à thiamine qu'on pouvait *a priori* désigner comme une « 2-déshydro-3-désoxy-D-gluconate décarboxylase », semblable à la pyruvate décarboxylase qui mène à l'éthanol chez la levure, nous parut idéalement à même de connecter le catabolisme des sucres à l'anabolisme de l'ADN. Le substrat de notre enzyme virtuel, le 2-déshydro-3-désoxy-D-glucose (KDG) est connu pour survenir dans le métabolisme du gluconate. Mettant en œuvre les cribles sélectifs et les tests biochimiques appropriés, l'équipe d'Evologic SA parvint à cloner deux enzymes, la première capable de former KDG à partir de gluconate, et la seconde de décarboxyler KDG en désoxyribose. Cette voie alternative d'accès à l'ADN paraît inédite dans le métabolisme des organismes connus, en dépit de sa compacité et de sa simplicité mécanistique.

Renouveler l'industrie de l'ADN et des antiviraux

Le procédé breveté est aujourd'hui développé par Madeleine Bouzon-Bloch à l'Institut de génomique du CEA. Il permet de convertir le gluconate, issu de l'agroconversion de l'amidon, et instaure ainsi à Genopole

une filière susceptible de renouveler l'industrie de l'ADN et des nucléosides pharmaceutiques. Les antiviraux AZT (azidothymidine) et DDI (didésoxyinosine), entrent dans la formulation des multithérapies, seul rempart sanitaire aujourd'hui disponible pour contenir la pandémie du sida. Ces deux molécules s'obtiennent par conversion chimique de la thymidine et de la désoxyadénosine, naguère obtenues par le fractionnement malaisé des produits d'hydrolyse de l'ADN naturel. C'est leur production massive et directe qu'autorise désormais la

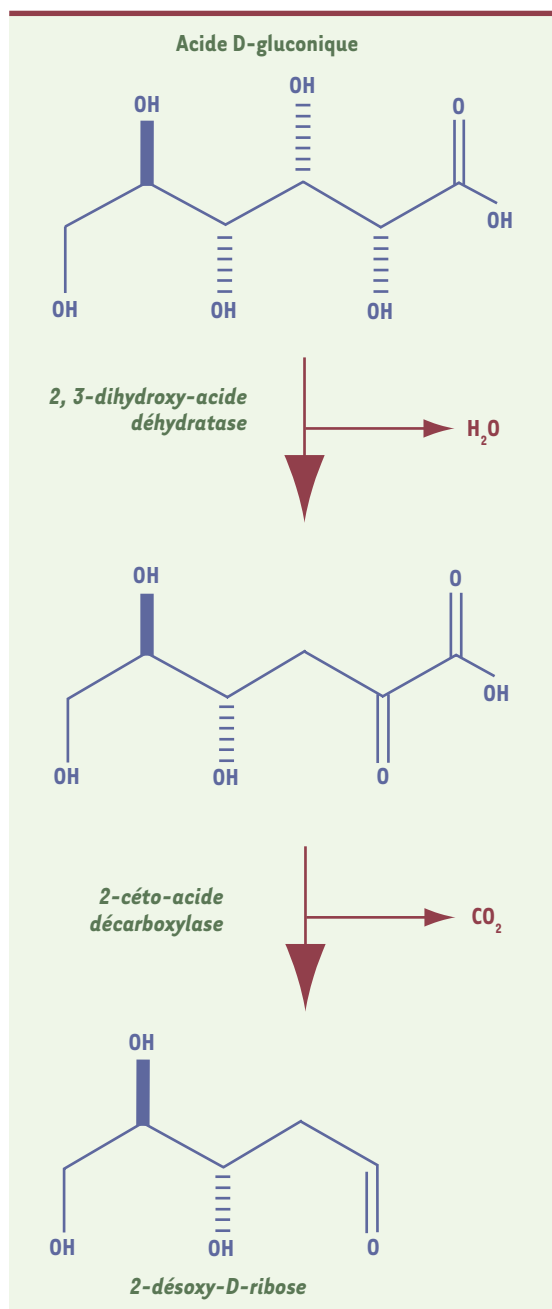


Figure 1. Rectification de la biosynthèse du désoxyribose.



bioconversion de gluconate en désoxyribose puis en désoxynucléosides, suivie d'étapes de conversion chimique en molécules thérapeutiques.

L'installation prochaine dans des cellules microbiennes de la voie rectifiée de bioproduction du désoxyribose et de la conversion de celui-ci en dNTP augure favorablement d'une filière simplifiant la préparation massive de chaînes d'ADN de séquence déterminée pour produire, si le besoin s'en fait sentir, des vaccins vétérinaires et médicaux ou des composants pour les montages nanotechnologiques. La disponibilité massive de dNTP ouvre potentiellement la voie à la production et au transfert génétique d'ADN aléatoire, dans lequel pourrait être recrutées de nouvelles fonctions, par exemple en co-exprimant l'enzyme du système immunitaire des mammifères, la terminal désoxyribonucléotide transférase. Les assemblages génétiques et métaboliques correspondants sont en cours de construction à l'Institut de génomique du CEA, en collaboration avec François Rougeon.

Vers la diversification des bases de l'ADN

Les nucléobases ATGC, hétérocycles facilement accessibles par synthèse organique et peu onéreux, sont, à ce premier stade de l'installation d'une voie alternative de biosynthèse de l'ADN, appelées à être fournies dans le milieu nutritif. Cependant, un schéma parallèle de bioproduction massive, chimiquement simplifiée et dérégulée, est exploré dans le but de dériver, à partir d'un métabolite simple mais non physiologique, à la fois les nucléobases puriques G et A et les nucléobases pyrimidiques C et U (uracile) suivant un nombre restreint d'étapes.

Il n'est pas extravagant de considérer l'évolution dirigée de la structure chimique de génomes descendant de celui d'*Escherichia coli* ou d'autres bactéries modèles comme *Acinetobacter baylyi*, de sorte qu'on puisse forcer le remplacement de l'une des bases soit par une base naturelle des bactériophages comme l'uracile ou la 2,6-diaminopurine, soit par une base synthétique comme les 5-halogéno-uraciles, les 8-azapurines, les 3-déazapurines ou les 7-déazapurines. La substitution *in vivo* des nucléobases de l'ADN génomique n'a pas encore fait l'objet d'une étude systématique, laquelle pourrait déboucher sur de nouveaux régimes de mutation ou de recombinaison, ainsi que sur la fonctionnalisation de l'ADN facilitant sa purification ou sa manipulation.

Les avancées dans la rectification du métabolisme de l'ADN et dans les techniques de culture devraient aboutir à des percées dans ces domaines. À terme, les quatre nucléobases ACGT pourraient subir un remplacement à l'échelle globale d'un génome bactérien par un analogue chimique, comme par exemple la 2,6-diamino-purine (A → D), l'uracile (T → U), la 5-méthyl-cytosine (C → mC) et l'hypoxanthine (G → Hx).

Vers la diversification du squelette de l'ADN

L'approche la plus radicale et la plus directe pour pousser les acides nucléiques *in vivo* dans des directions innovantes consistera à répliquer des plasmides, de telle sorte que l'enzyme en charge de cette répllication, la polymérase, ainsi que ses substrats, n'interfèrent pas dans la répllication du chromosome de l'hôte bactérien. Propager de tels réplicons de deuxième génération, en élaborant les biocatalyseurs

du transport, du métabolisme et de la polymérisation de leurs précurseurs nucléotidiques, se profile comme un domaine majeur de la biologie synthétique, en raison de ses enjeux fondamentaux et appliqués. La Commission européenne finance, dans le cadre du PCRD6, le projet collaboratif Orthosome, entre l'Institut Riga à Louvain, le MRC à Cambridge et l'Institut de génomique du CEA à Évry, qui vise à propager un troisième type d'acide nucléique (3NA) chimiquement distinct de l'ARN et de l'ADN chez un hôte bactérien.

C'est de la substitution chimique non des nucléobases mais du squelette phospho-désoxyribose qu'il s'agit dans ce cas. À son aboutissement, ce projet livrera un réplicon 3NA spécifiant au moins un gène indispensable au métabolisme de l'hôte, de manière à sélectionner sa persistance intracellulaire et sa prolifération contrôlée. Une fois cette tête de pont établie, c'est l'ensemble des gènes nécessaires assurant la répllication du 3NA (3NA répllicase), sa transcription en ARN (3NA transcriptase) et, éventuellement sa recombinaison, qui seraient appelés à être encryptés sous forme 3NA.

La structure chimique choisie pour instaurer *in vivo* un 3NA est le HNA (*hexitol nucleic acid*), dont le motif se distingue seulement du désoxyribose par l'interposition d'un groupe méthylène (CH₂) entre l'oxygène et le carbone lié à la nucléobase, menant à un cycle éther à six atomes doué d'une grande stabilité chimique. Il est maintenant expérimentalement établi que plusieurs squelettes portant des nucléobases canoniques ATGC sont capables de s'apparier soit exclusivement en homoduplex 3NA : 3NA, soit en homoduplex, ainsi qu'en hétéroduplex 3NA : DNA et 3NA : RNA. Le HNA en est un exemple, aux côtés du TNA, du GNA, du homoDNA. Sa structure autorise la formation d'appariements HNA : HNA, HNA : DNA et HNA : RNA de stabilité comparable sinon légèrement supérieure à celle des acides nucléiques naturels. Il n'est pas exagéré de supposer qu'une multitude de squelettes possède la capacité de servir de mémoire génétique une fois ceux-ci pourvus d'enzymes de répllication ajustées à leur structure, en dépit du fait qu'un infime échantillon de telles structures a été testé par l'évolution des espèces sur la terre.

Des expériences ont montré que certaines ADN polymérases sauvages, comme Poll, Taq, et Terminator, avaient une activité marginale permettant d'utiliser des analogues 3NA des substrats dNTP, amorce et matrice, et donc une capacité potentielle de 3NA répllicase. La transcription de 3NA en ARN par des ARN polymérases est de même susceptible d'être conduite par la T7 ARN polymérase. Il faut s'attendre, dans de brefs délais, à l'amplification PCR ou RCA d'autres squelettes que ceux de l'ADN ou de l'ARN.

D'autres expériences menées à l'aide d'extraits acellulaires ont montré que la traduction génétique des polynucléotides en polypeptides (PRN) pouvait marginalement s'effectuer en soumettant à des ribosomes sauvages des polynucléotides différents de l'ARN comme l'ADN, mais aussi des polymères artificiels comme le HNA. En redirigeant l'évolution des gènes de 3NA polymérase et des protéines et ARN ribosomiques vers des versions spécialisées capables d'engendrer et de traduire des 3NA messagers, il est maintenant visé de propager et d'exprimer des 3NA *in vivo*, dans un premier temps chez les bactéries.

Déborder le « dogme central » de la biologie moléculaire

C'est à un débordement du « dogme central de la biologie moléculaire » (ADN → ARN → PRN) en bonne et due forme qu'aboutira cette lignée technique, en ajoutant dans les cellules une enclave (3NA → ARN → PRN) ou même (3NA → PRN). D'autant plus que des acides aminés supplémentaires pourront être spécifiés dans les protéines codées dans des gènes écrits et propagés sous forme 3NA.

C'est aussi un contournement radical du risque de pollution génétique occasionné par les OGM que fait augurer la translittération des gènes dans des polynucléotides 3NA. Supposons qu'on souhaite introduire un transgène donné dans une espèce. La propagation et l'expression *in vivo* de ce gène dans sa forme 3NA exigera la présence d'enzymes spécialisées de réplication et de transcription (voire de traduction), mais surtout elle exigera l'apport nutritionnel des composants chimiques du 3NA. Si ceux-ci sont conçus pour ne pouvoir être dérivés des métabolites habituels de la cellule (comme c'est le cas pour HNA), l'information transgénique ne persistera que tant que les précurseurs

de 3NA seront procurés comme nutriments, et ne se disséminera pas dans les habitats naturels, où ces précurseurs sont absents. Ce confinement informationnel et nutritionnel sera renforcé en spécifiant les enzymes de réplication et de transcription du 3NA dans le plasmide 3NA lui-même.

Pour toutes ces raisons, il faut anticiper un effort systématique de diversification de la structure chimique des polynucléotides et de leurs précurseurs, allant dans le sens de la maximisation de leurs capacités de codage et de répliquabilité et simultanément du découplage de leur métabolisme et de leur prolifération vis-à-vis de ceux des acides nucléiques canoniques ADN et ARN.

Les moyens de cette exploration fonctionnelle et de cette diversification structurale dépendront de manière critique de la capacité à importer dans les cellules les molécules hétérodoxes que sont les polynucléotides 3NA et leurs précurseurs métaboliques, dans un premier temps dans des bactéries transgéniques, et à plus long terme dans des cellules eucaryotes.

Des enclaves répliquatives 3NA

La faisabilité d'enclaves répliquatives 3NA se trouve corroborée par une série de résultats récents obtenus en collaboration dans le laboratoire de Piet Herdewijn et démontrant que d'autres précurseurs activés que les désoxynucléoside triphosphates (dNTP) servent de substrats à la polymérisation de l'ADN. La transcriptase inverse du VIH est ainsi capable

de condenser en oligomères d'ADN des adduits phosphoramidate entre l'acide L-aspartique et les 5'-phospho-désoxynucléotides, en maintenant la discrimination entre les nucléobases. Cette activité catalysant l'attaque nucléophile du lien P-N au lieu du lien P-O ne représente que 1 % de l'activité de l'enzyme sauvage avec des substrats dNTP. Il est cependant permis d'envisager un basculement par lequel des mutations cumulatives produiront un nouveau type de polymérase spécialisée pour un autre groupe partant tel que l'aspartate, n'interférant plus dans le métabolisme de l'ADN ni de

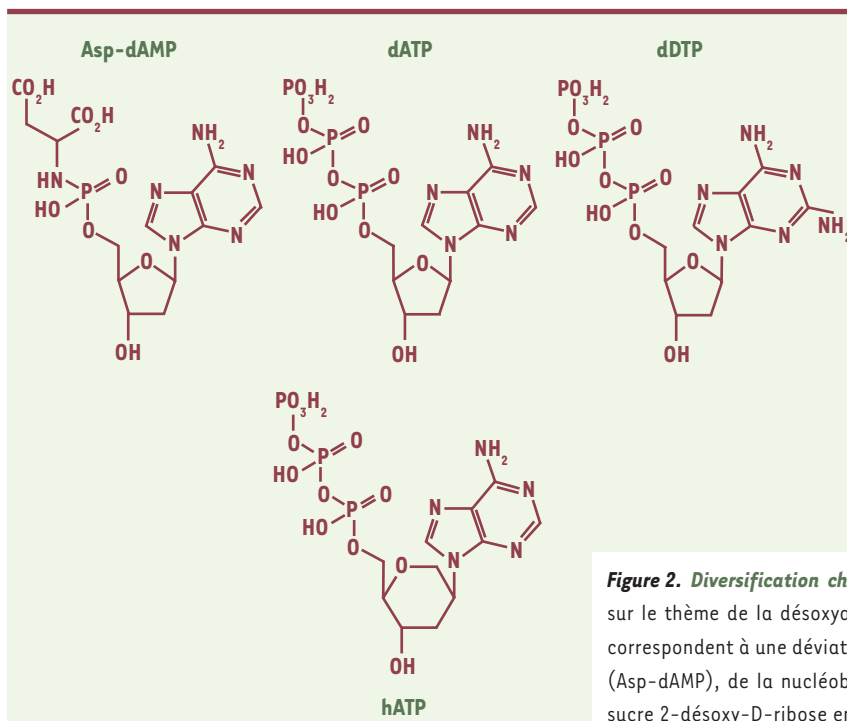


Figure 2. Diversification chimique des précurseurs de l'ADN. Trois variations sur le thème de la désoxyadénosine triphosphate (dATP) sont montrées, qui correspondent à une déviation du groupe partant pyrophosphate en aspartate (Asp-dAMP), de la nucléobase adénine en 2,6-diamino-purine (dDTP) et du sucre 2-désoxy-D-ribose en hexitol (hATP).



l'ARN, et tolérant une large gamme de variations structurales du sucre et de la base. Le même cheminement expérimental se conçoit pour activer les nucléotides avec d'autres métabolites que l'aspartate, qui puissent comme lui se prêter au recyclage métabolique, mais qui présentent une moindre ressemblance stérique et électronique avec le groupe pyrophosphate.

Renouveler le paradigme d'organisme génétiquement modifié

En bref, les audaces que l'expérimentateur s'autorise en tentant de répliquer et de métaboliser des formes chimiquement déviantes d'acides nucléiques semblent ne pas transgresser un interdit enzymatique ni physiologique. Une vaste plasticité fonctionnelle des acides nucléiques se présente aux *designers* des biotechnologies, pour surproduire et diversifier chimiquement ceux-ci, pour accomplir par leur propagation *in vivo* des bifurcations évolutives radicales, pour établir les têtes-de- pont de processus informationnels qui n'ont pas été explorés au cours de l'évolution spontanée des espèces. Au long cours, c'est le paradigme même d'organisme génétiquement modifié qui se trouvera renouvelé et l'interface sauvage/domestique préservée de la pollution génétique par l'introduction de catégories moléculaires répliquatives déviant de l'ADN et de l'ARN canoniques. Cette démarche a reçu l'appui du consortium Synbiosafe, subventionné par la Commission européenne dans le cadre du PCRD 6 pour baliser les risques de la biologie synthétique.

Grâces soient rendues à Pierre Tambourin et à Jean Weissenbach, qui animent l'un Genopole, l'autre l'Institut de génomique du CEA, d'avoir fondé un écosystème intellectuel propice à l'éclosion d'une biodiversité artificielle des génomes en France. ♦

TIRÉS À PART

P. Marlière

SUMMARY

Scientific horizon and industrial stake in the biosynthesis of nucleic acids

Nucleic acids offer an ample structural and functional plasticity enabling their industrial production and chemical diversification for therapeutic purposes as well as the establishment *in vivo* and *in vitro* of informational processes that did not spontaneously occur during evolution. ♦

BIBLIOGRAPHIE

1. Gibson DG, Benders GA, Andrews-Pfannkoch C, *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 2008 ; 319 : 1215-20.
2. El-Hajj HH, Wang L, Weiss B. Multiple mutant of *Escherichia coli* synthesizing virtually thymineless DNA during limited growth. *J Bacteriol* 1992 ; 174 : 4450-6.
3. Hewitt R, Sait JC, Billen D. Utilization of 5-bromouracil by thymineless bacteria. *J Bacteriol* 1967 ; 93 : 86-9.
4. Terrazas M, Marlière P, Herdewijn P. Enzymatically catalyzed DNA synthesis using L-Asp-dGMP, L-Asp-dCMP, and L-Asp-dTMP. *Chem Biodivers* 2008 ; 5 : 31-9.
5. Adelfinskaya O, Terrazas M, Froeyen M, *et al.* Polymerase-catalyzed synthesis of DNA from phosphoramidate conjugates of deoxynucleotides and amino acids. *Nucleic Acids Res* 2007 ; 35 : 5060-72.
6. Pochet S, Kaminski PA, Van Aerschot A, *et al.* Replication of hexitol oligonucleotides as a prelude to the propagation of a third type of nucleic acid *in vivo*. *CR Biol* 2003 ; 326 : 1175-84.
7. Lavrik IN, Avdeeva ON, Dontsova OA, *et al.* Translational properties of mHNA, a messenger RNA containing anhydrohexitol nucleotides. *Biochemistry* 2001 ; 40 : 11777-84.
8. Chaput JC, Szostak JW. TNA synthesis by DNA polymerases. *J Am Chem Soc* 2003 ; 125 : 9274-5.
9. Chaput JC, Ichida JK, Szostak JW. DNA polymerase-mediated DNA synthesis on a TNA template. *J Am Chem Soc* 2003 ; 125 : 856-7.
10. D'Abbadie M, Hofreiter M, Vaisman A, *et al.* Molecular breeding of polymerases for amplification of ancient DNA. *Nat Biotechnol* 2007 ; 25 : 939-43.
11. Doring V, Thibaut D, Kreimeyer A, Marlière P. Production of 2'-deoxynucleosides and 2'-deoxynucleoside precursors from 2-dehydro-3-deoxy-D-gluconate. WO2004113358.