FEBS Letters 586 (2012) 2049–2056

journal homepage: www.FEBSLetters.org

Review

Redesigning the leaving group in nucleic acid polymerization

P. Herdewijn [a,*], P. Marlière [b]

[a] Laboratory of Medicinal Chemistry, Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium
[b] Isthmus SARL, 13 rue Saint-Amand, F-75015 Paris, France

ARTICLE INFO

Article history:
Received 19 December 2011
Revised 22 February 2012
Accepted 22 February 2012
Available online 2 March 2012

Edited by Thomas Reiss and Wilhelm Just

Keywords:
Leaving group
Nucleotide
Pyrophosphate
Xenobiotics
Nucleic acid
Polymerase

ABSTRACT

Artificial nucleic acids have the potential to propagate genetic information in vivo purposefully insulated from the canonical replication and transcription processes of cells. Natural nucleic acids are synthesized using nucleoside triphosphates as building blocks and polymerases as catalysts, pyrophosphate functioning as the universal leaving group for DNA and RNA biosynthesis. In order to avoid entanglement between the propagation of artificial nucleic acids in vivo and the cellular information processes, we promote the biosynthesis of natural and xenobiotic nucleic acids (XNA) dependent on the involvement of leaving groups distinct from pyrophosphate. The feasibility of such radically novel biochemical systems relies on the systematic exploration of the chemical diversity of nucleic acid leaving groups that can undergo the catalytic mechanism of phosphotransfer in nucleic acid polymerization. Initial forays in this research area demonstrate the wide acceptance of polymerases and augur well for in vivo implementation and integration with canonical metabolism.

1. Introduction

The chemical constitution of biopolymers, found in living organisms, results from the diversity not only of the 'building blocks' i.e. the monomers present in the macromolecules once synthesized, but also from the diversity of the leaving groups used for activating the monomers (which are absent from the biopolymers and released during the synthetic process). This 'backstage scenery' of biosynthesis has so far not been subjected to a systematic study to define the potential for evolving artificial biodiversity and alternative bioprocesses. In the case of nucleic acids, pyrophosphate (PPi) is a universal leaving group that is released from nucleoside and deoxynucleoside triphosphates. No exception to the usage of pyrophosphate as leaving group has ever been reported to occur in any natural cell or virus. This situation contrasts dramatically with protein biosynthesis, where leaving groups occur as informational molecules, in the form of tRNAs, whose sequences vary in a cognate fashion with the chemical structure of the activated amino acids. Here, we explore the design principles of biochemical leaving groups, review the experimental attempts at diversifying leaving groups in nucleic acid biosynthesis in vitro and propose scenarios for introducing such alternative leaving groups in vivo.

Introduction of alternative leaving groups (ALG) distinct from pyrophosphate pertains to both the directed evolution of the natural nucleic acids RNA and DNA, and to the implementation of an additional category of nucleic acid in vivo i.e. XNA. The various configurations to be tackled from wild type (micro)organisms are summarized in Table 1.

2. Leaving group (bio) chemistry

Heterolytic cleavage of a molecule gives rise to two fragments, one of them is called the leaving group. When this leaving group takes away an electron pair it is called nucleofugal. This process usually occurs in the presence of an attacking reagent, in this case a nucleophile. The best leaving groups are the weakest bases, and protonation increases leaving group potential. This is a very common reaction in organic chemistry which is usually carried out in organic solvents. This reaction is also very common in biology

Abbreviations: 3-phosphono-L-Ala, 3-phosphono-L-alanine; ALG, alternative leaving groups; ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; CoASH, coenzyme A; dCMP, deoxycytidine monophosphate; dGMP, deoxyguanosine monophosphate; DNA, deoxyribonucleic acid; dNTP, deoxynucleotide triphosphate; dTMP, thymidine monophosphate; dTTP, thymidine triphosphate; HIV-1, human immunodeficiency virus-type 1; IDA, iminodiacetate; ILA-dAMP, β-imidazole lactic acid-dAMP; KF, Klenow fragment; Km, Michaelis constant; L-Asp, L-aspartic acid; L-Glu, L-glutamic acid; L-His, L-histidine; PAP, phosphoadenosine phosphate; PPi, pyrophosphate; RNA, ribonucleic acid; RT, reverse transcriptase; TPI, triphosphoiminodicarboxylate; tRNA, transfer-RNA; UDP, uridine diphosphate; V_{max}, maximum rate; XNA, xenobiotic nucleic acids

* Corresponding author. Fax: +32 16 337340.
E-mail addresses: piet.herdewijn@rega.kuleuven.be (P. Herdewijn), phmarliere@gmail.com (P. Marlière).

0014-5793/$36.00
doi:10.1016/j.febslet.2012.02.033

Table 1
Stepwise reassignment of leaving groups to canonical (DNA, RNA) and xenobiotic (XNA) nucleic acids in genetically reprogrammed living organisms.

Class of organism	RNA	DNA	XNA
0 (wild type)	PPi	PPi	–
I	ALG	PPi	–
II	PPi	ALG	–
III	PPi	PPi	PPi
IV	PPi	PPi	ALG

where it has to be performed in water, restricting the choice of leaving groups that can be used, for stability and solubility reasons. Several leaving groups are used in biology such as thiols (CoASAc), dialkyl sulphides (S-adenosyl methionin), enolethers (phosphoenolpyruvate), cyclic ureas (carboxylated biotin), but in most cases it is a phosphate or a pyrophosphate group (for example in UDP-glucuronic acid, 3′-phosphoadenosine-5′-phosphosulfate, ATP). UDP, for example, is the most common leaving group in polysaccharide synthesis. In some cases the leaving group may be a large molecule like tRNA in peptide synthesis. In all events, however, the leaving group itself is a common metabolite (CoASH, S-adenosyl homocysteine, pyruvate, biotin, UDP, PAP, ADP, AMP and so on) that can be reused or further processed. The example of protein synthesis is interesting, because it is an example where the leaving group (tRNA) covers information. This is necessary because during protein synthesis information needs to be translated from one biopolymer (RNA) to another (protein). tRNA is, likewise, considered as a crucial molecule in the origin of life and one can ask the question if the principle of putting information in the leaving group is not a general approach to exponential growth and diversification of information [1]. Another example of information transfer during polymerization is that of DNA and RNA synthesis. In this case, however, there is a direct readout of information and pyrophosphate functions as the leaving group. These polymers (nucleic acids and proteins) are central to biological studies and their properties and functions have been studied intensively. However, not much attention has been given to the potential role of the leaving group in metabolism, information and energy transfer. These issues become important when considering the use of artificial nucleic acids for storing information in a cell to be used for defined purposes such as the production of therapeutics or coding for a polymer indispensable for the survival of the cell.

3. Enzymatic nucleic acid synthesis

For developing artificial genes based on synthetic nucleic acids, one has to realize that in a living cell, many functions are carried out by similar molecules. The numbers of chemical principles that are used within a cell are limited. The most striking example is that of nucleoside triphosphates, characterized by the presence of high energy phosphoanhydride bonds that drive their chemistry in a biological context. ATP is involved in information storage and transfer (ATP as a building block for RNA), in energy transfer (ATP as an energy carrier to combine molecules), in regulatory functions (at the level of kinases and signal transduction) and in driving metabolic pathways (ATP as an activated carrier molecule). This multifunctionality of nucleoside triphosphates in general, can be a dangerous trap when trying to replace the natural genetic system by an artificial 'biopolymer' for defined purposes. Therefore it is important to dissect the role of nucleoside triphosphates in building the information system, from their role in other cellular functions (energy regulation, metabolism**),** by means of chemical modification. For safety reasons as well as efficiency reasons, the cellular chassis should function independently from the engineered device.

DNA and RNA are synthesized from nucleoside triphosphates. In this process, one nucleotide is attached (via its 5′-phosphate) to the 3′-hyroxyl group of the growing nucleic acid chain, using pyrophosphate as leaving group. The pyrophosphate is further hydrolyzed to release two phosphate molecules. The nucleoside triphosphates themselves are synthesized from the monophosphates through the action of nucleotide kinases and nucleoside diphosphate kinases. Because of the many functions of nucleoside triphosphates, there is ample room for interference with biological processes when triphosphates of modified nucleosides are brought into the cell (a process that is needed for the propagation of artificial nucleic acids in vivo). Indeed, it would be difficult to install additional nucleoside triphosphates without interfering with DNA and RNA metabolism, cell energy supply or substrate-level phosphorylation.

One way to avoid this intricacy is to use different types of building blocks for the enzymatic synthesis of the natural nucleic acids and for the enzymatic synthesis of the artificial nucleic acids. This can be realized by substituting the pyrophosphate moiety with alternative leaving groups in the precursors for the synthesis of artificial information systems. Indeed, the chemical structure of the leaving group can be elaborated leading to segregation of the polymerization of XNA from that of DNA and RNA and establishing an informational enclave. Another result, as a consequence, is disentangling the XNA polymerization from the phosphoanhydride economy of the cell and thus establishing an energetic enclave as well. In this case, these leaving-group-modified-nucleotides should not be accepted as substrate by regular natural polymerases and vice versa (the evolved polymerase should not accept natural nucleoside triphosphates as building blocks). The use of alternative leaving groups would result in genetic enclaves with metabolic independence, without having to physically separate precursors of the artificial nucleic acids from those of DNA and RNA. Indeed, the implementation of novel leaving groups in metabolism could overcome the need for physical compartmentalization of unnatural information transfers and enable the launching and sustaining of autonomous hereditary procedures in vivo.

The selection of new leaving groups for the enzymatic synthesis of nucleic acids should be accompanied by a selection of appropriate polymerases, which can use the modified nucleotides as building blocks for gene synthesis independent of the cellular gene-synthesis machinery. The activated nucleotides with alternative leaving groups should, likewise, not be recognized by other enzymes involved in cellular functions where nucleoside triphosphates play a role. Given the different biochemical mechanisms involved in polymerization, energy supply and metabolic regulations, this should be within reach. Another prerequisite of an alternative leaving group is that it is non-toxic. If it corresponds or could be converted to a common metabolite, it could fit into recycling pathways, and the accumulation of toxic by-products of the polymerization process could be prevented.

4. Building blocks for XNA synthesis

The ideal properties for such a nucleotide analogue (in vitro and in vivo) would be: water soluble, chemically and enzymatically stable, recognized by the active site of polymerases and serve as substrate for the polymerase; incorporation into the growing nucleic acid should be a mechanistic-based approach, chain elongation should occur, the leaving group should be non-toxic (for example a common cellular metabolite), and the leaving group should be actively degraded or recycled to make the polymerization process irreversible. This mechanistic-based approach is an important issue as it would imply that a relatively stable molecule can be used that is converted into a reactive transition state once