ChemTexts (2017) 3:9
DOI 10.1007/s40828-017-0046-8

LECTURE TEXT

Phosphorous chemistry in vivo: what makes the phosphoesters in DNA and RNA so diverse?

Sabine Müller[1]

Received: 28 March 2017 / Accepted: 17 May 2017 / Published online: 24 May 2017

Abstract Phosphate is an integral part of the molecular backbone of DNA and RNA, constituting the linkage by which the monomeric units are connected. The phosphodiesters in DNA and RNA have distinct stability, owing to the different nature of the sugar units (2′-deoxyribose versus ribose) they connect. This variation in stability can be explained by the 2′-hydroxyl group that is available in RNA, but not in DNA, acting as a neighbor group nucleophile attacking the scissile phosphate. This reaction in nature is supported by enzymes or ribozymes to accelerate the rate of reaction and to ensure specificity. The detailed mechanisms, in particular the geometry of the transition state and of the formed pentavalent phosphorous intermediate, can be rather diverse, involving attack of the nucleophile in line or adjacent to the leaving group. Phosphorothioates as useful analogs of the natural phosphate linkages have been applied to elucidate the stereochemical proceeding of the cleavage reaction allowing to draw conclusions on the mutual orientation of the reacting groups. The following article discusses the basic principles of chemistry inherently linked to the composition of the biomolecular reactants and shows that functionality of the biomolecules has evolved in close relation to their chemical nature.

Keywords Nucleic acid · Phosphodiester · Cleavage · Transesterification · Pseudorotation

✉ Sabine Müller
smueller@uni-greifswald.de

[1] Institute for Biochemistry, Ernst Moritz Arndt University Greifswald, Felix-Hausdorff-Str. 4, 17489 Greifswald, Germany

Introduction

The cleavage and formation of phosphoester bonds are fundamental reactions in biochemistry. For example, phosphorylation of proteins serves as signal for specific metabolic pathways, nucleic acid phosphodiester bonds are targets for enzymatic cleavage or are formed by ligation, and energy stored in phosphoester bonds or, even more importantly, in phosphoric anhydrides, is used for driving chemical reactions.

To understand the in vivo chemistry of phosphoesters, first, one has to look at the valence structure of phosphates (Fig. 1). Belonging to the fifth main group of the periodic table of elements, phosphorous has five outer electrons to provide for chemical binding. As such, in phosphate, phosphorous is bound to four oxygen atoms forming three single and one double bond (Fig. 1a). This results in an overall formal charge of −3, whereby the charge, and with it the double bond character, is delocalized and divided over all four oxygen atoms to ¼ each (Fig. 1b). Only in uncharged phosphotriesters the double bond becomes fully localized (Fig. 1c).

The chemistry of phosphoesters in many aspects is similar to the chemistry of carbonic acid esters, although differences in the electronic structure and hybridization of phosphorous as compared with carbon need to be considered (Fig. 2). Whereas, the carbon center in carbonic acid esters is sp^2-hybridized, hybridization of phosphorous in the corresponding phosphoesters is sp^3. This results in a tetrahedral geometry of the phosphoester as opposed to the trigonal planar geometry around the carbon center in carbonic acid esters. An equally important difference between the two molecules is the strength of the π-bond. In carbonic acid esters, the π-bond is formed by overlapping of 2p orbitals of oxygen and carbon. In contrast, in phosphoesters

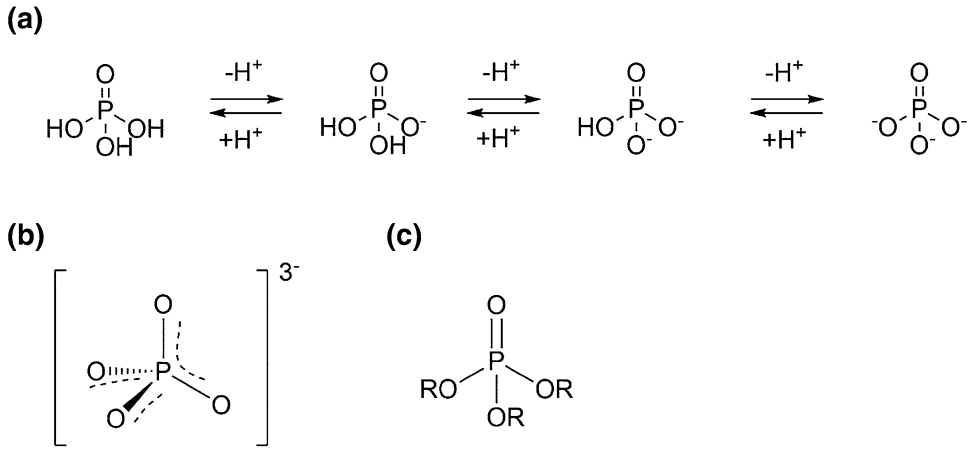

Fig. 1 Phosphoric acid and derivatives. **a** Protonation states of phosphoric acid; **b** valence structure of phosphate with delocalized double bond; **c** valence structure of a phosphotriester with localized double bond

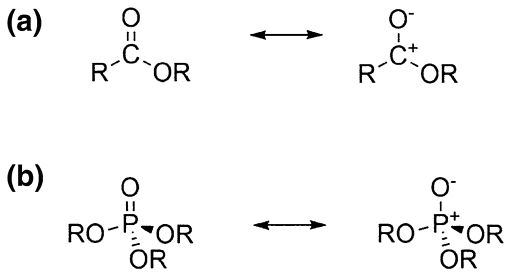

Fig. 2 Double bond character in carbonic acid esters (**a**) and phosphoesters (**b**). In phosphoesters, the ylid structure has higher relevance than in carbonic acid esters

a 2p orbital of oxygen overlaps with a larger 3d orbital of phosphorous to form the π-bond. In consequence, the π-bond in phosphoesters is less stable than in carbonic acid esters, such that for description of the electronic nature of phosphoesters, the mesomeric ylid structure needs to be stronger considered (Fig. 2).

Stability of the phosphodiester bond in nucleic acids

Looking at the two nucleic acids occurring in nature, one may ask the question why their functions are so diverse, and how this diversity is linked to the chemical structure. DNA as the carrier and storage of genetic information needs to be very stable, whereas RNA, being involved in so many regulatory processes in the cell, is required to be produced and degraded on a permanent level. The double helical structure of DNA that is stabilized by hydrogen bonds between nucleobases and between nucleobases and surrounding water molecules, as well as by base stacking between adjacent bases (additionally supported by the methyl group of thymine, which is missing in RNA) is certainly a strong factor contributing to the high stability of DNA. However, comparing the chemical structure of the sugar-phosphate backbones of DNA and RNA—where hydrolysis occurs—one must come to the conclusion that the 2′-OH group present in RNA, but missing in DNA, makes the difference. To evaluate this hypothesis, we first will have a look at the general modes of ester bond cleavage (Fig. 3). Carbonic acid esters as well as phosphoric acid esters can be attacked by nucleophiles at the carbon or phosphorous center; this reaction would be termed as acyl cleavage. Alternatively, a nucleophile can also attack the alcohol carbon of the ester; in this case one speaks of an alkyl cleavage.

Carbonic acid esters are mostly degraded by acyl cleavage, because owing to the carbon center being sp^2-hybridized, the carbon atom is well accessible. However, if soft nucleophiles such as thiols or thiolates are involved, preferentially alkyl cleavage occurs [1–3] (Fig. 3a). In phosphotriesters, the contribution of alkyl cleavage in general is higher, because owing to the tetrahedral geometry, the phosphorous center is stronger shielded in particular if bulky alkyl substituents are present. Thus, dependent on the nature of the nucleophile (soft or hard) and on the size of the substituents, the one or the other mode may be preferred. For example, trimethyl phosphate readily undergoes hydrolysis in neutral water mainly by alkyl cleavage (Fig. 3b). However, in the presence of 1 N NaOH, hydrolysis proceeds mainly by acyl cleavage with the harder nucleophile OH^- attacking the phosphorous center [4] (Fig. 3c). Reaction over 24 h at room temperature delivers the phosphodiester (dimethyl phosphate), which is surprisingly stable. The high stability of the phosphodiester can be explained by electrostatic repulsion between the negatively charged phosphodiester and the nucleophile, which in consequence cannot attack the phosphorous center. Further hydrolysis to the monophosphate can only proceed via alkyl cleavage, which; however, due to the hardness of the nucleophile is very slow (Fig. 3c). Accordingly, dimethyl phosphate in the presence of 1 N NaOH and at 100 °C has a half-life of 16 days

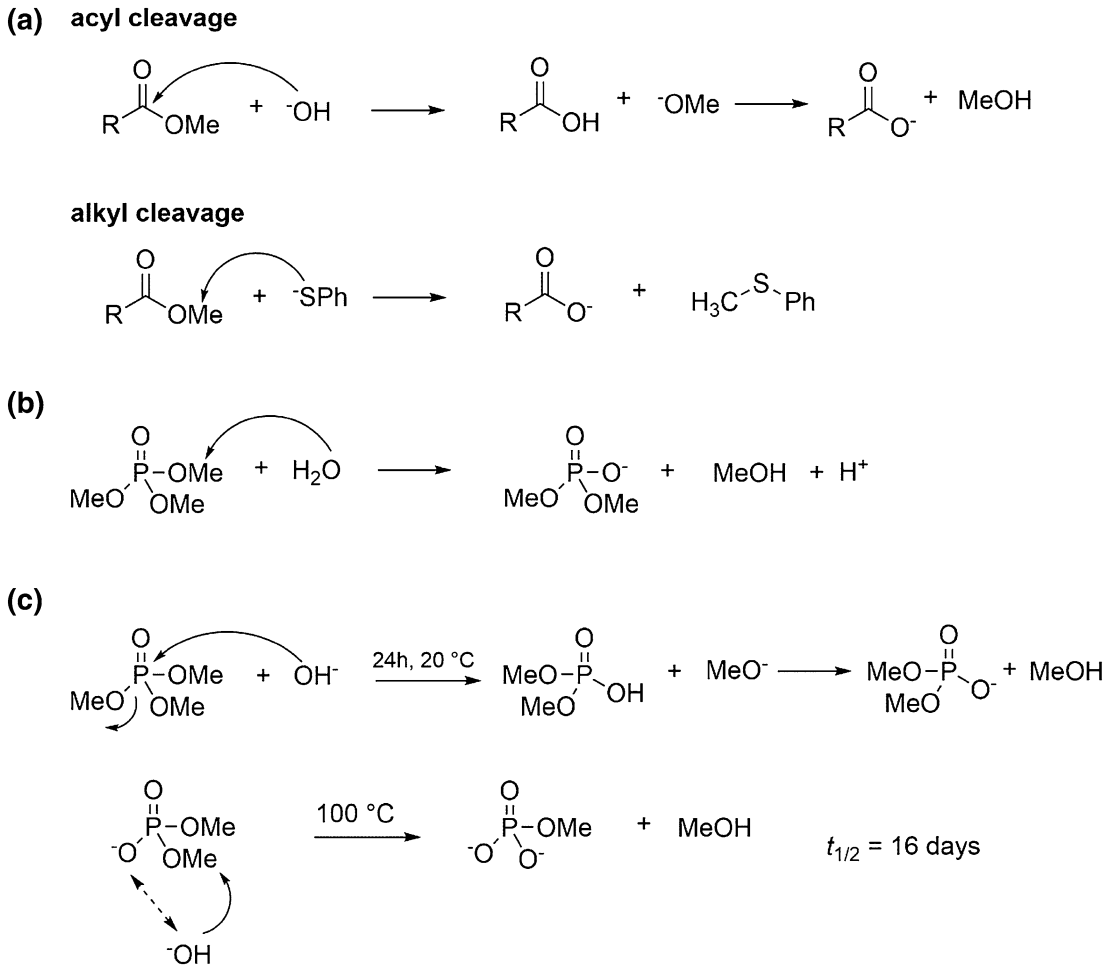

Fig. 3 Cleavage of carbonic acid esters (**a**) and phosphoric acid esters (**b**, **c**)

[4, 5]. This result can be transferred to DNA: in an exemplary experiment, no degradation of phosphodiester bonds is observed when DNA is treated with 1 N NaOH over 1 h (Fig. 4a).

What about RNA? Here, a completely different picture appears: Incubation of RNA in 0.1 N NaOH at room temperature leads to significant cleavage of the phosphodiester bond [6] (Fig. 4b).

Cleavage of phosphodiester bonds in RNA: model studies

As already mentioned above, obviously, the 2′-OH group makes the difference. In the language of the organic chemist, the contribution of the 2′-OH group is named neighboring group participation effect. As a model compound mimicking the situation in the RNA sugar-phosphate backbone, 2-hydroxyethyl methyl phosphate was incubated in 1 N NaOH at 25 °C, and cleavage of the methyl ester was observed to occur with $t_{1/2}$ of 25 min [7]. The explanation for this rather fast degradation follows from the mechanism (Fig. 5). Attack of the phosphorous center by the hard nucleophilic hydroxide ions is hampered due to electrostatic repulsion as explained above, and direct attack of the methyl group as soft electrophile is very slow. However, hydroxide can act as base to abstract a proton from the hydroxyethyl group, thus enhancing the nucleophilicity of the oxygen, which then attacks the phosphorous center. Owing to the intramolecular fashion of the attack, reaction is entropically less expensive and proceeds to a pentavalent intermediate, from which the methoxy group is leaving. The obtained ethylene phosphate with two of the oxygens being bridged and thus forming a five-membered ring is sufficiently energy rich (owing to the ring strain energy) that hydroxide now can attack the phosphorous center, opening the ring and setting the hydroxyethyl group free again. Accordingly, hydrolysis of an RNA phosphodiester bond (Fig. 4b) proceeds via the same mechanism, with the 2′-hydroxyl group taking part as a neighbor group nucleophile.

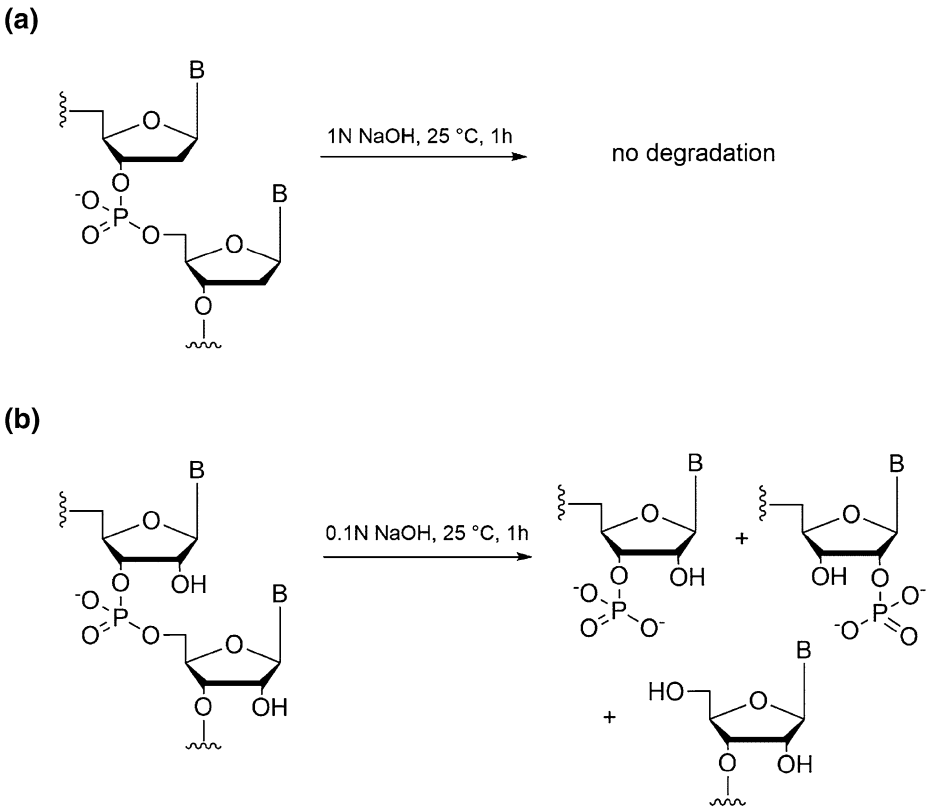

Fig. 4 Distinct stability of DNA (**a**) and RNA(**b**) in sodium hydroxide solution

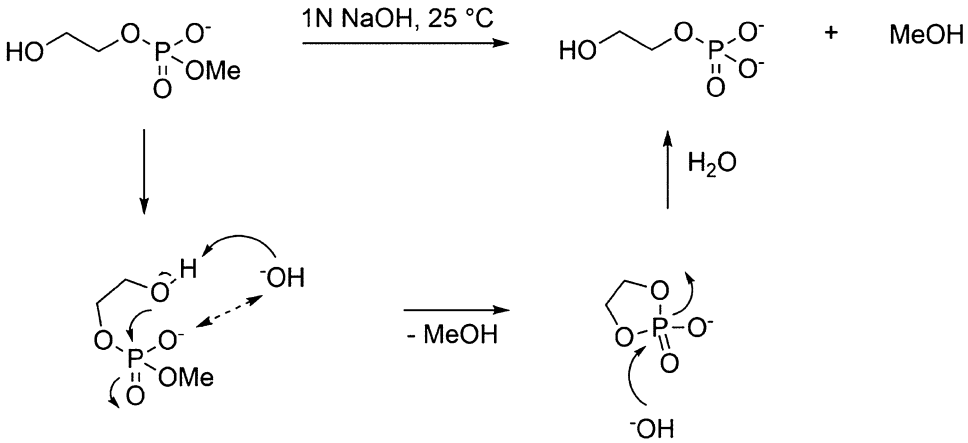

Fig. 5 Ester hydrolysis by neighbor group assistance. Model studies with 2-hydroxyethyl methyl phosphate [6]

Looking at this reaction, a characteristic feature of phosphoester chemistry becomes visible: the hydrolysis of phosphoesters proceeds via a pentavalent intermediate. Basically one may discuss three different mechanisms for hydrolysis (attacking nucleophile is water or OH^-) or transesterification (attacking nucleophile is an alcohol or RO^-) of phosphoesters [8] (Fig. 6). Assuming an S_N2-like mechanism as it occurs for nucleophilic substitution of a carbon center, bond formation and bond cleavage would proceed in a concerted way with the nucleophile attacking from the back side of the tetrahedral phosphorous compound, via the triangular plane formed by the three substituents A, B and O, forming a trigonal–bipyramidal transition state and defining OR′ as leaving group (Fig. 6a). Reaction would proceed with inversion of the configuration of the phosphorous as it occurs in carbon S_N2 reactions. However, in contrast to carbon that cannot form five bonds and, therefore, can arrange five substituents only in a high energy transition state, phosphorous has five valences, and thus can indeed form a more stable pentavalent intermediate. Therefore, in a mechanism called associative hydrolysis/transesterification, the nucleophile can attack the phosphorous center and form the intermediate before the leaving group is released

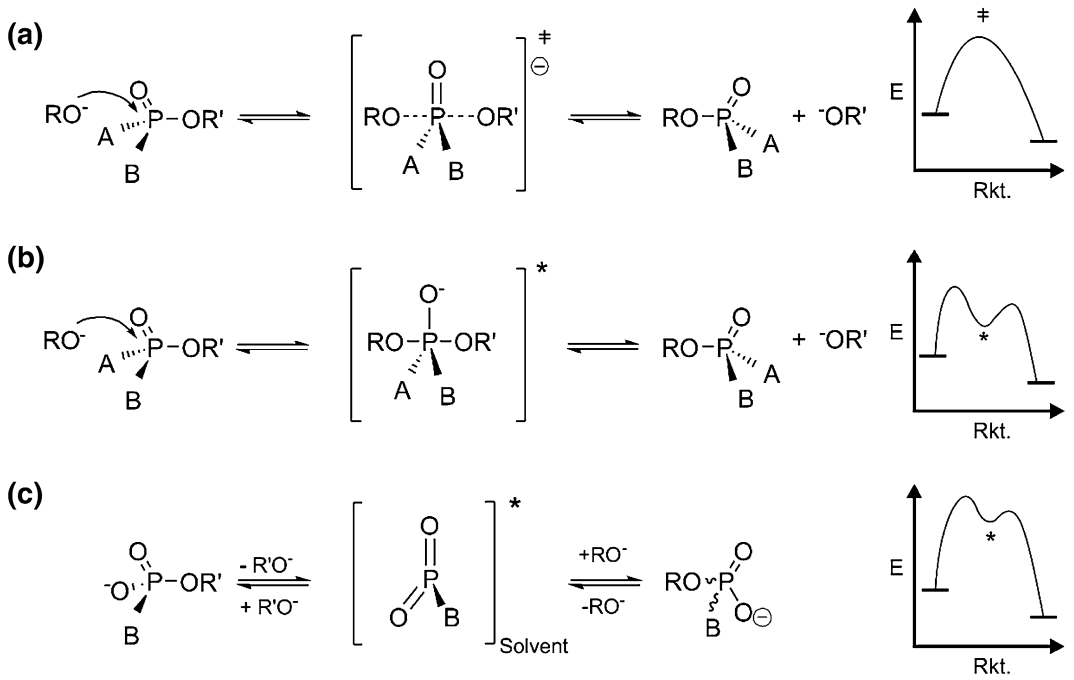

Fig. 6 Cleavage of phosphoesters following a concerted (**a**), associative (**b**) or dissociative (**c**) mechanism, including the energetic reaction profiles

(Fig. 6b). This mechanism resembles the addition–elimination pathway proceeding in carbonic acid reactions, but from the sterical perspective, it is very close to the concerted hydrolysis/transesterification; it also proceeds with inversion of the configuration of the phosphorous center. However, the pentavalent state would be a reactive intermediate rather than a transition state as mirrored in the energetic profiles of the two alternative reaction paths (Fig. 6a, b). A third possibility is a dissociative mechanism. Here, bond cleavage and release of $R'O^-$ would occur prior to attack of the nucleophile and bond formation (Fig. 6c).

This reaction path resembles carbon S_N1 chemistry. The trivalent phosphorous intermediate is trigonal planar, such that the nucleophile can attack from both sides forming a racemic mixture of enantiomers.

Looking at the transition states/intermediates of these three hypothetical reaction pathways, the associative mechanism is the most plausible, owing to the favorable energetic profile. The pentavalent intermediate is thermodynamically favored over the trigonal–bipyramidal transition state, making the concerted mechanism less likely. A dissociative path is least plausible for naturally occurring phosphoesters. The high energy trigonal intermediate would require further stabilization by organic solvents, but is insufficiently stable in water (the natural solvent).

Nucleophilic attack in line or adjacent?

Summing up what we have learned above, the hydrolysis/transesterification of phosphoesters proceeds by nucleophilic attack preferentially on the phosphorous center, and in case of RNA profits from the neighbor group participation effect of the 2′-OH functionality. The favored reaction mechanism is addition of the nucleophile to form a pentavalent intermediate with trigonal bipyramidal geometry and subsequent release of the leaving group (associative mechanism). In analogy to carbon S_N2 chemistry, attack of the nucleophile proceeds in line with the leaving group, meaning that both, attacking nucleophile and leaving group, are positioned at the apical positions of the bipyramid (Fig. 6).

For hydrolysis of a cyclic phosphate (Fig. 7), this scenario would define one of the bridging oxygens as leaving group, thus opening the strained ring and producing 2-hydroxyethyl methyl phosphate as the main product. However, careful inspection of the product mixture revealed that in addition to 2-hydroxyethyl methyl phosphate, cyclic ethylene phosphate is formed, although as minor product [6]. At first glance, this could be partially explained by the attack of the nucleophile in line with the methoxy substituent. However, this scenario would force the two oxygens linked by the ethylene bridge in equatorial positions of the bipyramide, which is sterically highly demanding

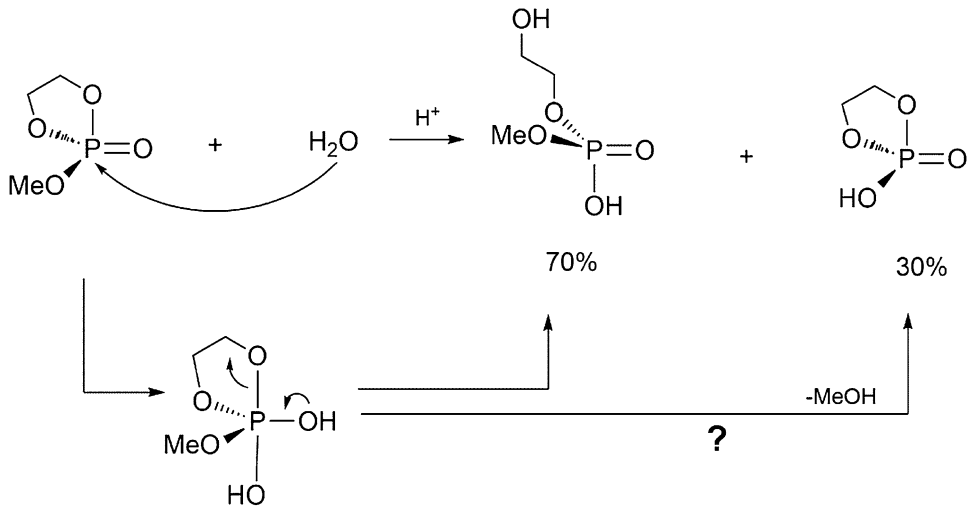

Fig. 7 Hydrolysis of cyclic methyl ethylene phosphate

and thus associated with a high energy barrier. The bipyramidal intermediate would preferentially form with one of the bridging oxygens located in equatorial position, the other in apical position, to reach an unstrained and thus energetically more favorable geometry (Fig. 7). How then the occurrence of cyclic ethylene phosphate may be explained?

Some insight comes from studies of pentafluoro phosphine—a compound with trigonal–bipyramidal geometry, phosphorous in the center and covalently bound fluorine atoms in the triangular plane (equatorial positions) and at the two peaks (apical positions) (Fig. 8). The F–P–F angles in the triangular plane are 120°, the two apical fluorine atoms are perpendicular to the plane. Bond lengths between phosphorous and the three equatorial fluorine atoms are 153 pm, and between phosphorous and the two apical fluorine atoms are 158 pm. Thus, the geometry of the bipyramide defines two sorts of fluorine atoms (equatorial and apical). Nevertheless, as measured by ^{19}F-NMR, only one signal instead of two is obtained for the five fluorine atoms [9]. This phenomenon may be explained by pseudorotation—a phenomenon that inherits the fast dynamic change of atomic positions in molecules without bond breaking. The fluorine atoms permanently change their positions in the bipyramide owing to bending vibrations. This process is too fast to be resolved by NMR, such that only a time-averaged signal is obtained.

As depicted in Fig. 8a, atoms number 4 and 5 at the apical positions bend towards each other, thus shortening the bond lengths and decreasing the bond angle from 180° to 120°. Equivalently, atoms number 2 and 3 shift apart, thus increasing the distance to each other and to the phosphorous, and widening the bond angle from 120° to 180°. The remaining fluorine atom (#1) in equatorial

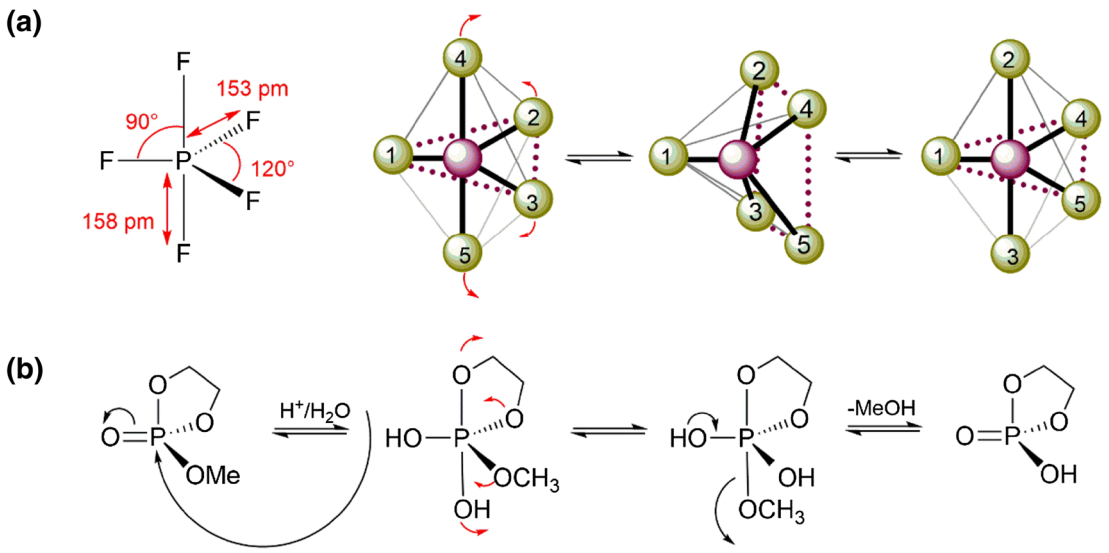

Fig. 8 Pseudorotation at the example of pentafluoro phosphine (**a**) and during hydrolysis of cyclic methyl ethylene phosphate (**b**)

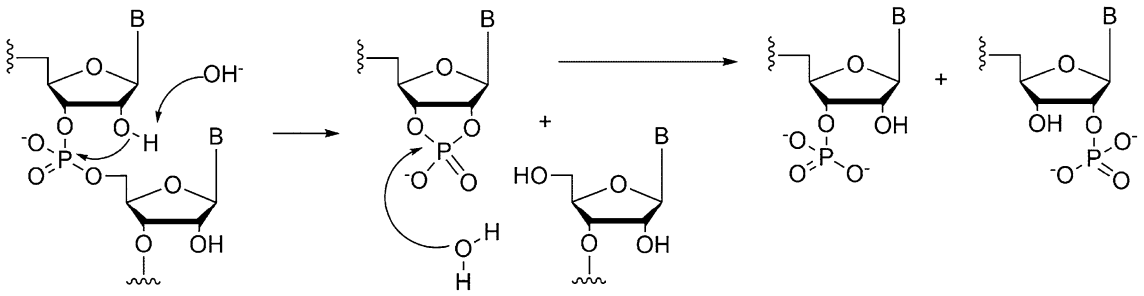

Fig. 9 RNA cleavage in two steps

position serves as the pivotal point, to which the other atoms move in relation. In the process, a transition state with quadratic pyramidal geometry is formed. The product is again a trigonal bipyramide; however, apical substituents are now in equatorial positions and two of the former equatorial substituents are now apical. Each of the three equatorial substituents can be the pivot, such that in the aggregate all fluorine atoms change positions between equatorial and apical and vice versa [10].

Pseudorotation plays also a role in the pentavalent intermediates formed during phosphoester hydrolysis/transesterification. Going back to the reaction shown in Fig. 7, the attacking nucleophile and one of the bridging oxygens are located at the apical positions of the bipyramide. As already discussed above, this geometry would define the bridging oxygen in the apical position, in line with the nucleophile, as leaving group and result in the formation of hydroxyethyl methyl phosphate as product. However, pseudorotation can lead to rearrangement of atoms: OH and the bridging oxygen in apical positions become equatorial substituents and consequently, two of the former equatorial substituents (for sterical reasons one of them must be the other bridging oxygen, see above) change to apical positions. In the resulting bipyramide, now the methoxy group is at an apical position and can leave the intermediate to form the cyclic ethylene phosphate as product (Fig. 8b). In comparison to the first reaction path, where nucleophilic attack occurred in line to the leaving group, here the water nucleophile attacks the phosphorous not in line, but adjacent to the methoxy leaving group, and the formation of the products can be

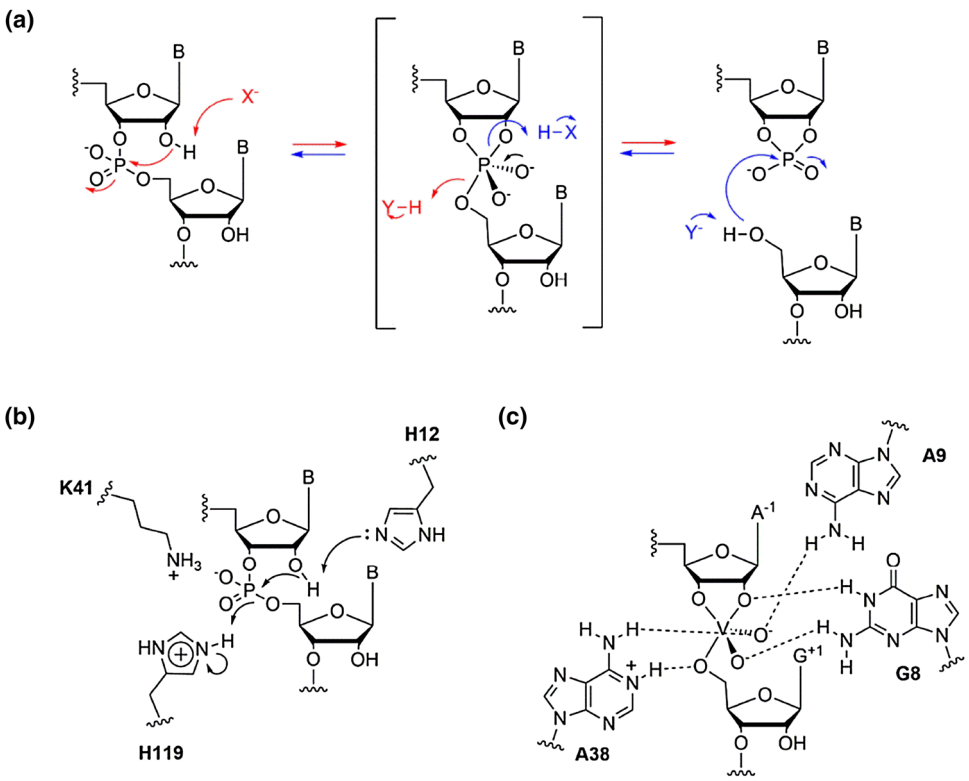

Fig. 10 RNA cleavage by enzymes and ribozymes. **a** General mechanism of enzyme or ribozyme supported transesterification leading to RNA cleavage (from *left* to *right*, *red colored* electron transfer), or RNA ligation (from *right* to *left*, *blue colored* electron transfer), X^-/Y^-: general base, Y–H/X–H: general acid; **b** active site chemistry of RNase A [11]; **c** active site of the hairpin ribozyme with a vanadate transition state analog [15]

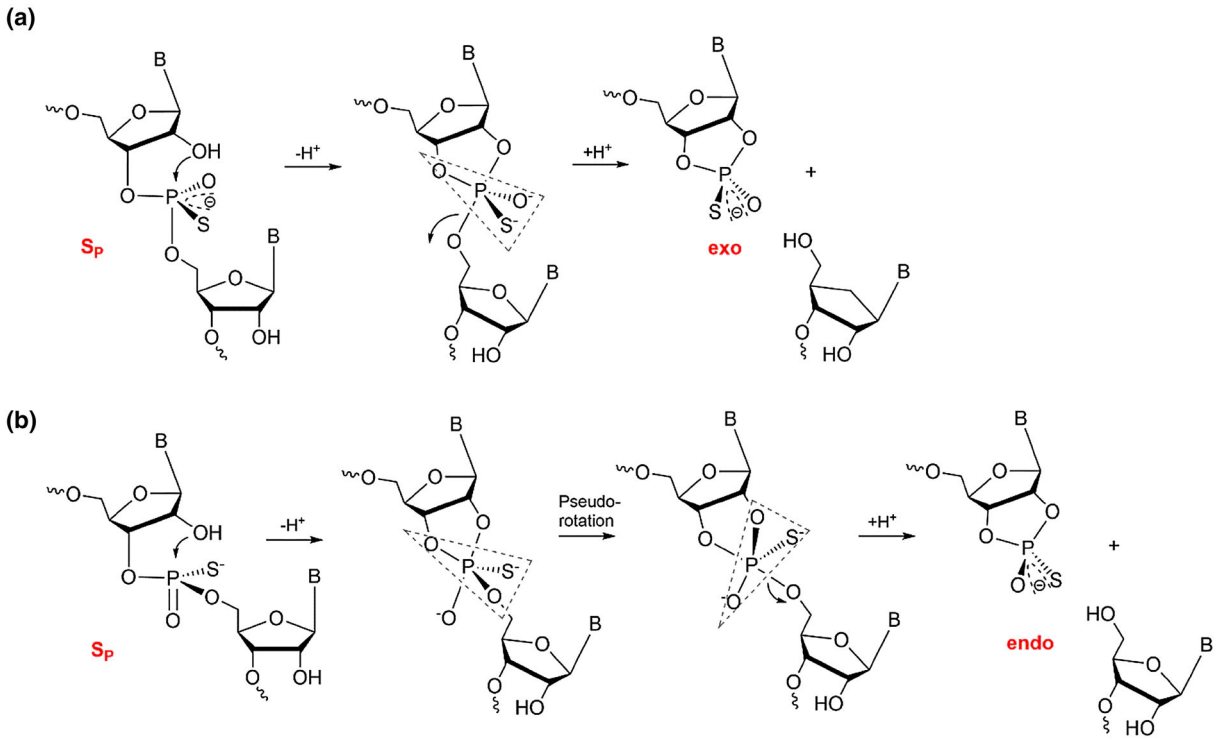

Fig. 11 Cleavage of a phosphorothioate bond by in line (**a**) or adjacent (**b**) attack of the 2′-OH on the neighboring phosphorous atom. In line attack (**a**) proceeds with inversion of the configuration at the P-center, such that the *exo*-configured cyclic thiophosphate is produced from the S-epimer (Sp). Adjacent attack (**b**) including pseudorotation proceeds with retention of configuration, the *endo*-isomer is produced from the S-epimer

explained only by taking into account the process of pseudorotation.

Chemical and enzymatic RNA cleavage

Projecting what we have discussed above on the cleavage of a natural RNA phosphodiester bond, the initial products are a fragment with terminal 2′,-3′-cyclic phosphate and a fragment with terminal 5′-OH group (Fig. 9). This first transesterification step is followed by hydrolysis of the cyclic phosphate, which owing to its strained structure is still sufficiently reactive. Thus, a water molecule attacks the phosphorous of the cyclic phosphate, whereby one of the two oxygens, either 2′-O or 3′-O, is released as a leaving group. As discussed for the first transesterification step, hydrolysis also can proceed by in line or adjacent attack of the water nucleophile.

The chemical cleavage of an RNA phosphodiester backbone can be easily induced by hydroxide to abstract the proton of the 2′-OH group followed by nucleophilic attack onto the neighboring phosphorous as explained above. Owing to the more favorable energy landscape, in line attack is the preferred reaction as compared with adjacent attack and pseudorotation. In single stranded RNA, the 2′OH group is transiently in an in line position to the 5′-O-leaving group, such that transesterification can take place. Double-stranded RNA is more rigid and thus more resistant to base-induced cleavage.

In nature, RNA is cleaved by a number of enzymes, many of them supporting the mechanisms described above and shown in Fig. 10a. For example, RNase A employs histidine residues as general base to abstract the 2′-OH proton, and also as general acid to deliver a proton to the 5′-O-leaving group [11] (Fig. 10b). In addition, enzyme catalysis includes conformational strategies to orient the substrate in the preferred in line or eventually adjacent geometry, and equally important, charged transition states and intermediates are stabilized by electrostatic neutralization of the additional negative charge by suitable amino acid side chains in the catalytic center of the enzyme [12]. In addition to proteinaceous enzymes for RNA cleavage (RNases), nature also harbors a number of RNA catalysts (ribozymes) for specific RNA cleavage [13]. RNA cleaving ribozymes use the same strategies for catalysis as protein enzymes do: general acid–base catalysis, conformational effects to support in line or adjacent geometry and stabilization of transition states and of the pentavalent intermediate. The role of the amino acid side chains acting as Brönsted or Lewis base or acid or providing positive

charges for electrostatic stabilization, in ribozymes are taken over by the nucleobases. For example, the hairpin ribozyme- a naturally occurring small RNA- uses a deprotonated guanosine in the catalytic core as general base, and a protonated adenosine as acid to support RNA cleavage [14] (Fig. 10c).

Investigation of the enzymatic mechanism involving in line or adjacent attack of the nucleophile onto the phosphorous can be conducted by stereochemical analysis. Phosphodiesters are non-chiral. Therefore, for stereochemical analysis, the natural phosphodiester in RNA is replaced by a phosphorothioate at the site of interest. This can be easily done by chemical assembly of the RNA chain, which; however, delivers a diastereomeric mixture with R or S configuration of the phosphorothioate (R_P, S_P) (Fig. 11). The diastereomers can be separated by HPLC to obtain the two epimers. Enzymatic (or ribozymatic) cleavage of the S_P epimer would produce the cyclic phosphate, in case of an in line mechanism with inverted configuration at the phosphorous, thus producing a cyclic thiophosphate in *exo*-configuration. In case of an adjacent mechanism reaction proceeds with retention of the phosphorous configuration, producing the cyclic thiophosphate in *endo*-configuration [16, 17] (Fig. 11). Using this strategy, it was proven that both steps (formation of 2′,3′-cyclic phosphate and subsequent hydrolysis to 3′-phosphate) of RNase A mediated cleavage of RNA proceed by an in line mechanism [18, 19].

In summary, the hydrolysis of phosphate esters is a key reaction in biology, catalyzed by enzymes and modulated by the chemical nature of the phosphate esters themselves. The high stability of phosphate diesters under physiological conditions is an important property of nucleic acids, and in particular with respect to DNA, a vital feature of their biological role: the storage and preservation of genetic information. In RNA, the bridging phosphate is an ester of a 1,2-diol, with the consequence that the vicinal hydroxyl group enormously enhances the rate of hydrolysis. Thus, activation of a specific 2′OH group in RNA by enzymes or ribozymes leads to facile cleavage of the neighboring phosphodiester. This characteristic feature is linked to the biological role of RNA, being a molecule that underlies spatially and temporarily controlled production and degradation processes. Thus, natural transformations underlie the basic principles of chemistry inherently linked to the molecular composition of the reactants. Accordingly, functionality of the biomolecules has evolved in close relation to their chemical nature.

References

1. Bunton CA, Wood JL (1955) Tracer studies on ester hydrolysis. Part II. The acid hydrolysis of tert-butyl acetate. J Chem Soc 1522–1525
2. Bruice TC, Fife TH (1961) A facile base-catalyzed ester hydrolysis involving alkyl-oxygen cleavage. the mechanism of hydrolysis of esters of 4(5)-hydroxymethylimidazole. J Am Chem Soc 83:1124–1128
3. Cohen SG, Schneider A (1941) Cleavage of the alkyl-oxygen bond in the hydrolysis of esters. t-Butyl 2,4,6-trimethylbenzene. J Am Chem Soc 63:3382–3388
4. Barnard PWC, Bunton CA, Llewellyn DR, Vernon CA, Welch VA (1961) The reactions of organic phosphates. Part V: the hydrolysis of triphenyl and trimethyl phosphates. J Chem Soc 2670–2676
5. Thatcher GRJ, Kluger R (1989) Mechanism and catalysis of nucleophilic substitutions in phosphate esters. Adv Phys Org Chem 25:99–265
6. Dugas H (1999) Bioorganic Chemistry, 3rd edn. Springer, New York
7. Oivanen M, Kuusela S, Lönnberg H (1988) Kinetics and mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by Bronsted acids and bases. Chem Rev 98:961–990
8. Anslyn EV, Perrault DM (1997) Unifying the current data on the mechanism of cleavage-transesterification of RNA. Angew Chem Int Ed 36:432–450
9. Gutowsky HS, McCall DW, Slichter CP (1953) Nuclear magnetic resonance multiplets in liquids. J Chem Phys 21:279–292
10. Cass M, Hii KK, Rzepa HS (2006) Mechanisms that interchange axial and equatorial atoms in fluxional processes: illustration of the Berry pseudorotation, the turnstile and the lever mechanisms via animation of transition state normal vibrational modes. J Chem Ed 83:336–336
11. Yang W (2011) Nucleases: diversity of structure, function and mechanism. Q Rev Biophys 44:1–93
12. Marshall GR, Feng JA, Kuster DJ (2008) Back to the future: ribonuclease A. Biopolymers 90:259–277
13. Jimenez RM, Polanko JA, Luptak A (2015) Chemistry and biology of self-cleaving ribozymes. Trends Biochem Sci 40:648–661
14. Ferré-D'Amaré AR (2004) The hairpin ribozyme. Biopolymers 73:71–78
15. Rupert PB, Massey AP, Sigurdsson S, Ferre-D'Amaré AR (2002) Transition state stabilization by a catalytic RNA. Science 298:1421–1424
16. Taira K (1987) Stereoelectronic control in the hydrolysis of RNA by imidazole. Bull Chem Soc Jpn 60:1903–1909
17. Eckstein F (1983) Phosphorothioate analogues of nucleotides. Tools for the investigation of biological processes. Angew Chem Int Ed 22:423–506
18. Deakyne CA, Allen LC (1979) Role of active site residues in the catalytic mechanism of ribonuclease A. J Am Chem Soc 101:3951–3959
19. Usher DA, Ehrenrich ES, Eckstein F (1972) Geometry of the first step in the action of ribonuclease A. Proc Natl Acad Sci 69:115–118