Chemical Configurations: Proteins and DNA

Sarah Maurer
Central Connecticut State University
Department of Chemistry and Biochemistry

Prebiotic mixture of chemicals

Origin of life

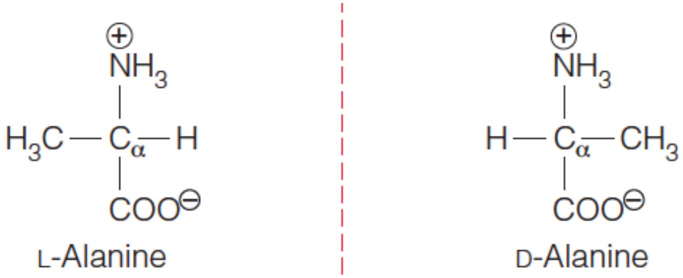
20 L - α - amino acids
5 D – nucleotides

Table 1. Yields (based on carbon) of products following passage of a spark discharge through a mixture of CH_4, NH_3, H_2O, and H_2 (total yield: 15.2% [22]).

Compound	Yield (%)	Compound	Yield (%)
Formic acid	4.0	α-Hydroxybutyric acid	0.34
Glycine	2.1	Succinic acid	0.27
Glycolic acid	1.9	Sarcosine	0.25
Alanine	1.7	Iminoacetic-propionic acid	0.13
Lactic acid	1.6	*N*-Methylalanine	0.07
α-Alanine	0.76	Glutamic acid	0.051
Propionic acid	0.66	*N*-Methyl urea	0.051
Acetic acid	0.51	Urea	0.034
Iminodiacetic acid	0.37	Aspartic acid	0.024
α-Amino-*n*-butyric acid	0.34	α-Aminoisobutyric acid	0.007

Podlech, J. (2001). Origin of organic molecules and biomolecular homochirality. *Cellular and Molecular Life Sciences*, *58*(1), 44–60.

Chirality



L-amino acids – enriched by meteorites

Table 2. Enantiomeric distributions of the *Allo* isoleucine and isoleucine diastereomers in selected CR2 meteorites

	LAP*	EET	MET	PCA	QUE	MIL	GRA1*
Ile L-ee	3.6	26†	50	19	50†	46†	14.0
*Allo*ile D-ee	2.2	21	60	19	34.5	18	12.1
Allo/ile	1.4	1.8†	2.2	2.0	1.1†	1.4†	2.3

*From ref. 3.
†Shows L-proteinogenic amino acid excesses in the extracts, could be in part contaminant and ile/*allo* ratios were estimated from L-*allo*/D-ile.

D - sugars (also in nucleotides) – enriched by L - amino acids

Pizzarello, S., Schrader, D. L., Monroe, A. A., & Lauretta, D. S. (2012). Large enantiomeric excesses in primitive meteorites and the diverse effects of water in cosmochemical evolution. *PNAS*, *109*(30), 11949–11954. https://doi.org/10.1073/pnas.1204865109

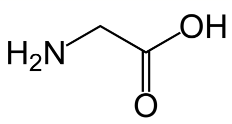
Amino acids

Biological	Alternatives
Amino acid	hydroxy acids, thio acids, amino sulfonic- or amino phosphinic acids
Residues at the alpha carbon	β-, γ-, or δ-amino acids, or other dervatives
20 exactly	More or less than 20
Our specific set of 20	Other amino acids that were available prebiotically

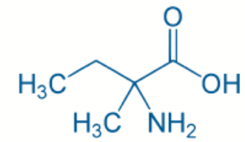
Constraints on origin of amino acids:

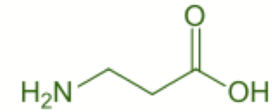
- Prebiotic availability
- Metabolic accessibility/compatibility
- Evolutionary history/functional utility

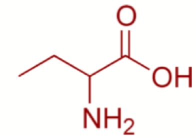
Prebiotic amino acids

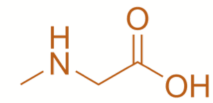


Standard
Dialkyl-amino acid
β- γ- substituted
Non-standard
Amino substituted





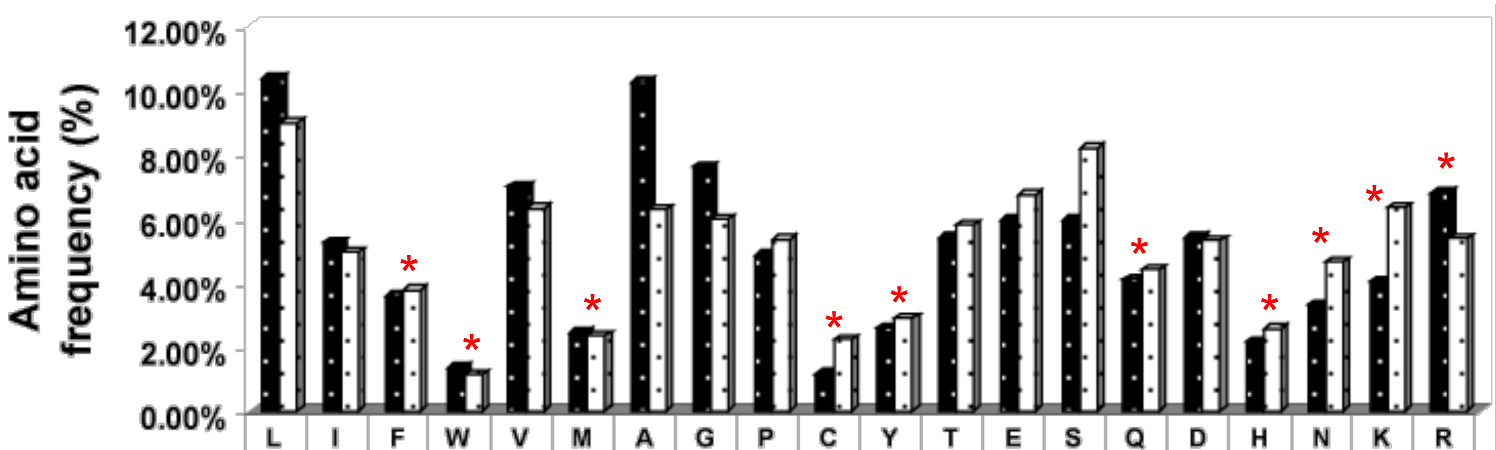




Amino acid	Murchison	Spark discharge
Glycine	++++	++++
Alanine	++++	++++
α-Amino-n-butyric acid	+++	++++
α-Aminoisobutyric acid	++++	++
Valine	+++	++
Norvaline	+++	+++
Isovaline	++	++
Proline	+++	+
Pipecolic acid	+	< +
Aspartic acid	+++	+++
Glutamic acid	+++	++
β-Alanine	++	++
β-Amino-n-butyric acid	+	+
β-Aminoisobutyric acid	+	+
γ-Aminobutyric acid	+	++
Sarcosine	++	+++
N-Ethylglycine	++	+++
N-Methylalanine	++	++

Cleaves, H. J. (2010). The origin of the biologically coded amino acids. *J. Theor. Biol.*, *263*(4), 490–498. https://doi.org/http://dx.doi.org/10.1016/j.jtbi.2009.12.014

Reduced alphabet



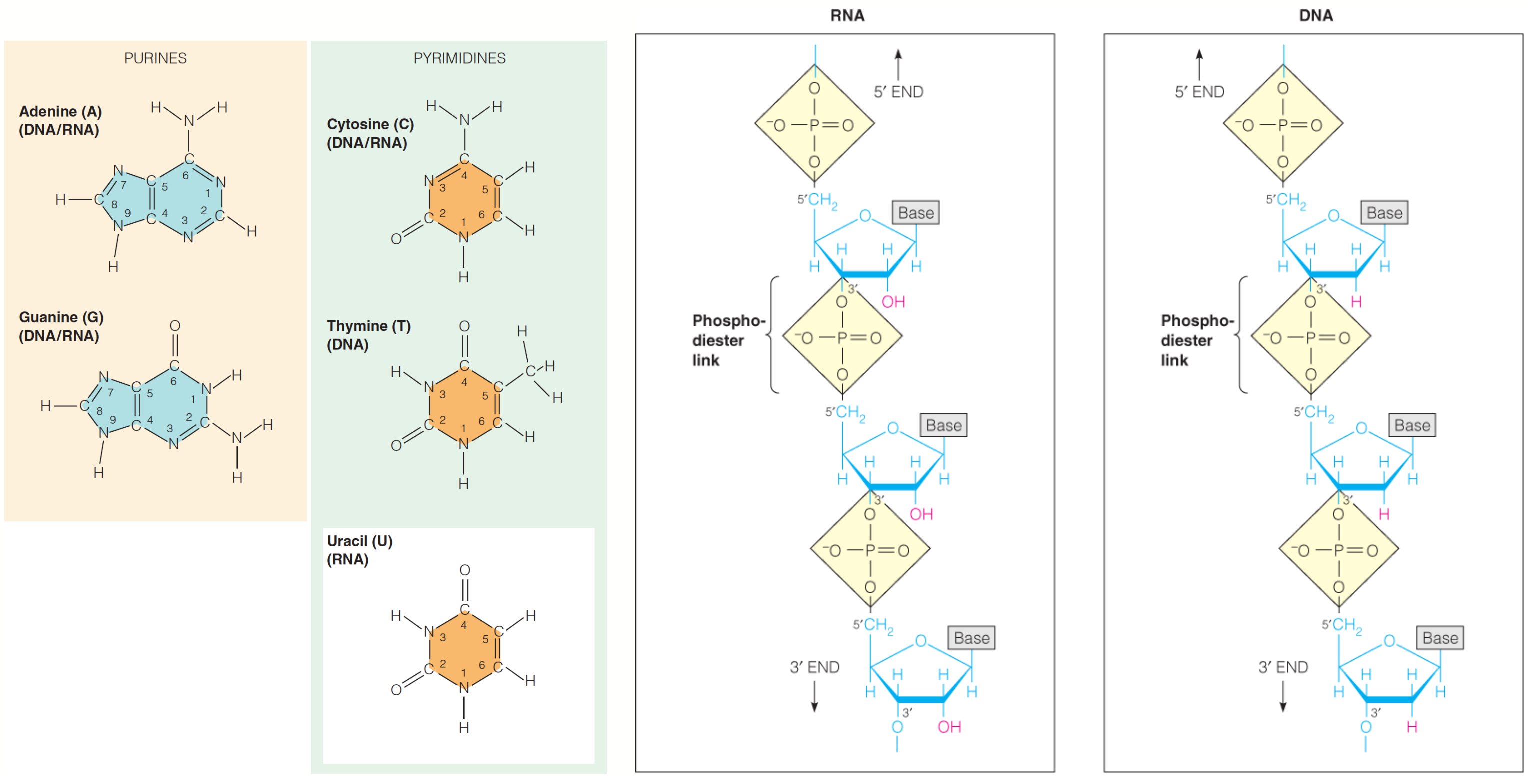
	L	I	F	W	V	M	A	G	P	C	Y	T	E	S	Q	D	H	N	K	R
■ Prok	10.4	5.26	3.60	1.35	7.05	2.44	10.3	7.65	4.87	1.15	2.60	5.43	5.98	5.98	4.10	5.43	2.16	3.33	4.07	6.83
□ Euk	9.01	4.98	3.78	1.14	6.34	2.36	6.32	6.01	5.37	2.21	2.91	5.81	6.76	8.23	4.42	5.35	2.55	4.65	6.38	5.41
Meteorite	+	+	-	-	+	-	+	+	+	-	-	+	+	+	-	+	-	-	-	-
Spark discharge	+	+	-	-	+	-	+	+	+	-	-	+	+	+	-	+	-	-	-	-
HCN polymerization	+	*	-	-	+	-	+	+	-	-	-	+	+	+	-	+	*	-	*	-

+ Detected with high confidence, * detected with low confidence, – not detected.

More complex amino acids were likely selected based on other criteria (not prebiotic availability), possibly with later adaptation into biochemistry

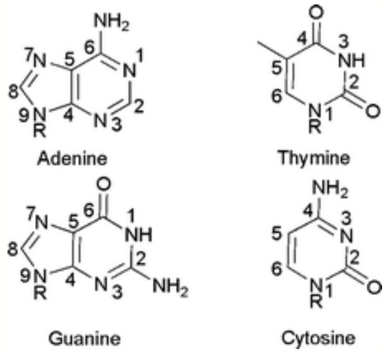
- Sulfur containing
- Aromatics
- Nitrogen containing

Nucleic acid structure



phosphate Base
Sugar
phosphate Base
Sugar
phosphate Base
Sugar

Natural xNA



- Modification of expression
- Timing replication
- Controlling DNA transposition
- Protect from host degradation
- Regulation of repair

TABLE 1. *Modified DNA bases*

Base	Structure of substituent	Organism	Percentage replacement of standard base	References
5-Methylcytosine	CH_3	higher eukaryotes	5–30%	48, 80
		Xanthomonas oryzae phage XP12	100%	81
		Other phages (e.g., ϕX174)	0.2–0.5%	80, 82
		Dinoflagellates	2–17%	60
		Chlamydomonas	0.7%	59
		Fungi	1–5%	Reviewed in 53
		Bacteria	0.3–2%	Reviewed in 80
N^6-Methyladenine	CH_3	*Tetrahymena* and other ciliates	0.8–2.5%	49; reviewed in 59
		Dinoflagellates	10%	60
		Chlamydomonas	0.5%	59
		Bacteria	0.3–3%	Reviewed in 80
		Phages (e.g., T_2, T_4)	0.5–2%	80, 83
5-Hydroxymethylcytosine (hexosylated)	CH_2OR R_1: H R_2: α-glucose R_3: β-glucose R_4: β-glucose-α-glucose	*Escherichia coli* phages T_2, T_4, T_6	$R_1+R_2+R_3+R_4$ = 100%	8, 15
5-Hydroxymethyluracil	CH_2OH	*Bacillus subtilis* phages SP8, ϕe, SPO1, H1, SP82G, 2C, ϕ25	100%	16; reviewed in 28
		Dinoflagellates	12–68%	50, 60
Uracil	—	*B. subtilis* phage PBS2	100%	84
α-Putrescinylthymine	$NH(CH_2)_4NH_2$	*Pseudomonas acidovorans* phage ϕW14	50%	13
Sugar-substituted 5-dihydroxypentyluracil	$CH_2(CH_2)_2CHORCH_2OR$ R: glucose[a] R: glucuronolactone-1-phosphate[a]	*B. subtilis* phage SP15	62%	9, 31
a-Glutamylthymine	COOH NHCH $(CH_2)_2$ COOH	*B. subtilis* phage SP10	15–20%	See ref 1
7-Methylguanine	CH_3	*Shigella sonnei* phage DDVI	1%	85
2-Aminoadenine	NH_2	*S. elongatus* phage S-2L	100%	20
N^6-carbamoylmethyladenine	CH_2CONH_2	*E. coli* phage Mu	15%	17
N^4-methylcytosine	CH_3	bacteria	0.5–2%	86, 87
Hexosylated 5-hydroxycytosine	OR R_1: (D-gal)-1-α-D-glc-6→1-α-D-glc R_2: (D-gal)-1-α-D-glc R_3: D-gal	*Rhizobium phage* RL38JI	100%	10
β-D-glucosyl-hydroxymethyluracil	CH_2OR R: β-D-glucose	*Trypanosoma brucei*	0.4%	51

[a] It has not been established yet whether glucose is attached to the 4' position and glucuronolactone-1-phosphate to the 5' position or vice versa.

Why ribose?

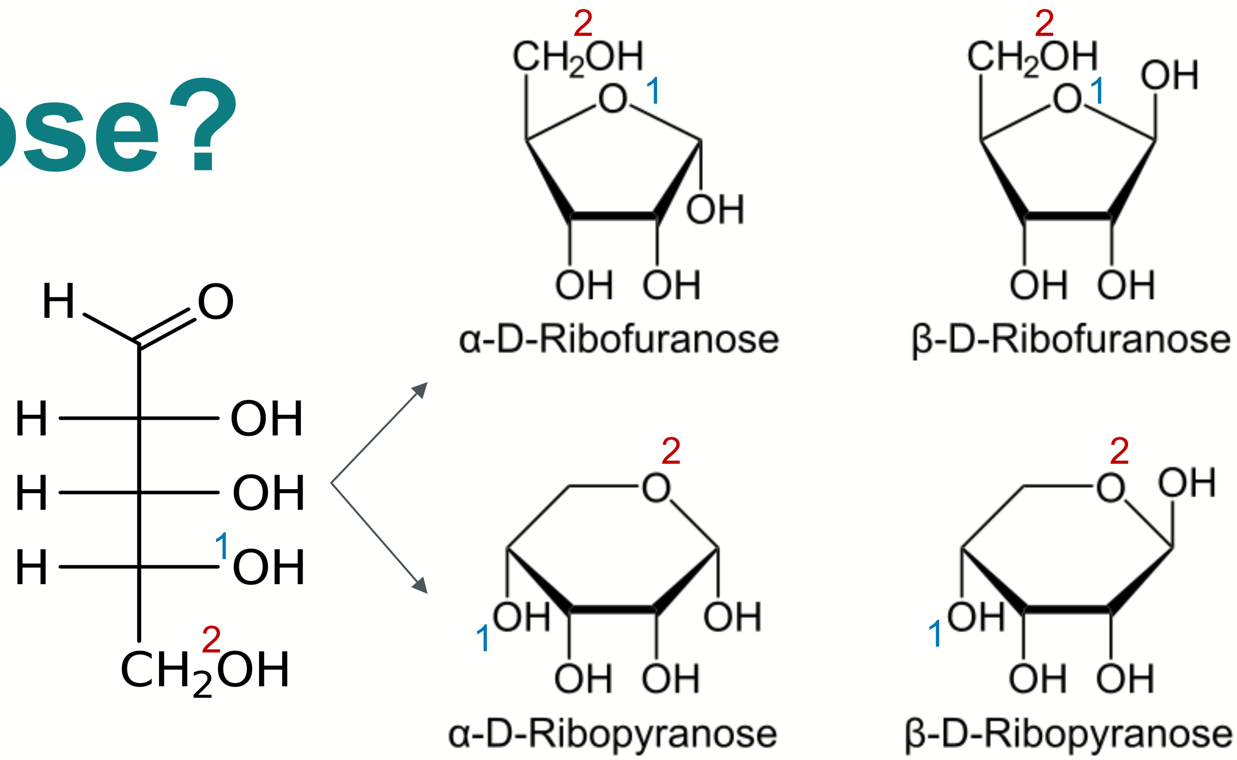


Table 1. Populations of Different Forms of Aldopentoses in Aqueous Solution[40]

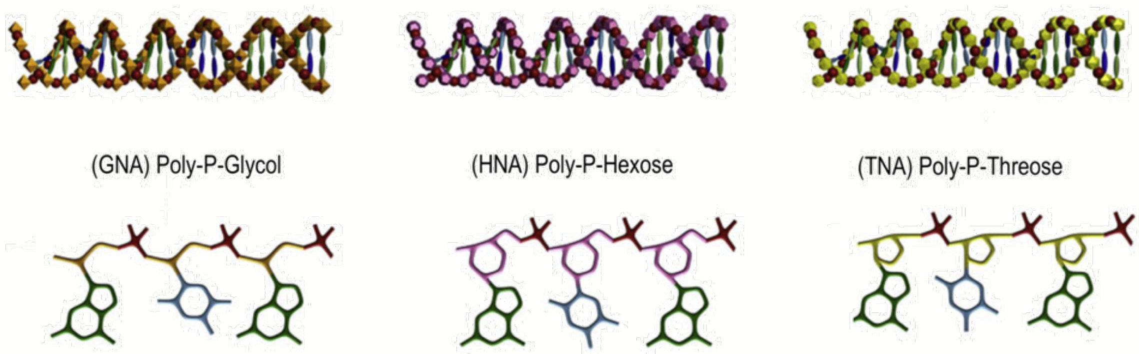
	pyranose		furanose	
	α-form (%)	β-form (%)	α-form (%)	β-form (%)
D-ribose	21.5	58.5	6.5	13.5
D-xylose	36.5	63.0	0	0
D-arabinose	60.0	35.5	2.5	2

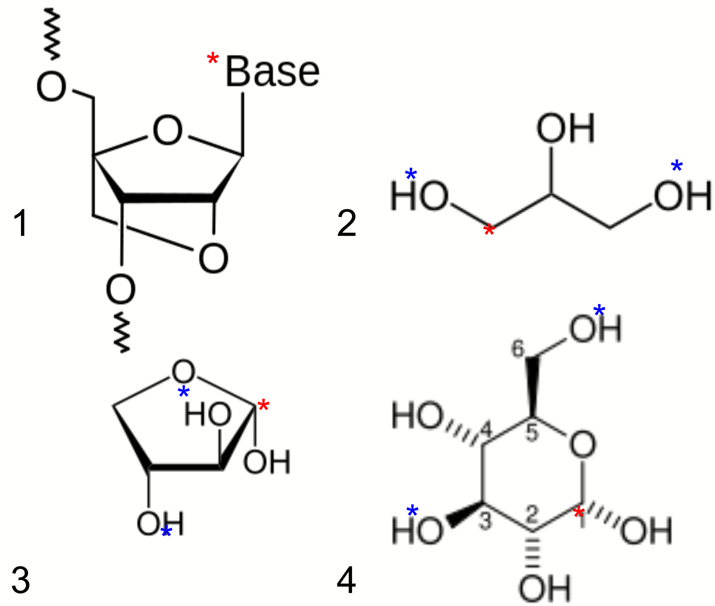
Wei, C. Y., & Pohorille, A. (2009). Permeation of Membranes by Ribose and Its Diastereomers. *JACS*, *131*(29), 10237–10245.

Why ribose?

Carbons	
3	Glycerol§
4	Erythritol§ DL-Threitol
5	Adonitol Arabitol Xylitol
6	Dulcitol Mannitol
7	Sorbitol
5	Arabinose Lyxose Ribose D-Xylose L-Xylose Ribulose
6	Galactose Glucose§ Mannose Fructose§
7	L-Sorbose

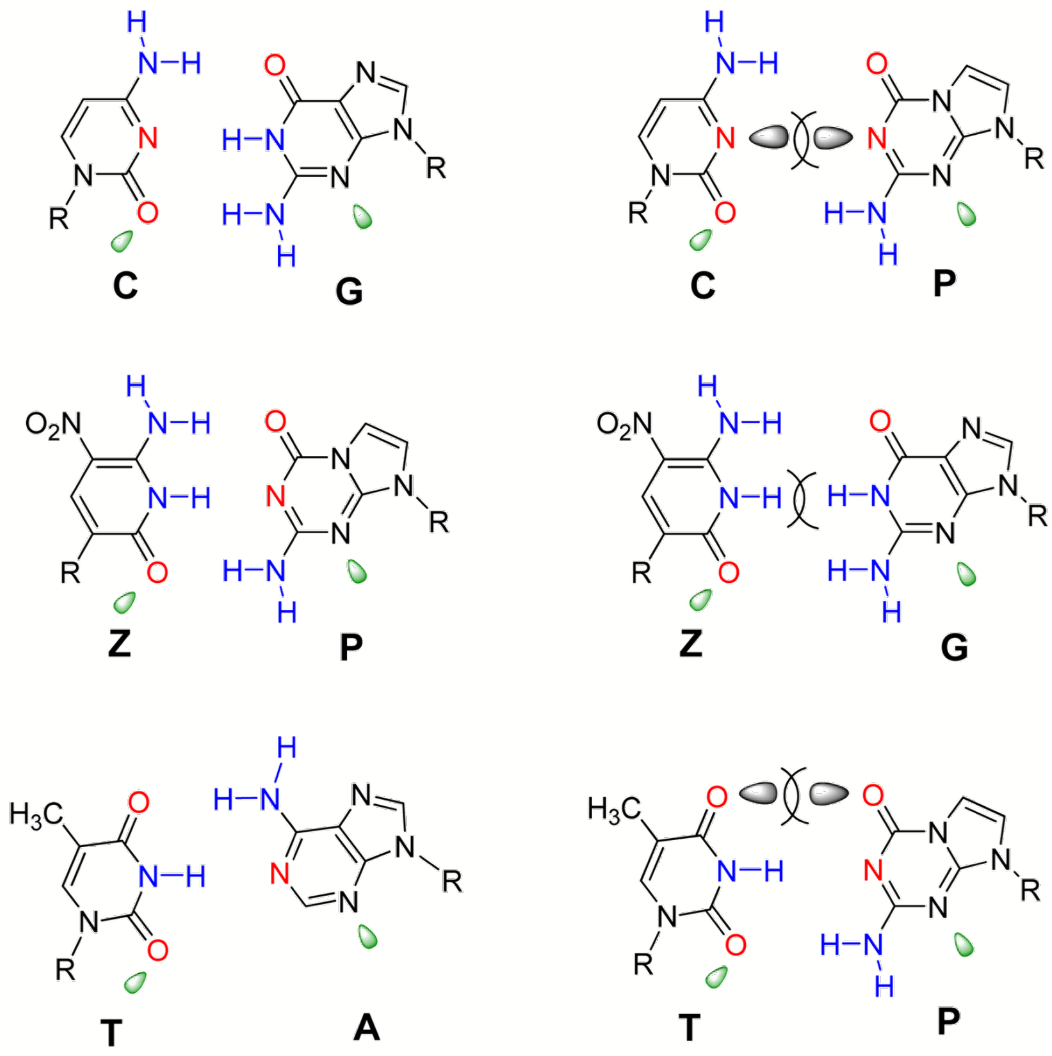
Alternative Sugars





Sugar subsitution:
1. Locked NA
2. Glycerol NA
3. Threose NA
4. Hexose NA

Expanding the genetic code

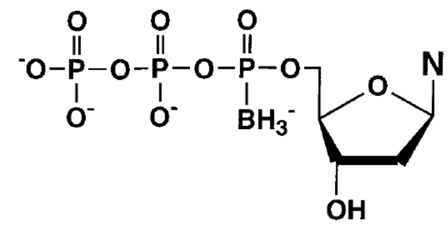


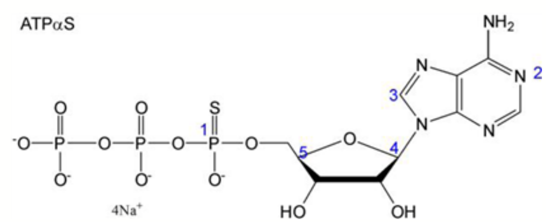
Phosphate modification

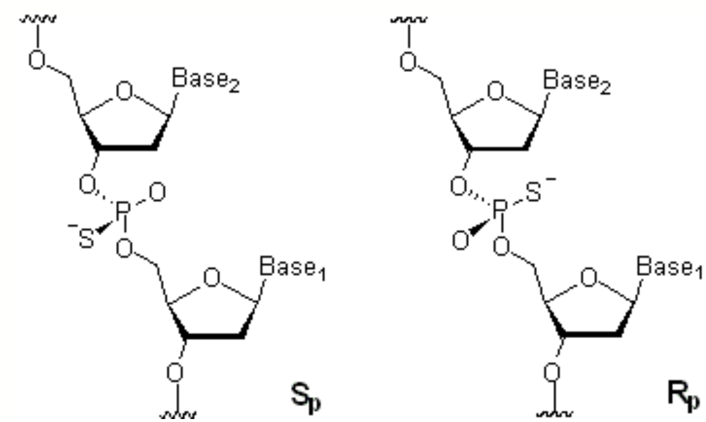
Oxygen-substitution generates chirality

- Boron
- Sulfur
- Selenium
- Hydrogen

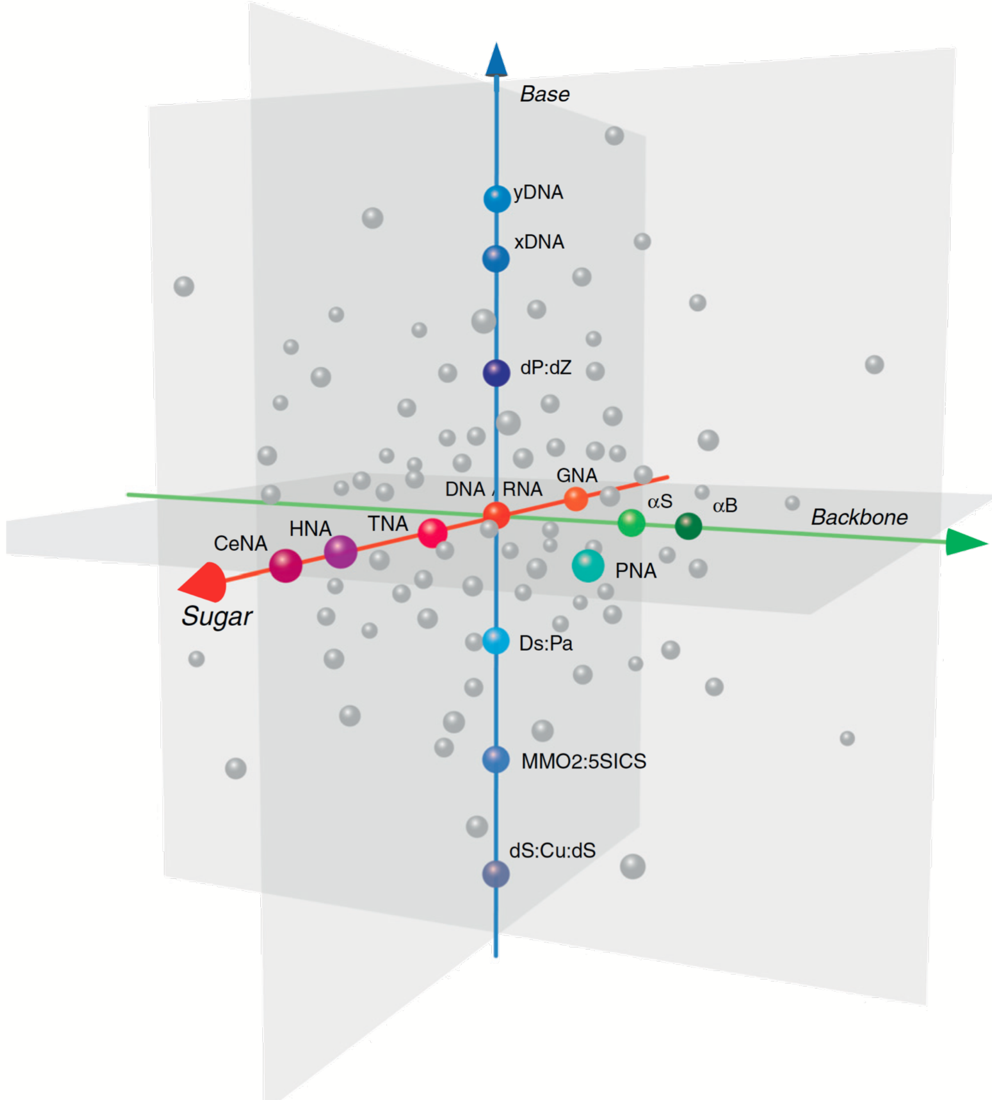
Improves stability *in vivo*







Possible modifications



How did life choose?

Prebiotic selection:

Availability

Stability of monomers

Stability of polymers

Biotic selection:

Cost of biosynthesis

Increased functionality

Would it choose these twice?