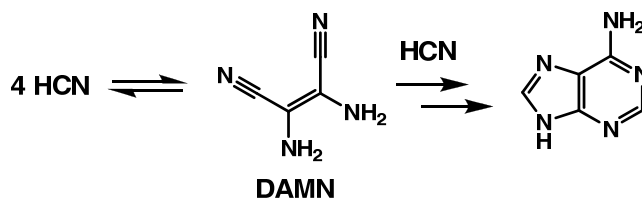
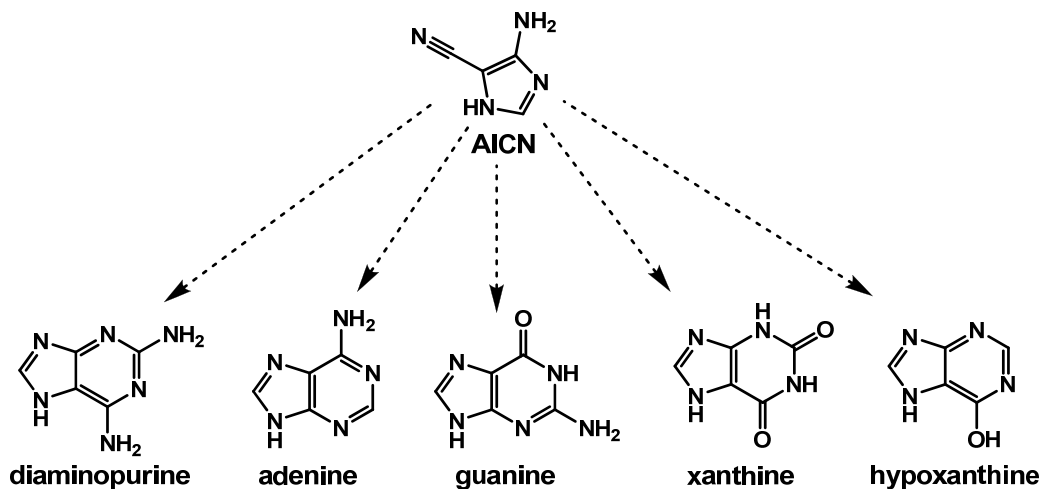


H, C, O and N are the first, third, fourth, and seventh most abundant elements in the universe. The stable molecules that are composed of these elements include nitriles, water, aldehydes, amines, carbonates, and hydrocarbons. Biomolecules can be interpreted as polymers of these building blocks. Indeed, we have seen the prebiotic synthesis of peptides and sugars that are explicitly constructed from them. Nucleobases may also be polymers of these simple compounds. For example, adenine has the molecular formula $C_5H_5N_5$, which is suggestive of a 5 HCN origin. Similarly guanine is $C_5H_5N_5O$, suggestive of a 5 HCN origin with a C-OH bond (from water) in place of a C-H bond.

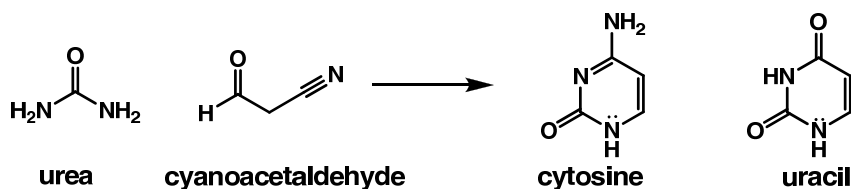
The pre-cellular synthesis of purines is thought to have progressed through a DAMN intermediate by a mechanism that will be covered in lecture. However, the aromaticity of nucleobases provides a thermodynamic driving force for their formation, and multiple reaction pathways produce them. Purine nucleobases have been found in both Miller-Urey experiments and in carbonaceous meteorites.



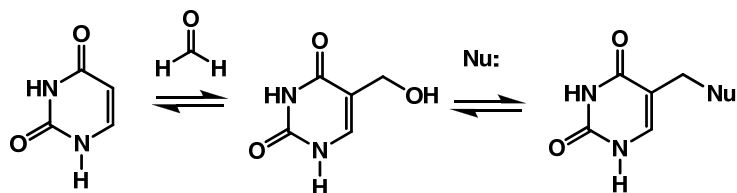
At least 5 different purine nucleobases are produced by cyanide polymerization. Their last common chemical precursor is likely **4-amino-imidazole-5-carbonitrile (AICN)**. The mechanism by which only guanine and adenine became incorporated into nucleic acids is unknown.



Interestingly, pyrimidines have not been detected in HCN polymerization reactions. However, inspection of the products from the Miller-Urey reaction revealed the formation of cyanoacetaldehyde and urea. Urea is a hydrolysis product of cyanamide (H_2NCN). Cyanoacetaldehyde is the coupling product of glycolaldehyde (a two carbon sugar from the formose reaction) and cyanide. Reaction of these two molecules, in concentrations higher than those produced in Miller-Urey, produces **cytosine** and **uracil**. The mechanism of this reaction will be covered in lecture.

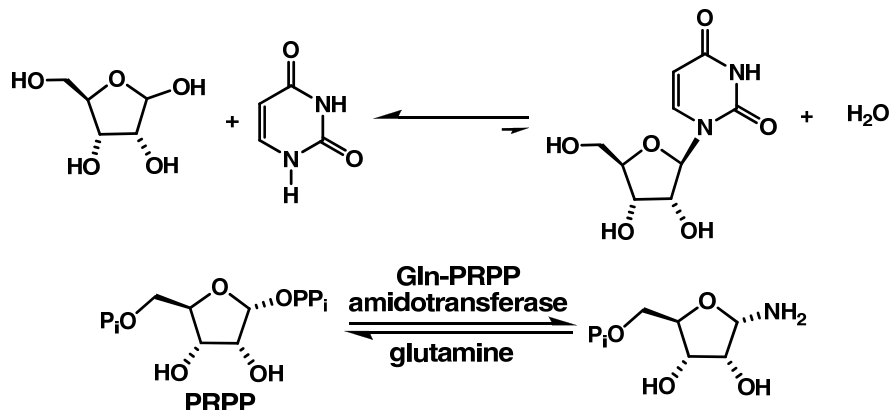


An interesting reaction of pyrimidines with formaldehyde may have helped to increase the catalytic potential of an RNA World before proteins arose. Uracil contains a nucleophilic **enamine** that can react with formaldehyde. This produces a serine-like side chain on the non-Watson-Crick side of the nucleobase; thus base pairing would still have been possible. Available nucleophiles may have displaced this hydroxyl to produce a variety of catalytically-active residues reminiscent of modern proteins. Heating formic acid (formed from cyanide hydrolysis) causes it to eliminate a hydride and produce carbon dioxide. This hydride can add to the Michael acceptor to produce thymidine, the final nucleobase required for the synthesis of DNA.



The formation of a covalent bond between a nucleobase and ribose to form a nucleoside is not straightforward, and there are a number of unsolved problems in describing its pre-cellular synthesis. The coupling reaction requires a thermodynamically disfavored condensation that employs a poorly nucleophilic aromatic nitrogen. Modern

cells generate the ribose-nucleobase bond by first pyrophosphorylating a 1'-ribose hydroxyl with ATP to generate 5-phosphoribosyl-1-pyrophosphate (PRPP). The pyrophosphate is then displaced by a nitrogen from glutamine. The nucleobases are subsequently built upon this 1-amino-ribose, thus avoiding direct coupling of a nucleobase to ribose.



Pre-cellular syntheses have been hypothesized that use similar strategies. For example, ribose reacts with cyanamide to form a bicyclic intermediate. Subsequent reaction with cyanoacetylene (HCCCN) followed by hydrolysis generates the cytidine, through a mechanism we will discuss in lecture. This synthesis solves regiochemical and kinetic problems that arise from a direct ribose-nucleobases coupling. For example, a furanosyl ribose ring, as seen in cellular nucleic acids, results from the reaction of ribose with cyanamide through a ring closing mechanism; the reaction of nucleobases with free ribose generates almost exclusively pyranosyl nucleosides. Additionally the reaction is facile in neutral aqueous solution, while direct nucleobases-ribose coupling is low-yielding even when the surrounding aqueous solvent is boiled away.

