## The molecular origins of life



## L5 SoSe 2020 Zbigniew Pianowski

#### Cyanosulfidic chemistry



#### **Prebiotic route to pyrimidine nucleotides**





Prof. John Sutherland Laboratory of Chemical Biology Cambridge, UK



Dr. Matthew Powner University College London, UK

# Prebiotic synthesis of activated pyrimidine nucleotides

Catalysis, and reaction control through pH and chemical buffering, is indicated by dashed lines.



M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239–242

#### Cyanosulfidic chemistry

Rainfall on higher ground (left) leads to rivulets or streams that flow downhill, sequentially leaching feedstocks from the thermally metamorphosed evaporite layer. Solar irradiation drives photoredox chemistry in the streams. Convergent synthesis can result when streams with different reaction histories merge (right), as illustrated here for the potential synthesis of arabinose aminooxazoline (5) at the confluence of two streams that contained glycolaldehyde (1), and leached different feedstocks before merging.









Table 1 | Yields for the part of the reaction network shown in Fig. 1b.

Conversion	Number of steps	Yield (%)	Conversion	Number of steps	Yield (%)
<b>4</b> → <b>17</b>	1	59	<b>26</b> → <b>28</b>	1	57
<b>17</b> → <b>18</b> +	1	29	<b>28 → 29</b>	1	75
19		34			
18 → 24	2	62	<b>26</b> → <b>29</b>	2	43
<b>24</b> → <b>25</b>	1	41	<b>29</b> → <b>30</b>	2	66
<b>25</b> → <b>26</b>	2	78	<b>30</b> → <b>31</b>	1	42
<b>26</b> → <b>27</b>	1	42	$19 \rightarrow 21 +$	2	31
			22		40

Table 2 | Yields for the parts of the reaction network shown in Fig. 1c,d.

Conversion	Number of steps	Yield (%)	Conversion	Number of steps	Yield (%)
33 → 34	1	83	$38 \rightarrow 41+$	1	30
			42		60
<b>34</b> → <b>35</b>	1	55	38 → 44	2	70
<b>3</b> 4 → 37	2	77	<b>44</b> → <b>47</b>	2	32
<b>3</b> 4 → 36	1	45	45 → 46	1	90
37 → 39	1	77	$6 \rightarrow 48 +$	1	50
			49+		25
			50		16
<b>37</b> → 40	2	~100	48 → 51	1	90
<b>37</b> → <b>43</b>	3	~70	51 → 52	1	89
<b>37</b> → <b>45</b>	5	~50	52 → 53	1	~100
36 → 38	1	~100	52 → 54	2	~70

#### Cyanosulfidic chemistry system



#### Cyanosulfidic chemistry system



#### Enantiomeric excess in the cyanosulfidic chemistry



**a**, In the presence of an enantioenriched L-proline ( $\underline{30}$ ), the diastereoselective formation of a three-component side product ( $\underline{6}$ ) effectively sequesters the unnatural L-glyceraldehyde ( $\underline{L-1}$ ).

**b**, The side reaction acts as a kinetic resolution of glyceraldehyde, giving enantiorichment of greater than 90% e.e.  $\underline{D-1}$ , which reacts with  $\underline{2}$  to form the enantioenriched amino-oxazoline RNA precursors  $\underline{D-4}$  and  $\underline{D-5}$ . e.e. values are  $\pm 2\%$ .

J. E. Hein, E. Tse, D. G. Blackmond, Nature Chem., 2011, 3, 704-706

#### **Enantiomeric excess in the cyanosulfidic chemistry**

Table 1 | Formation of enantioenriched amino-oxazolines in the presence of L-amino acids.

Amino acid	Three-component product* 6	Ribose amino- oxazoline D-4	Arabinose amino- oxazoline D-5
$A _{2}$ (32)		80	81
Ara ( <b>3b</b> )		11	72
Arg (30)	++	11	7.5
Asir (SC)	+	21	0.5
Asp (30)	+	2.1	1.4
Cys ( <b>3e</b> )	+++	n.a.	1.4
Gln ( <b>3f</b> )	+	1.2	1.1
Glu ( <b>3</b> g)	+	0.8	0.1
Gly (3h)	++	-	-
His (3i)	++	7.5 (L)	8.1 (L)
lle ( <b>3j</b> )	+	2.1	0.5 (L)
Leu (3k)	+	1.1	2.1
Lys ( <b>3</b> I)	+++	n.a.	n.a.
Met (3m)	+++	n.a.	n.a.
Phe ( <b>3n</b> )	+++	2.5	5.4
Pro ( <b>3o</b> )	++	55	58
Ser ( <b>3p</b> )	+++	3.0	1.9
Thr ( <b>3q</b> )	++	1.1	2.6
Trp ( <b>3r</b> )	++	10.2	9.8
Tyr (3s)	+	0.5	2.6
Val (3t)	++	2.0	1.0 (L)

\*Yield of side product 6: +, low; ++, medium; +++, high. n.a., no products isolated or observed by chiral LC



1% e.e. L-proline (30) is suspended in solvent (either  $CHCl_3$  or EtOH). After equilibration, the remaining solid is removed and the solvent is evaporated from the supernatant. Racemic glyceraldehyde DL-1 and amino-oxazole **2b** are then added and the mixture is dissolved in water. The ensuing reaction produces amino-oxazolines **4** and **5** in 20–80% e.e. Cooling the mixture to 4 °C induces crystallization of enantiopure ribo-amino-oxazoline crystals.

J. E. Hein, E. Tse, D. G. Blackmond, Nature Chem., 2011, 3, 704-706

#### Chiral sugars drive enantioenrichment in prebiotic aminoacid synthesis



D. G. Blackmond et al., ACS Cent. Sci., 2017, 3, 322-328



#### Strategies toward Enantio-enriched Glyceraldehyde and Ribonucleotide Precursors

### **Overcome of the Formation of Prebiotic Clutter.** 29 + 30HC 20 14 HO $H_2N \longrightarrow N$ HO 33 31 22 22

The synthesis of activated pyrimidine ribonucleotides **29** and **30** is dependent on the controlled formation of pentose aminooxazolines **31** (black), but the synthesis of **31** is wholly reliant on the ordered introduction of pure glycolaldehyde **14** (to cyanamide 33) and glyceraldehyde **20** (to 2-aminooxazole **32**) to prevent the formation of numerous deleterious byproducts (red). Ribonucleotide synthesis fails without the adherence to this order of synthetic steps. Glyceraldehyde **20** is highly susceptible to equilibration with dihydroxyacetone **22**, especially in phosphate buffer, which results in diminishing amounts of pentose aminooxazolines **31** being formed (inset).

S. Islam, M. W. Powner Chem 2017, 2, 470-501

#### 2-Aminothiazole-Controlled Aldehyde Sequestration



S. Islam, M. W. Powner Chem 2017, 2, 470-501

#### 2-Aminothiazole-Controlled Aldehyde Sequestration



S. Islam, M. W. Powner Chem 2017, 2, 470-501

#### Systems Chemical Analysis of Amino Acid and Nucleotide Syntheses



Analysis of the prebiotic amino acid and nucleotide syntheses reveal that glycolaldehyde **14**—a serine and ribonucleotide precursor—lies at a generational node between these two metabolite classes. The same analysis applied to cysteine suggested that b-mercaptoacetaldehyde **47** would be as important as glycolaldehyde **14** and that the reactivity of 2-aminothiazole **44** might have key implications for the concomitant prebiotic synthesis of amino acid and nucleotides

S. Islam, M. W. Powner Chem 2017, 2, 470-501







S. Stairs, A. Nikmal, D-K. Bucar, S-L. Zheng, J. W. Szostak, M. W. Powner Nature Commun. 8:15270





S. Stairs, A. Nikmal, D-K. Bucar, S-L. Zheng, J. W. Szostak, M. W. Powner Nature Commun. 8:15270



S. Stairs, A. Nikmal, D-K. Bucar, S-L. Zheng, J. W. Szostak, M. W. Powner Nature Commun. 8:15270







#### Canonical purine nucleoside synthesis via cyanosulfidic chemistry

#### Canonical purine nucleoside synthesis via cyanosulfidic chemistry

beta-Ribofuranosyl-pyrimidine nucleotide assembly and potential stepwise, regioselective beta-ribofuranosyl-purine assembly Pathway via the intermediacy of tetrahydroimidazo[1',3']-2"-aminooxazolo[1',2']-pyrimidinesa



M. W. Powner, J. D. Sutherland, J. W. Szostak J. Am. Chem. Soc. 2010, 132, 16677-16688

#### Prebiotic synthesis of deoxyribonucleosides





proposed multicomponent deoxyribonucleotide syntheses



M. W. Powner, S.-L. Zheng, J. W. Szostak J. Am. Chem. Soc. 2012, 134, 13889-13895

#### Purine nucleoside synthesis - alternatives



#### **Prebiotic synthesis of purine nucleosides –FaPY pathway**



T. Carell, Nature 2016, 352(6287), 833-836

#### Prebiotic syntheses of aminopyrimidines



T. Carell, Nature 2016, 352(6287), 833-836

#### **Prebiotic synthesis of purine nucleosides –FaPY pathway**



#### **Prebiotic synthesis of purine nucleosides –FaPY pathway**



T. Carell, Nature 2016, 352(6287), 833-836



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82



#### S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82
#### Unified prebiotic synthesis of pyrimidine and purine ribonucleotides



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82

### Unified prebiotic synthesis of pyrimidine and purine ribonucleotides



S. Becker, J. Feldmann, S. Wiedemann, ..., T. Carell, Science 2019, 366, 76-82



#### Prebiotic phosphorylations and the origins of protometabolism

### Selective Phosphorylation of Glycolaldehyde and Aldol Reactions of Glycolaldehyde Phosphate



S. Islam, M. W. Powner Chem 2017, 2, 470-501



Prebiotic Reconstruction of the Triose Glycolysis Pathway by Selective a-Phosphorylation of Sugars

S. Islam, M. W. Powner Chem 2017, 2, 470-501



н

H

Purines

### **Prebiotic soup - summary**

Fischer-Tropsch chemistry - lipids



Strecker chemistry - aminoacids



Dehydrating agents (COS, NO) – condensation of AAs to peptides





Phosphorus reactivity - phosphates



The Traditional Modular Retrosynthetic Analyses Disconnect RNA to Ribofuranosyl Sugar, Inorganic Phosphate, and Canonical RNA Nucleobases.



S. Islam, M. W. Powner Chem 2017, 2, 470-501

The Three Pillars of Prebiotic Chemistry

(A) The spark discharge aminonitrile synthesis (The Miller-Urey experiment.),

(B) Nucleobase synthesis by HCN oligomerization, and

(C) Sugar synthesis by the formose reaction.



S. Islam, M. W. Powner Chem 2017, 2, 470-501



S. Islam, M. W. Powner Chem 2017, 2, 470-501

*Summary of the Prebiotic Syntheses of the Activated Pyrimidine Ribonucleotides* 



S. Islam, M. W. Powner Chem 2017, 2, 470-501

Conversion of Ribose Aminooxazoline to Activated Pyrimidine Ribonucleotides



S. Islam, M. W. Powner Chem 2017, 2, 470-501

Simultaneous pH-Controlled Multicomponent Assembly of Purine and Pyrimidine Nucleotide Precursors



HCN tetramers AICA **40** and AICN **41** participate in a high-yielding pH-dependent three-component reaction with glyceraldehyde **20** and 2-aminooxazole **32**. This produces potential purine ribonucleotide precursors **39**. The Mannich-type reactivity results in N9-purination with absolute regiospecificity. At pH 6–6.5, both purine **39** and pyrimidine **31** ribonucleotide precursors are observed, suggesting that a divergent synthesis of purine and pyrimidine ribonucleotides from within one pool of reagents is an enticing prospect.

S. Islam, M. W. Powner Chem 2017, 2, 470-501

Origin of the Universe – stars, planets, elements

Origin of biorelevant monomers – primordial soup

Complex chemical processes on the way to living systems

**Protocells and LUCA** 

## **Condensation of aminoacids into peptides**



## Biochemical condensation of aminoacids into peptides



Nature Publishing Group, www.nature.com/nrg/multimedia



## Spontaneous vs. assisted dehydratation



Rode, B. M.; Fitz, D.; Jakschitz, T. Chem. *Biodiversity* **2007**, *4*,2674.

Activating agent	Hydrolysis/ hydration product	$\Delta G^{o\prime}/{ m kJ\ mol^{-1}}$
NH <sub>2</sub> CONH <sub>2</sub>	$CO_2 + NH_3$	$-16^{a}$
COS (g)	$CO_2 + H_2S$	$-17^{a}$
Pyrophosphate	Phosphate	$-19^{b}$
CO (g)	HCO <sub>2</sub> H	$-16^{a}$
HNCO	$CO_2 + NH_3$	$-54^{a}$
HCN	$HCO_2H + NH_3$	$-75^{a}$
RCN	$RCO_2H + NH_3$	$-80^{c}$
NH <sub>2</sub> CN	Isourea	$-83^{d}$
HNCNH	Isourea	$-97^{d}$
HCCH (g)	CH <sub>3</sub> CHO	$-112^{a}$

Danger, G.; Plasson, R.; Pascal, R. *Chem. Soc. Rev.* **2012**, *41*, 5416.

## **Condensation of aminoacids into peptides**





## Amyloid world



#### Schematic of amyloid formation.

Native proteins are in dynamic equilibrium with their less-structured, partially folded and/or unfolded states. One of these states initiates amyloid fibril formation by assembling into oligomeric species. Oligomeric species can then assemble further to form higher-order oligomers, one or more of which can form a fibril nucleus, which, by rapidly recruiting other monomers, can nucleate assembly into amyloid fibrils.

M. G. Iadanza et al. Nature Rev. Mol. Cell Biol. 2018, 19, 755-773



Typically, the repeating unit of the amyloid fibrils consists of two tightly packed layers of  $\beta$ -sheets with side chains within the bilayers forming a dry interdigitating zipper interface.

C. P. J. Maury Cellular and Molecular Life Sciences 2018 75, 1499–1507



Electron micrograph of a polymorphic fibrillar amyloid network self-assembled from a prebiotically relevant 9-mer peptide (EGGSVVAAD) in aqueous environment.

> Maury CPJ, Liljeström M, Zhao F, *J Biol Res* **2012**, *18*, 332–335



# Amyloid world

An initial **slow nucleation** process is followed by a fast polymerization phase where peptide monomers are added to the growing end of the protofilament. Fragmentation generates new seeds that can initiate repeated replication cycles. The same peptide monomer can give rise to different amyloid structures and molecular rearrangements are possible. Specific conformational changes can be replicated in the fibril/protofibril-catalyzed cycle II. Amyloid is also able to direct the synthesis of its own constituent peptides. The β-sheet conformers and ribonucleotides interact dynamically and cooperatively, and the amyloidbased supramolecular fibrillar assemblies can function as a primitive metabolic apparatus catalyzing the formation metabolite precursors.

> C. P. J. Maury Cellular and Molecular Life Sciences **2018** 75, 1499–1507

# Amyloid world

### Schematic representation of the self-replicating cycles of amyloid.



An initial **slow nucleation phase (a)** is followed by a kinetically **fast elongation phase (b)** where monomers (or oligomers) are added sequentially to the growing end of the protofiber. **Breakage** of the fiber results in **new seeds (c)** and repeated replication cycles. Importantly, molecular **rearrangements** and conformational changes in amyloid may occur **(d)** that, by a **templated** *conformational* replication mechanism **(e)**, can faithfully be transmitted to other amyloid conformers. The pool of the environmentally fittest variant(s) then **expands (f)**.

C. P. J. Maury J. Theor. Biology 2015 382, 292-297

## The amyloid world model



From one type of prebiotic peptide monomer a spectrum of amyloid conformers may be formed  $(T_1, T_2, T_3)$ . By templated conformational replication, the pool of the environmentally fittest type  $(T_2,$ Selection 1) rapidly expands. A change in the environment (e.g., pH, temperature, radiation) induces conformational changes in  $T_2$  ( $T_{2a}$ ,  $T_{2b}$ ,  $T_{2c}$ ). The fittest conformer  $(T_{2h})$  is selected (Selection 2) and undergoes templated conformational replication cycles expanding the  $(T_{2h})$  pool. The environmentally less suitable variants are decomposed and recycled. The environment-induced variations in the amyloid conformations combined with faithful replication of the selected amyloid conformers (variants) and repeated selection cycles allow evolution to occur.

C. P. J. Maury J. Theor. Biology 2015 382, 292-297

## **Peptide self-replication**





### **Peptide self-replication**





K. Severin, D. H. Lee, A. J. Kennan and M. Reza Ghadiri Nature 1997, 389, 706-709

## **Peptide self-replication**



K. Severin, D. H. Lee, A. J. Kennan and M. Reza Ghadiri Nature **1997**, 389, 706-709

### Nucleotide polymerization

*Regioselective formation of 3'-5' phosphodiester bonds between nucleotides* 



## *Nucleotide polymerization – sequence control*





dxline.info/img/new\_ail/dna-polymerase\_1.jpg







www.neb.com

## **Products of chemical condensation of nucleotides**



(A) Reaction of an activated mononucleotide (N<sub>i+1</sub>) with an oligonucleotide (N<sub>1</sub>–N<sub>i</sub>) to form a 3',5'-phosphodiester (left), 2',5'-phosphodiester (middle), or 5',5'-pyrophosphate linkage (right).

(B) Typical oligomeric product resulting from chemical condensation of activated mononucleotides

# Degradation of activated nucleotides



hydrolysis

3',5'-cyclization

### Template-directed synthesis



#### Montmorillonite





B = adenine, guanine, cytosine or uracil



(Top) Structure of ribonucleotide 5'-phosphoimidazolides (left) and ribonucleotide 5'-phosphoro-1-methyladeninium (right). (Bottom) Unit cell of montmorillonite and phosphodiester bond formation within the clay interlayers, as proposed by Ferris and coworkers (right). XH, depicted in blue in the cartoon, is any undifferentiated protic species inside the clay galleries. Joshi, P. C.; Aldersley, M. F.; Delano, J. W.; Ferris, J. P. J. Am. Chem. Soc. **2009**, *131*, 13369

### Intercalating agents





Rate increase by three orders of magnitude vs. ligation without proflavine

N. V. Hud et al. Angew. Chem. Int. Ed. 2004, 43, 2004 – 2008

### **Template-directed synthesis**

no example demonstrated yet, where single activated nucleotides would form a complementary strand on an RNA (or DNA) template without enzymatic support



Current experiments focus on ,primer extention' or ,filling abasic sites'— sequence-selective complementary nucleobase addition to a pre-existing strand (or between two pre-existing strands) already hybridized on a template. Here, pre-organization provided by the existing base-pairing network supports selection of the correct nucleoside to be joined.

Complementary approaches are regioselective ligation reactions of short oligonucleotides on templates, or dynamic covalent chemistry, where nucleobase-containing components would be added sequence-specifically to a pre-existing ,empty' backbone on a template
#### First non-enzymatic self-replicating system CH<sub>3</sub>O CH,O-Product inhibition – HO ĤΟ problematic for longer templates 1 HO-HO 3 EDC ĤO-Η̈́O HO ĤO-ĤΟ-2 ΗO• HO EDC: H<sub>3</sub>C.N CH<sub>3</sub> CH, N=C=N

V. Patzke, G. von Kiedrowski *ARKIVOC* **2007** 293-310 D. Sievers, G. von Kiedrowski *Nature* **1994** *369(6477)*, 221-224 G. von Kiedrowski *Angewandte Chemie* **1986** *98(10)*, 932-934

#### SPREAD – Surface-Promoted Replication and Exponential Amplification of DNA Analogues



V. Patzke, G. von Kiedrowski ARKIVOC 2007 293-310

#### Nonenzymatic primer extension



#### Ribozyme-catalyzed primer extension



Design of a 5'-nucleotidyl transferase for N>p's.

(a) Two-way junction HPz (small hairpin ribozyme), which catalyzes reversible RNA ligation using a 2',3'-cyclic phosphate.

(b) Redesign of HPz into 5NTz (nucleotidyl transferase).

(c) Structural model of the substrate-binding pocket of 5NTz (based on PDB1M5V).

(d) 5NTz catalyzes 5'-adenylation in ice (2 mM A>p, 2  $\mu$ M 5NTz, 1  $\mu$ M 3'-FITC-labeled AS, 72 h in ice at -7 °C).

An engineered hairpin ribozyme catalyzes addition of all four N>p's (2',3'-cyclic A-, G-, U-, and CMP) to the 5'hydroxyl termini of RNA strands (eutectic ice phase formation at -7 °C). 5' addition of 2',3'-cyclic phosphateactivated  $\beta$ -nicotinamide adenine dinucleotide (NAD>p), as well as ACA>p RNA trinucleotide, and multiple additions of GUCCA>p RNA pentamers was also observed.

H Mutschler, P. Holliger J. Am. Chem. Soc. 2014, 136, 5193-5196





Protection of the 2'-OH group of 1-3'P facilitates rapid template-directed 3',5'-ligation after electrophilic phosphate activation. The 3'-OH group of 1-2'P is protected to a lesser extent, such that 1>P is the major product of phosphate activation and slow template-directed 2',5'-ligation follows.

X = leaving group, Y = leaving group generated by electrophilic activation of phosphate oxygen with or without a subsequent nucleophilic displacement



Treatment of adenosine-3'phosphate (A3'P) (100 mM) with sodium thioacetate 3 (100 mM) and cyanoacetylene 4 (200 mM) in D<sub>2</sub>O at neutral pD for 24 hours results in selective acetylation of the 2-OH group.



Treatment of A3'P (80 mM) and A2'P (20 mM) as given before results in the exclusive 2-acetylation of the former nucleotide. Partial <sup>1</sup>H NMR spectrum of the reaction products.



Additional electrophiles 6–8 shown to drive the acetylation of ribonucleotides with thioacetate 3. Direct acetylation with 9 is also possible, as is oxidative activation of 3 with ferricyanide 10 to afford ferrocyanide 11 and a dimeric acetylating agent 12. Curly arrows indicate electrophilic activation/acetylation steps.

Chemoselective acetylation of 3'P-oligoribonucleotides expedites templated ligation



Sequences and reaction conditions employed for acetylation (i) and subsequent templated ligation (ii).

The acetylation mixture contained 80 mM primer and 50 mM NAI; the ligation mixture contained 4 mM primer from the acetylation reaction, 25 mM template, 30 mM ligator, 200 mM imidazole nitrate buffer (pH 6.2), 10 mM MnCl<sub>2</sub> and 100 mM NCl.

#### Dynamic oligonucleotide analogue sequence-specific assembly



M. R. Ghadiri et al. Science 2009, 325, 73-77

## Correction mechanism 2'-5' $\rightarrow$ 3',5'



# Correction mechanism 2'-5' $\rightarrow$ 3',5'



A. Mariani, J. D. Sutherland Angew. Chem. Int. Ed. 2017, 56, 6563-6566

#### Dynamic oligonucleotide analogue sequence-specific assembly



M. R. Ghadiri et al. Science 2009, 325, 73-77