A system-level viewpoint on the chemical origins of life

MARTHA GROVER

SCHOOL OF CHEMICAL & BIOMOLECULAR ENGINEERING

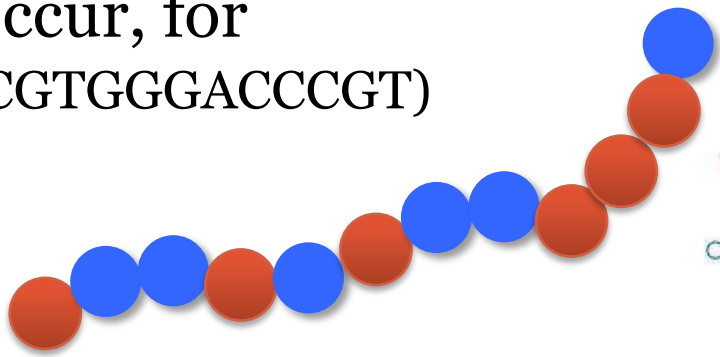
GEORGIA INSTITUTE OF TECHNOLOGY

SEPTEMBER 1, 2017

Goal

- Design and demonstrate prebiotically plausible minimal systems that can
  - Polymerize
  - Store information
  - Transfer information
  - Catalyze reactions (or other functions)
  - Undergo selection and evolution

- Ideally, we would like to find a common environment in which these steps would occur, for
  - Nucleic acid polymers (e.g. ACGTGGGACCCGT)
  - Peptides (e.g. KLVFFAE)
  - Other important biopolymers



CENTER FOR CHEMICAL EVOLUTION

What does engineering have to do with origins of life?

- Research goal: Design a polymer system that can evolve.
- Design: a methodical series of steps that engineers use in creating functional products and processes
  - Wikipedia, “engineering design”
- Design involves tradeoffs
- Design involves specifying an objective

Design principles

- Robust function
  - Modularity
  - Redundancy
  - Feedback
    - Positive: amplification and selection
    - Negative: regulation and robustness
- Complex function
  - Diversity (division of labor)
  - Cooperation

Emergence?

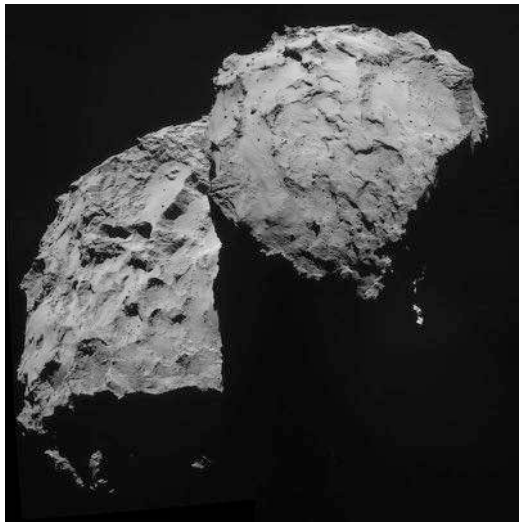
Optimization and Evolution

- Optimization
  - Maximizing an objective by adjusting design variables within constraints
- Evolution (Darwinian)
  - Change in the heritable characteristics of a population over successive generations
- Similarities
  - Genetic algorithms for engineering design
- Differences
  - Evolution: Selection by the environment
  - Optimization: Engineer specifies an objective function

The Organic Inventory

Rosetta Spacecraft - Philae Lander – Atmospheric MS Data

Name	Formula	Molar mass (u)	MS fraction	Relative to water
Water	H_2O	18	80.92	100
Methane	CH_4	16	0.70	0.5
Methanenitrile (hydrogen cyanide)	HCN	27	1.06	0.9
Carbon monoxide	CO	28	1.09	1.2
Methylamine	CH_3NH_2	31	1.19	0.6
Ethanenitrile (acetonitrile)	CH_3CN	41	0.55	0.3
Isocyanic acid	HNCO	43	0.47	0.3
Ethanal (acetaldehyde)	CH_3CHO	44	1.01	0.5
Methanamide (formamide)	$HCONH_2$	45	3.73	1.8
Ethylamine	$C_2H_5NH_2$	45	0.72	0.3
Isocyanomethane (methyl isocyanate)	CH_3NCO	57	3.13	1.3
Propanone (acetone)	CH_3COCH_3	58	1.02	0.3
Propanal (propionaldehyde)	C_2H_5CHO	58	0.44	0.1
Ethanamide (acetamide)	CH_3CONH_2	59	2.20	0.7
2-Hydroxyethanal (glycolaldehyde)	CH_2OHCHO	60	0.98	0.4
1,2-Ethanediol (ethylene glycol)	$CH_2(OH)CH_2(OH)$	62	0.79	0.2



Goesmann, F., Rosenbauer, H., Bredehöft, J. H., Cabane, M., Ehrenfreund, P., Gautier, T., ... & Ulamec, S. (**2015**). Organic compounds on comet 67P/Churyumov-Gerasimenko revealed by COSAC mass spectrometry. *Science*, *349* (6247).

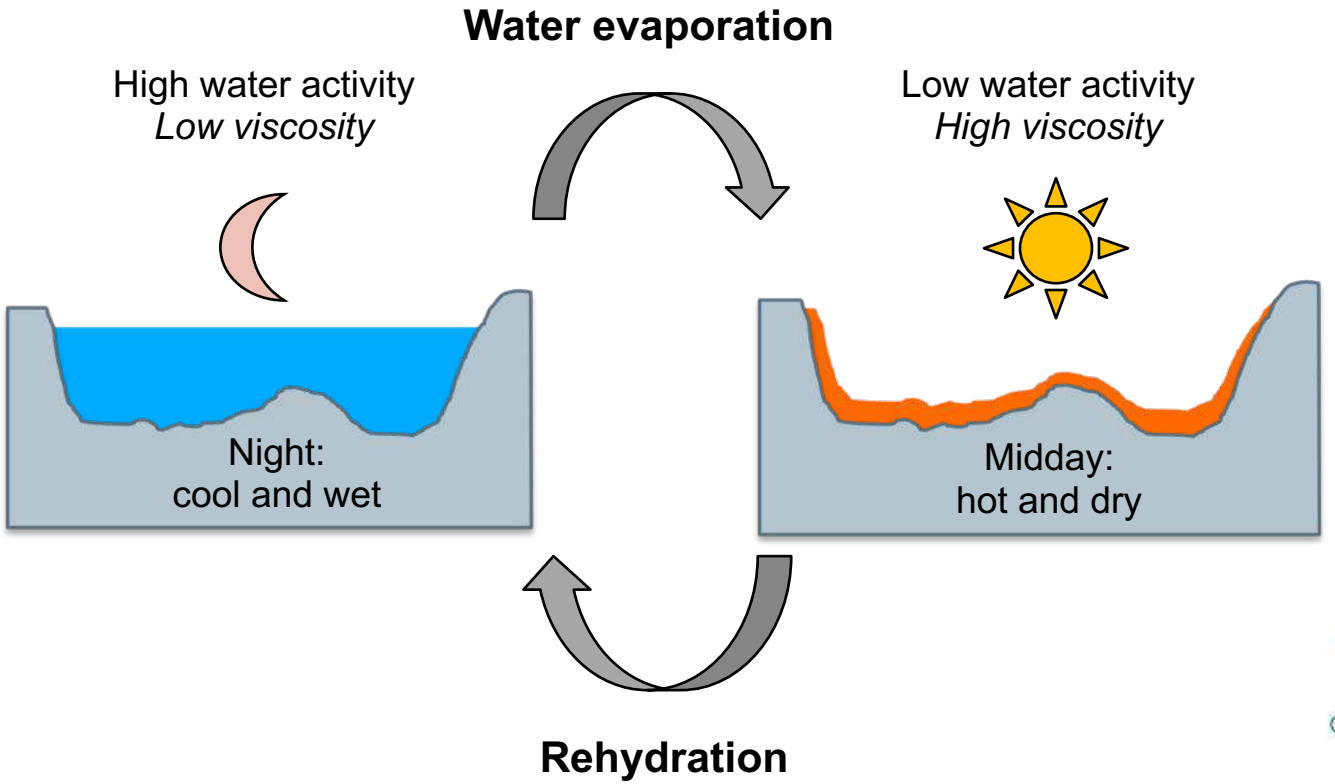
CENTER FOR CHEMICAL EVOLUTION

Need to address outstanding problems

- The water problem
  - Biopolymers are formed through condensation reactions, which are not favored in water.
- The strand inhibition problem
  - Heating drives duplex separation, but cooling promotes formation over replication.
- The single-winner scenario
  - Selection for the fastest replicator does not lead to a productive outcome.
    - Need to select for function.
    - Selection and evolution require sustained diversity.

Concept for the environment

Coupling between the chemistry and the environment determines system level behavior



Environmental cycles

- Prebiotically plausible (e.g. daily, tidal, seasonal)
- Types include hot/cold, wet/dry, pH swings
- Drives non-equilibrium behavior through the influx of solar energy
- Induces reversible phenomena over the course of the cycle
  - polymerization and hydrolysis
  - duplex formation and separation

CENTER FOR CHEMICAL EVOLUTION

Non-aqueous solvents

- Many organics present in the prebiotic inventory
- Solvents possibly created from non-volatile organics, after water evaporation
- Features
  - Drive condensation polymerization forward
  - Control differential mobility via viscosity
  - Promote intramolecular folding by suppressing intermolecular interactions

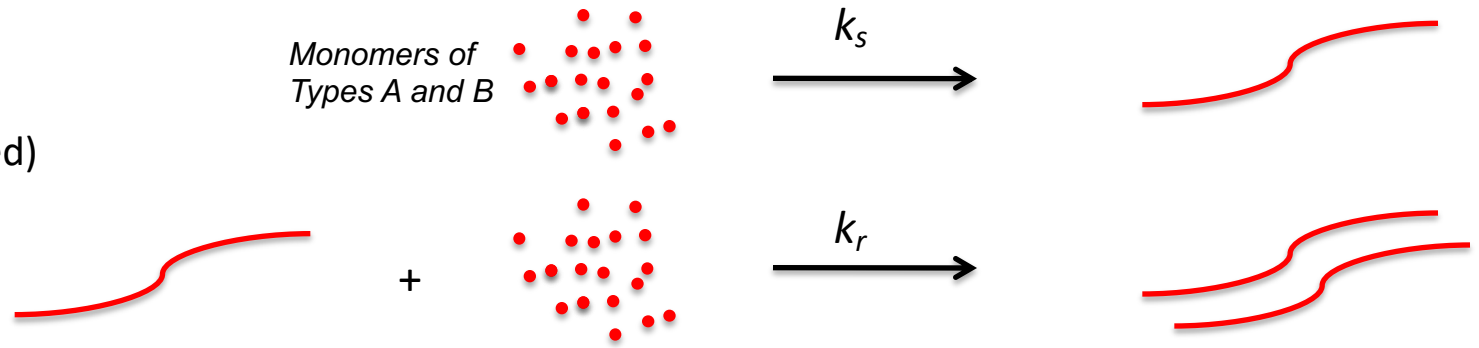
System-level behavior

- The interaction between the environment and the chemistry is critical to prediction of performance.
- Modeling approach
  - Use mathematical modeling to predict behavior and design promising experiments
  - It is difficult to predict the tradeoffs between simultaneously occurring phenomena without a mathematical model.
- A case study
  - The first functional biopolymer could have been the monomer synthetase.
  - Selection is not based on the inherent replication rate, but rather on the local resources in the environment.

A model of chemical evolution

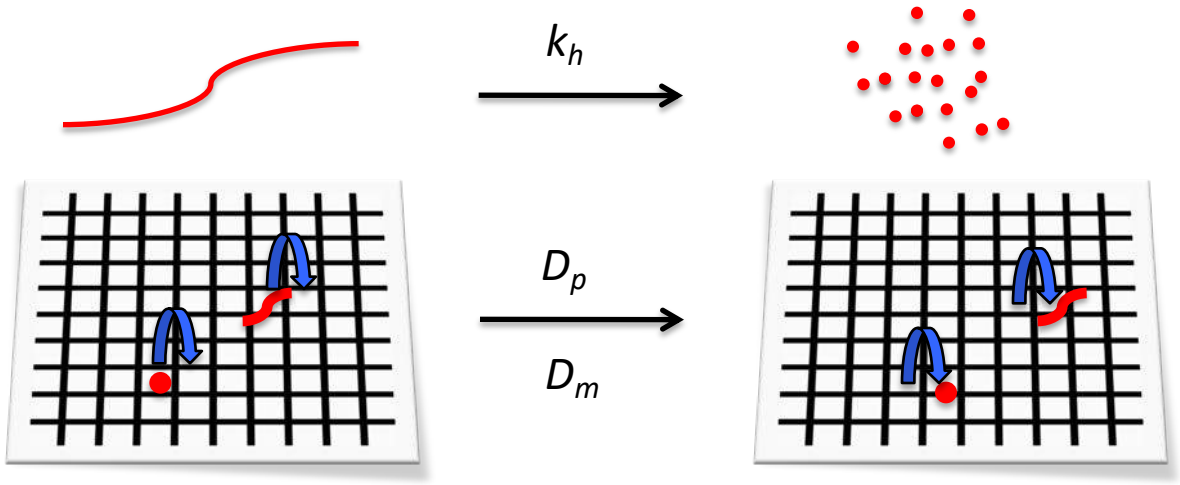
Five events and parameters define the system dynamics.

Day Phase
(dehydrated)



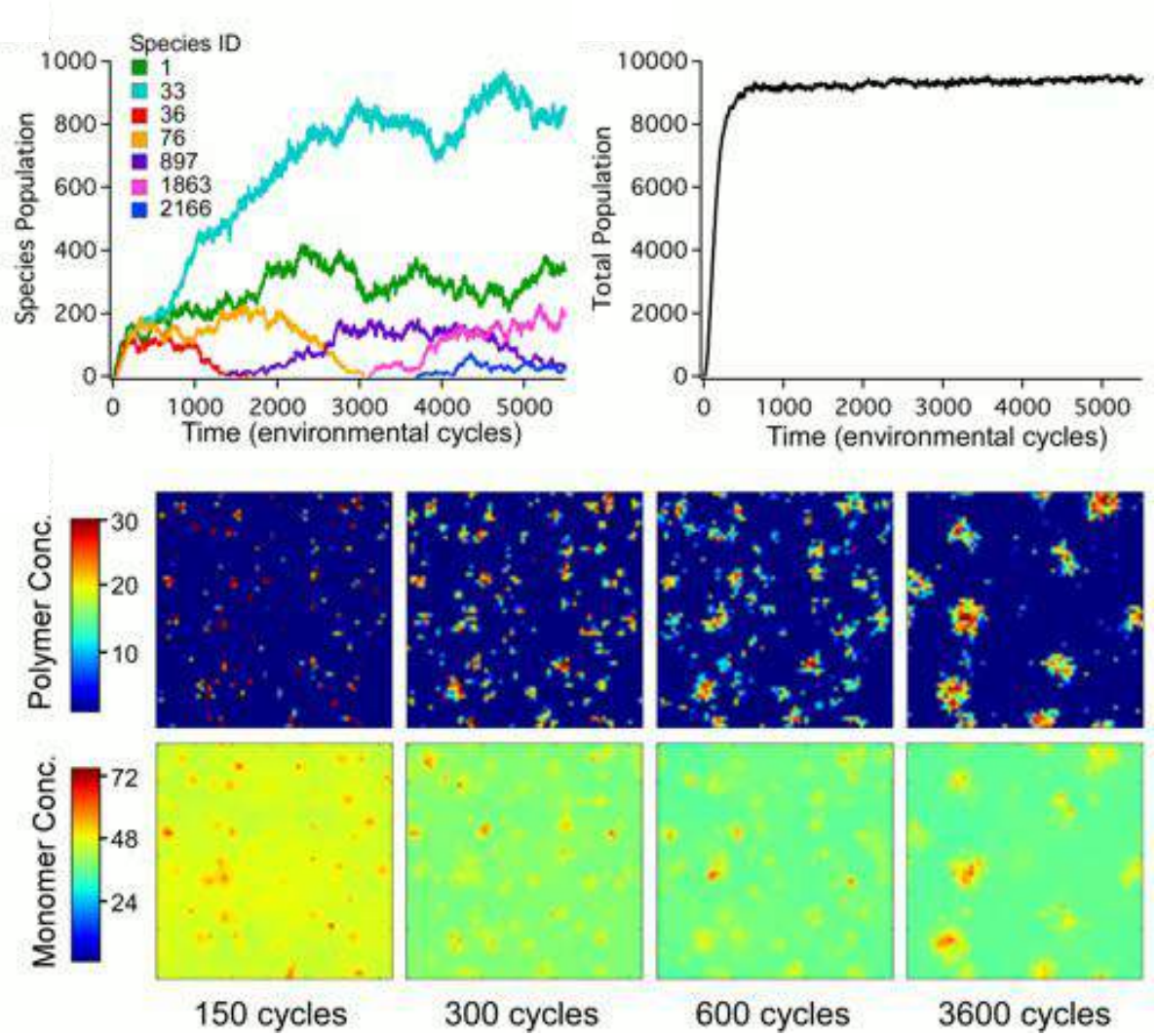
Rates of polymer formation are dependent on local resources.

Night Phase
(hydrated)



CENTER FOR CHEMICAL EVOLUTION

Sequence evolution of the pool



$k_s = 10^{-7}$

$k_r = 10^{-4}$

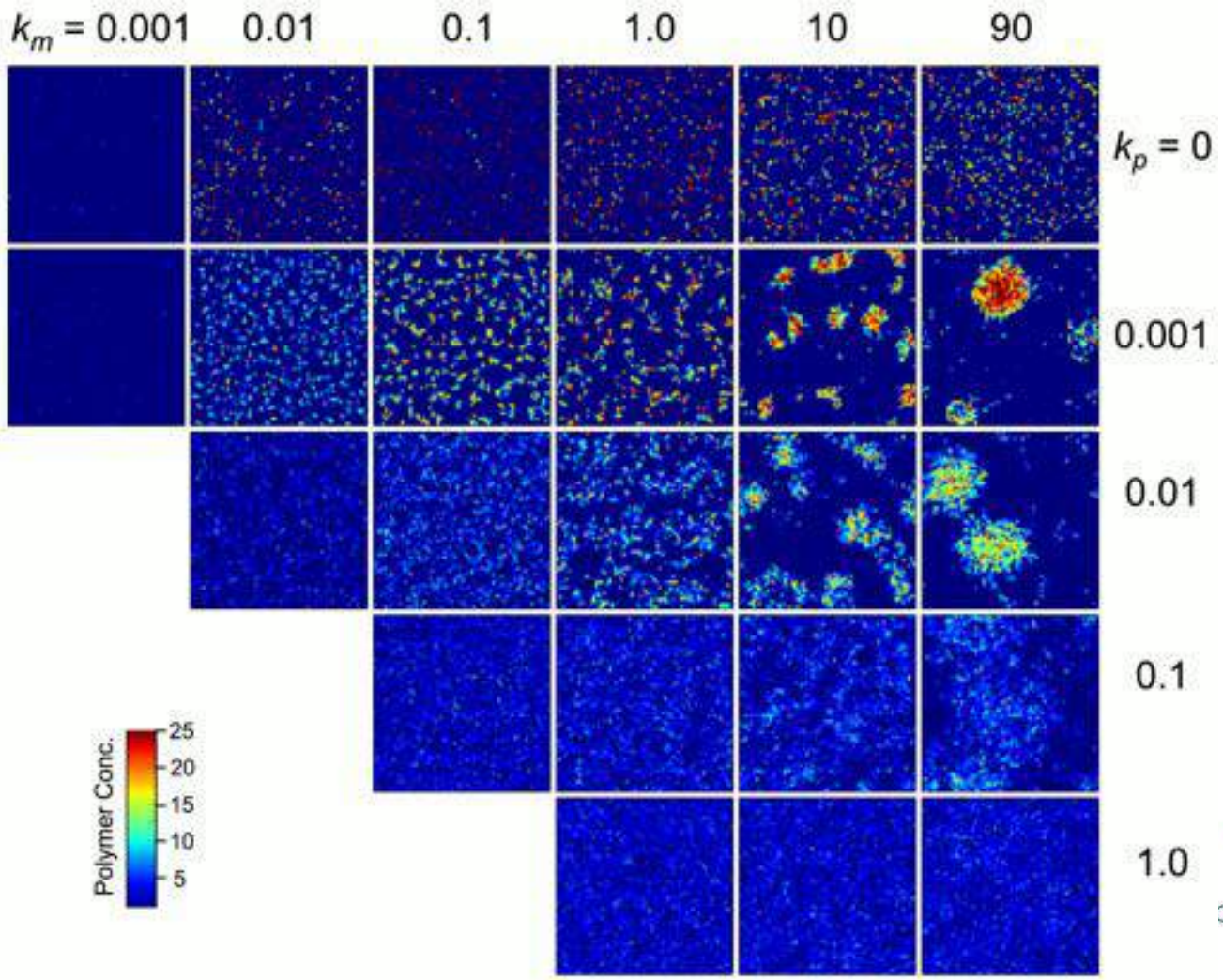
$k_h = 0.1$

$k_p = 0.001$

$k_m = 10.0$

CENTER FOR CHEMICAL EVOLUTION

Spatial patterning



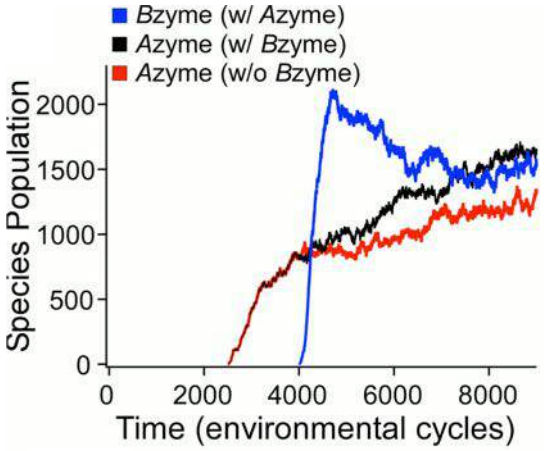
$k_s = 10^{-7}$

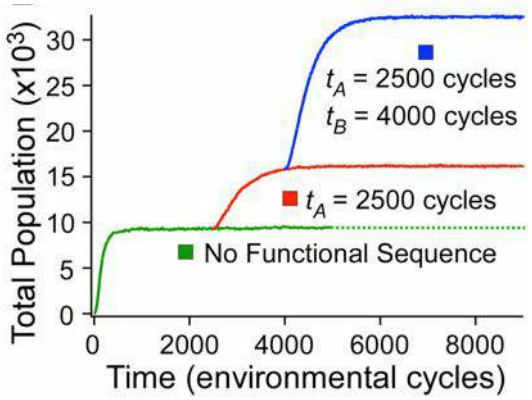
$k_r = 10^{-4}$

$k_h = 0.1$

CENTER FOR CHEMICAL EVOLUTION

Cooperativity over time and space





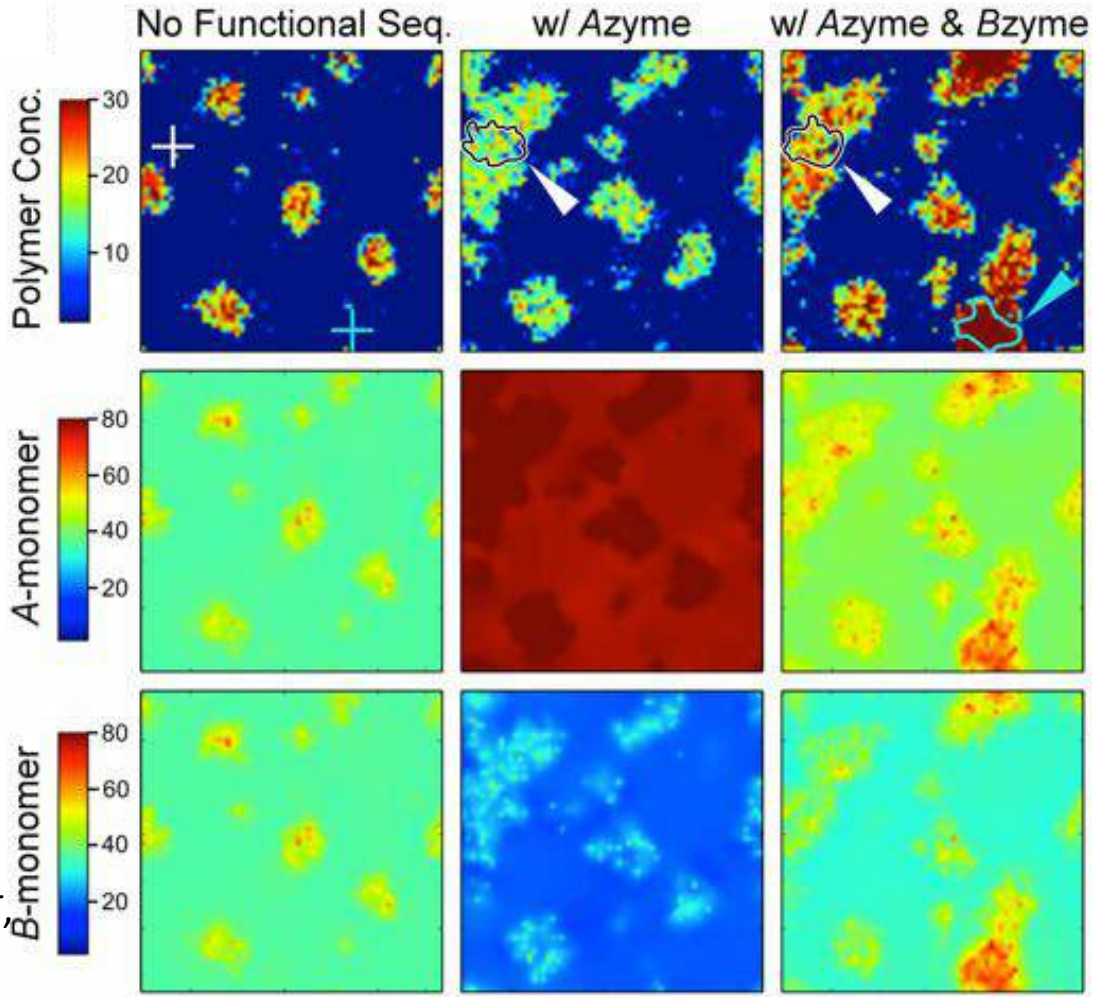
$k_s = 10^{-7}$

$k_r = 10^{-4}$

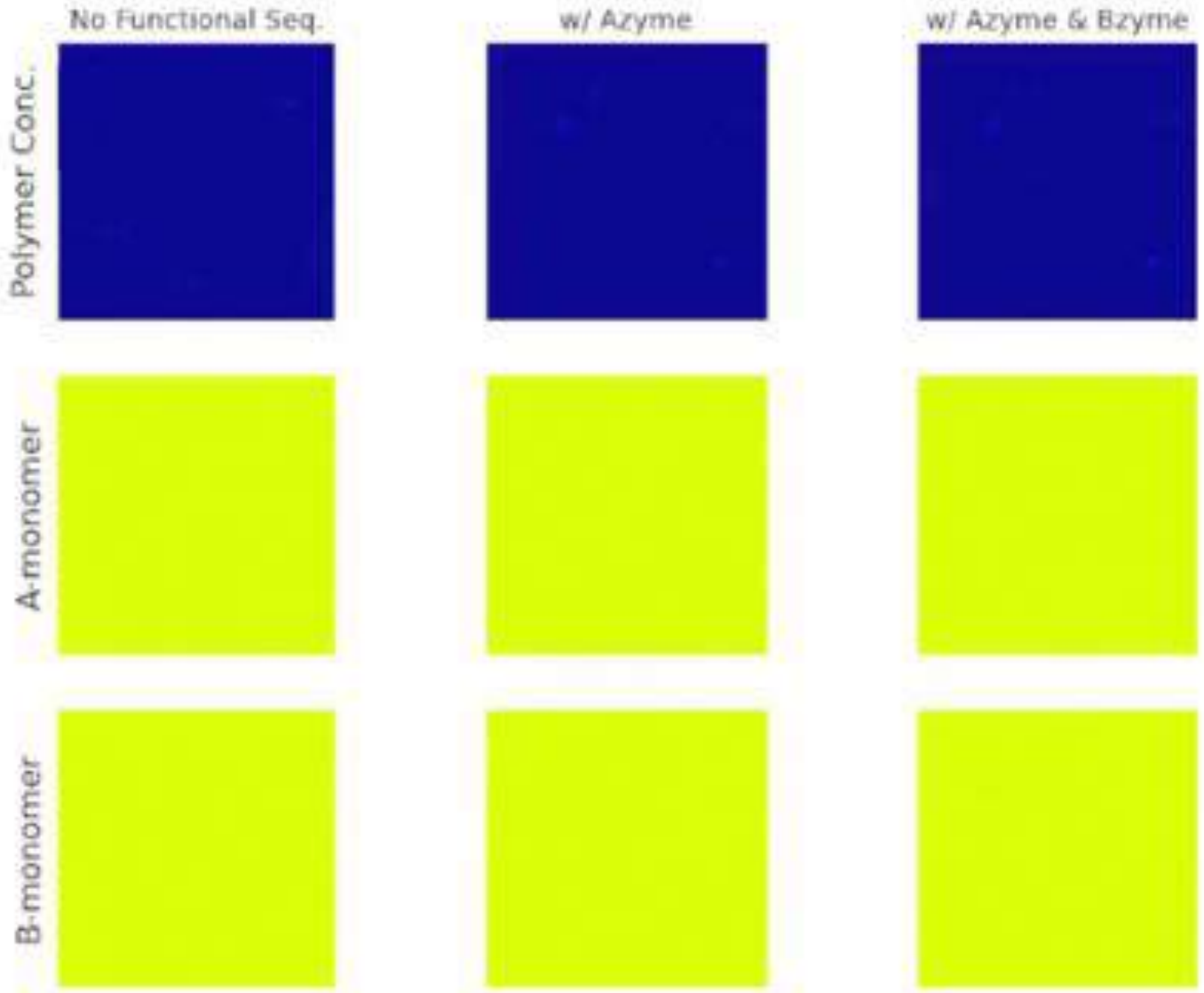
$k_h = 0.1$

$k_p = 0.01$

$k_m = 1.0$



Walker, Grover, Hud, *PLoSONE*, 7:e34166 (2012).


No Functional Seq.
w/ Azyme
w/ Azyme & Bzyme
Polymer Conc.
A-monomer
B-monomer

New insights

- Optimal system behavior occurs at “sweet spots.”
  - Some reversibility of polymerization is critical for the generation of diversity, but with too much, heredity is lost.
  - Some diffusion of species is needed for resource allocation, but limiting diffusion promotes diversity and selection.
- Cooperative networks could emerge step-wise.
- Emergent clustering
  - Did not expect to see clusters since the polymers have no explicit interaction in the model.
  - Clustering is driven by monomer recycling.
  - Limited diffusion on surfaces could provide early compartmentalization prior to lipid-based protocells.

CENTER FOR CHEMICAL EVOLUTION

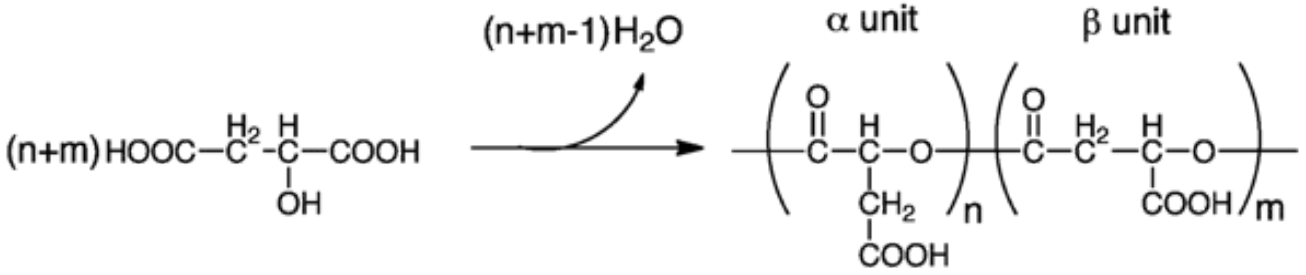
Features to be included in real systems

- Backbone reversibility and monomer recycling
- Environmental cycling
  - Drive condensation polymerization and hydrolysis
  - Drive duplex separation and replication
- Limited diffusion to bias mobility
- Selection based on function

CENTER FOR CHEMICAL EVOLUTION

Reversible linkages: a candidate

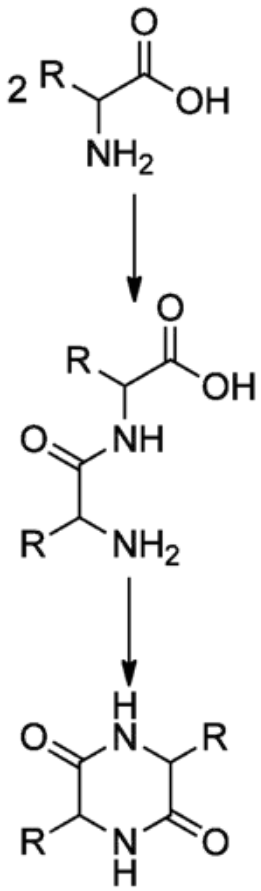
- Hypothesis: Polyesters are prebiotic precursors to peptides
  - Ribozome catalysis of α-hydroxy acid coupling
  - Polymer interactions defined by sidechains, not backbone hydrogen bonds
  - Ester bond polymerizes more readily than amide bond
- Polyesters can be synthesized under day/night cycles



Orgel, L. E., "Some consequence of the RNA world hypothesis." *Origins of Life and Evolution of the Biosphere*, **33**(2), 211-218 (2003).

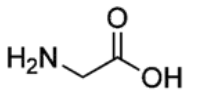
CENTER FOR CHEMICAL EVOLUTION

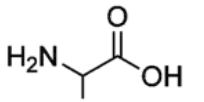
Amino and hydroxy acids

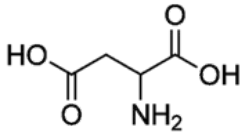


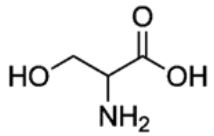
- Peptide problem
  - Bond formation free energy: 2-4 kcal/mol
  - Diketopiperazine sink
- Ester solution
  - Bond formatino free energy: -1 kcal/mol
  - Cyclization is reversible

α-amino acids

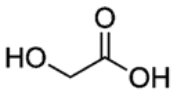

glycine

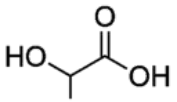

alanine

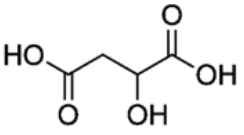

aspartic acid

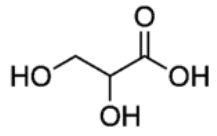

serine

α-hydroxy acids


glycolic acid


lactic acid


malic acid


glyceric acid

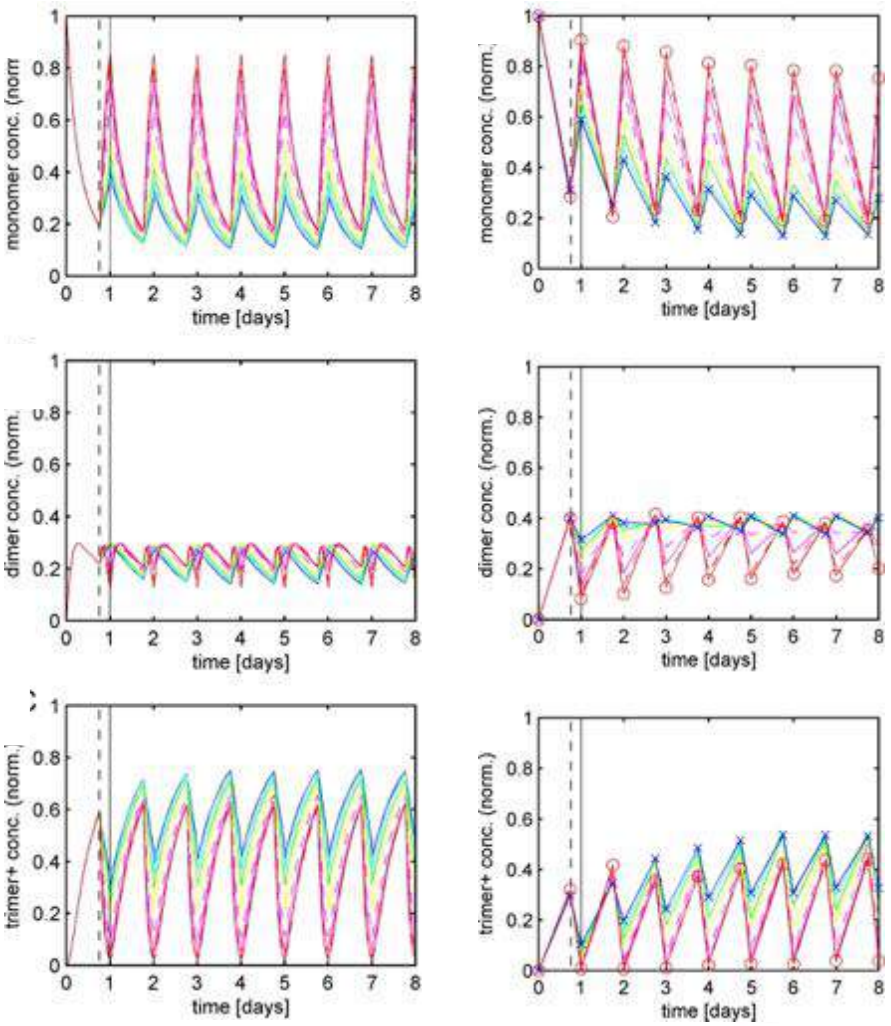
Martin RB. *Biopolymers* (1998)
Williams RJ *et al.*, JACS (1928)
Houk KN *et al.*, JOC (2007)

CENTER FOR CHEMICAL EVOLUTION

Polymerization under cycling

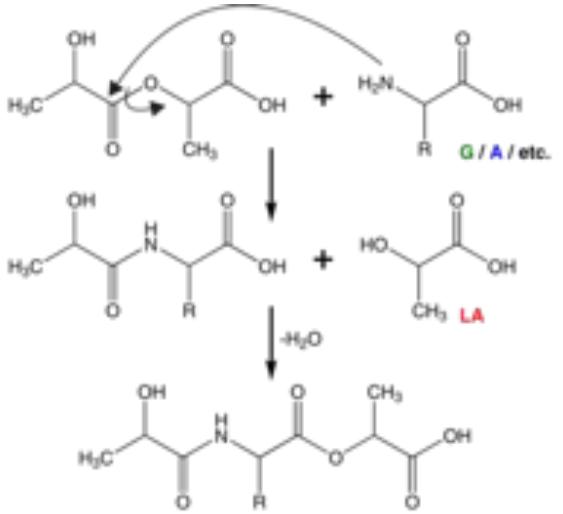
- 8 cycles of 1 day each
  - 18 hours open/dry
  - 0.5 hours rehydrate and sample
  - 5.5 hours capped
- 25 mM L-malic acid
  - No buffering/pH control
- Characterization by GPC
- Cross-over behavior seen in dimer
  - Monomer-dimer exchange
  - Dimer-trimer exchange
- Ratchet followed by cyclic steady state

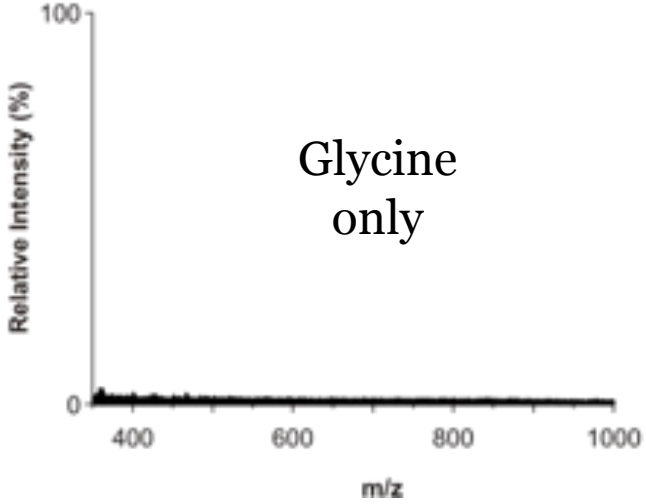
Mamajanov, *Macromolecules* (2014)



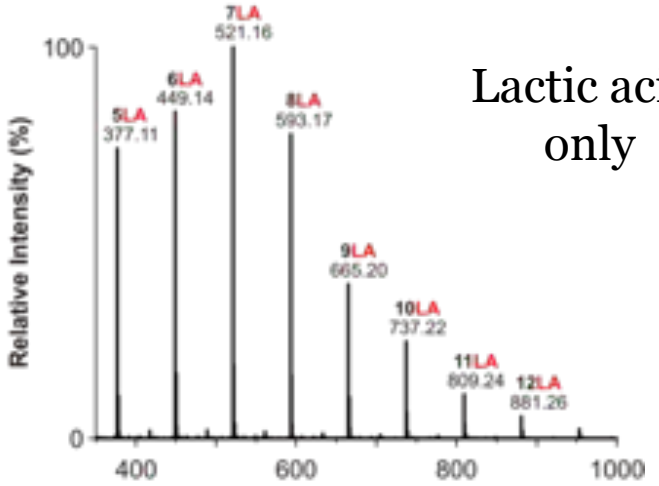
Dry T = 85° C, Wet T = 60-85° C

Ester-amide exchange: a path back to peptides

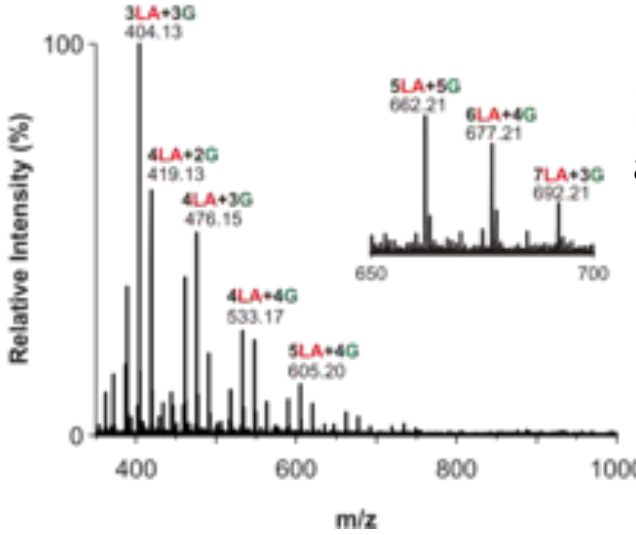




Glycine only



Lactic acid only

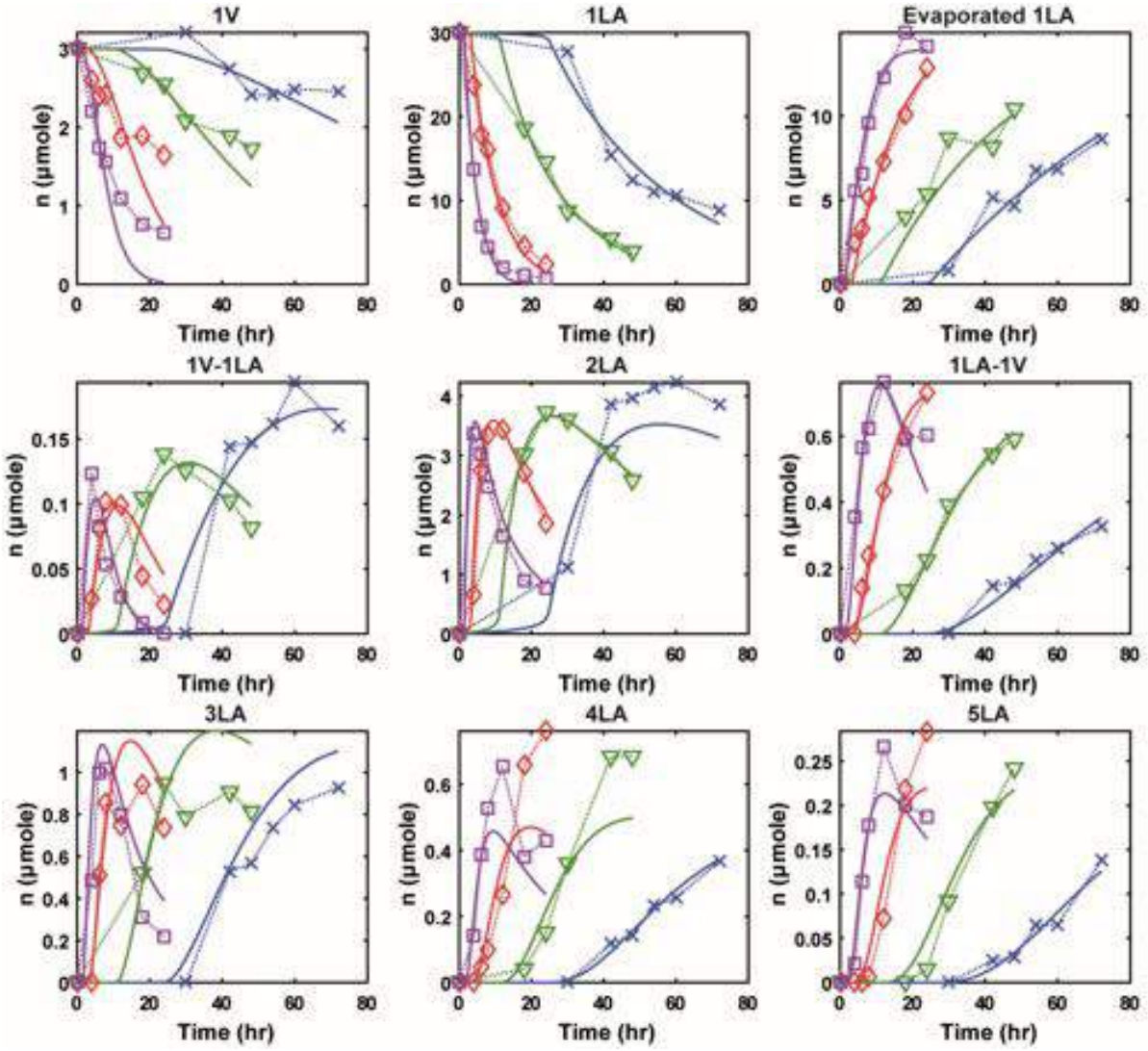


Lactic acid and glycine

CENTER FOR CHEMICAL EVOLUTION

Forsythe, Yu, *Angewandte Chemie* (2015)

A system-level model



Our theoretical results and the experimental data achieved agreement qualitatively and quantitatively.

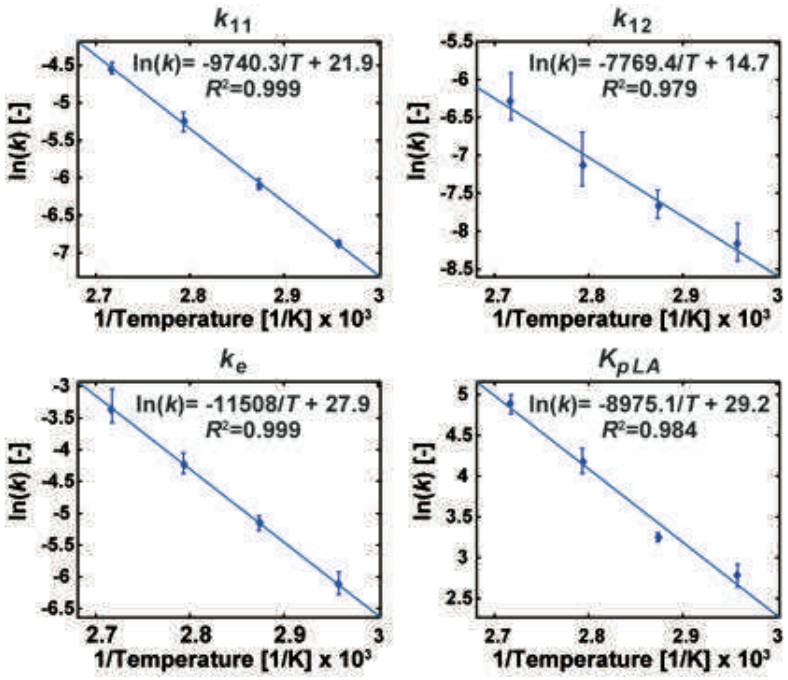
This allows us to pursue more in-depth understanding of the reaction.

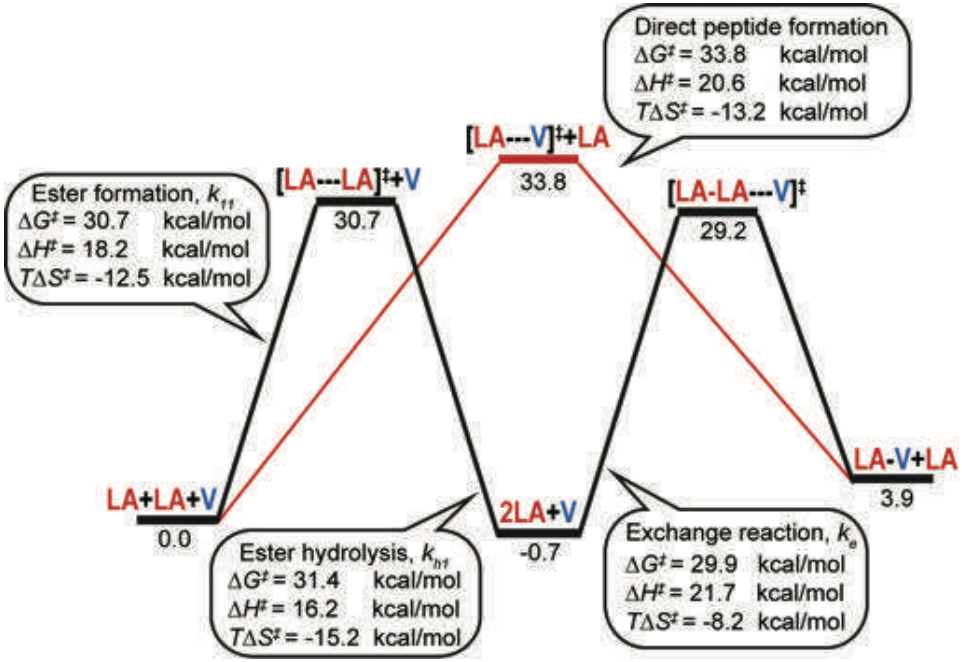
Reaction at four different temperatures:
95 ^{0}C, 85 ^{0}C, 75 ^{0}C and 65 ^{0}C.

Yu, *Phys Chem Chem Phys* (2016)

CENTER FOR CHEMICAL EVOLUTION

Analysis of the model





The rate constants follow the empirical Arrhenius equation even the reaction proceeds under nearly dry state.

Yu, *Phys Chem Chem Phys* (2016)

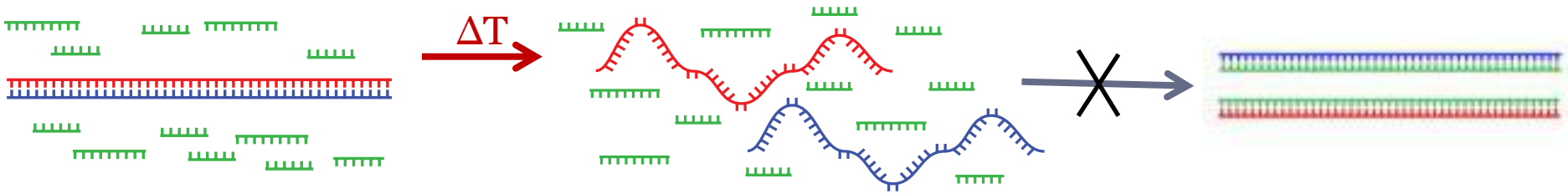
Calculation of reaction activation from the Arrhenius equation reveals the **ester-mediated pathway** is indeed more favorable than the direct amide bond formation.

New insights

- Original idea: Hydroxy acids could play the role of amino acids in prebiotic chemistry
- Turn in the road: Hydroxy acids may be the *catalyst* for amino acid polymerization
  - Ester-amide exchange with polyesters
  - Peptides form as the more stable product over time.
  - Depsipeptides could also be important co-polymer evolutionary intermediates.
- Life uses condensation polymers *because* they were able to recycle in water, and therefore evolve. Not a "problem."

The strand inhibition problem

- Before the appearance of coded proteins, duplex separation likely occurred by heating.



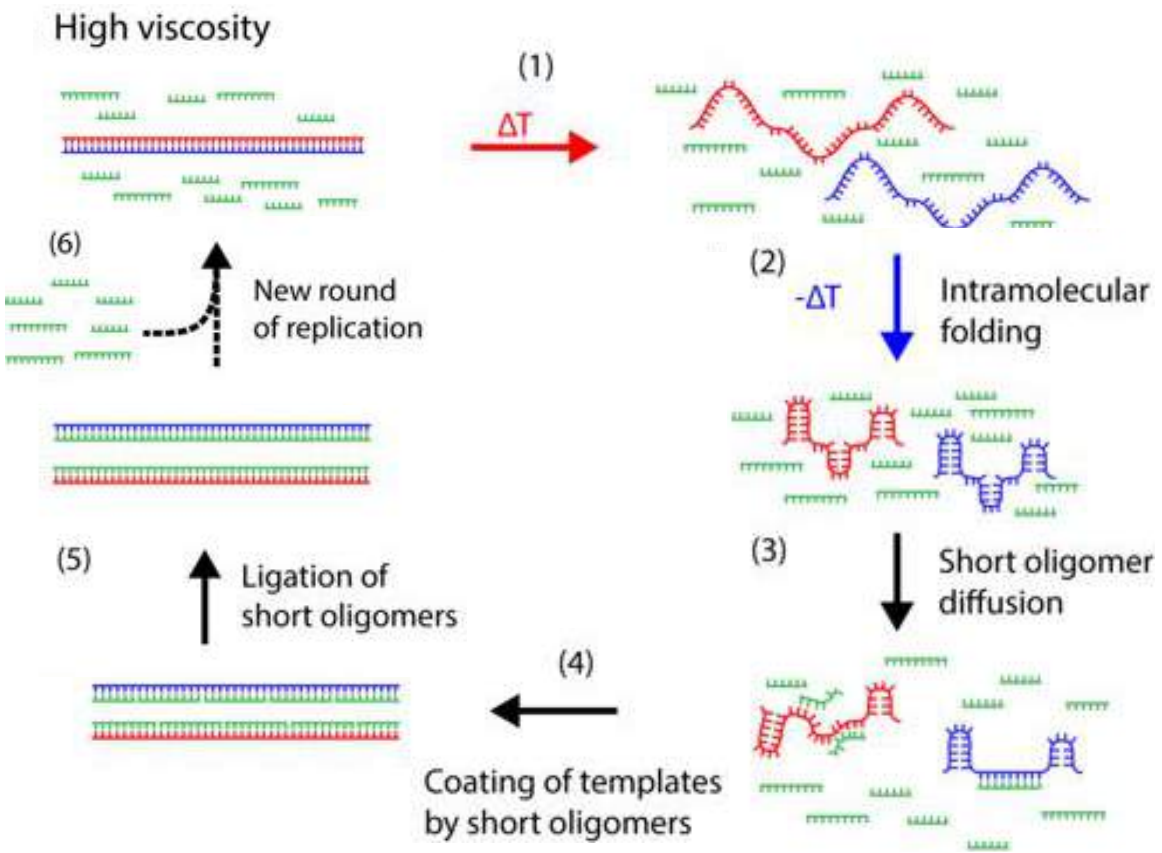
- However, at lower temperatures, the template duplex reforms before oligonucleotides have the opportunity to associate with the single stranded templates.
- Synthesis on a template strand is inhibited by the presence of the complementary template strand.

Szostak, *Journal of Systems Chemistry* (2012)

CENTER FOR CHEMICAL EVOLUTION

Hypothesis

Thermal cycling in viscous environments can be utilized to overcome strand inhibition and promote template-directed nucleic acid synthesis.

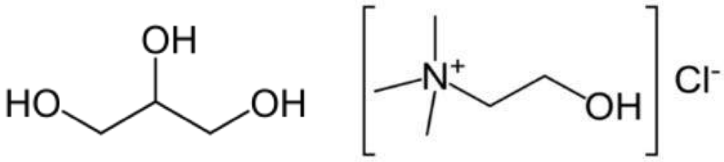


Viscosity-mediated replication can promote copying from long (gene-length) templates of arbitrary sequence.

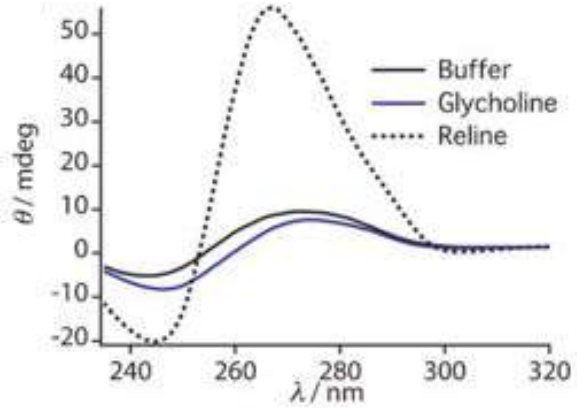
He, *Nature Chemistry* (2017)

Choosing a viscous environment

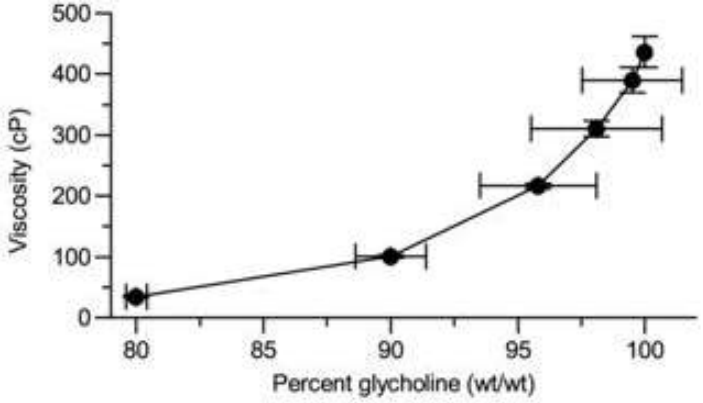
Glycholine: glycerol and choline chloride in a 4:1 molar ratio



B form DNA is retained



Highly hygroscopic and miscible with water

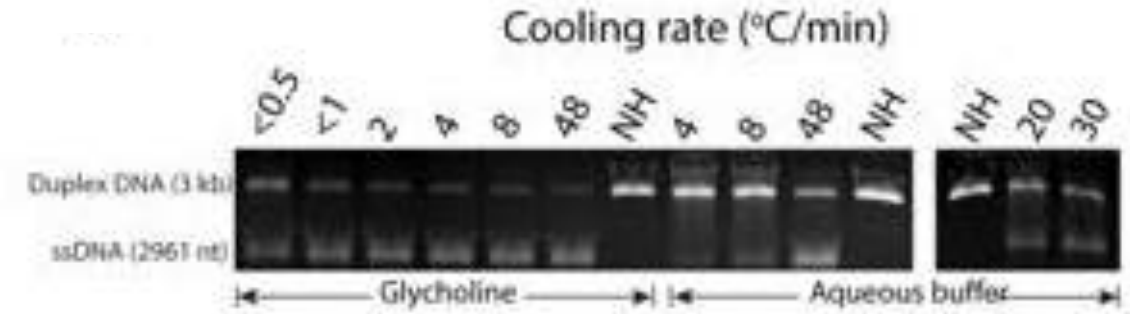


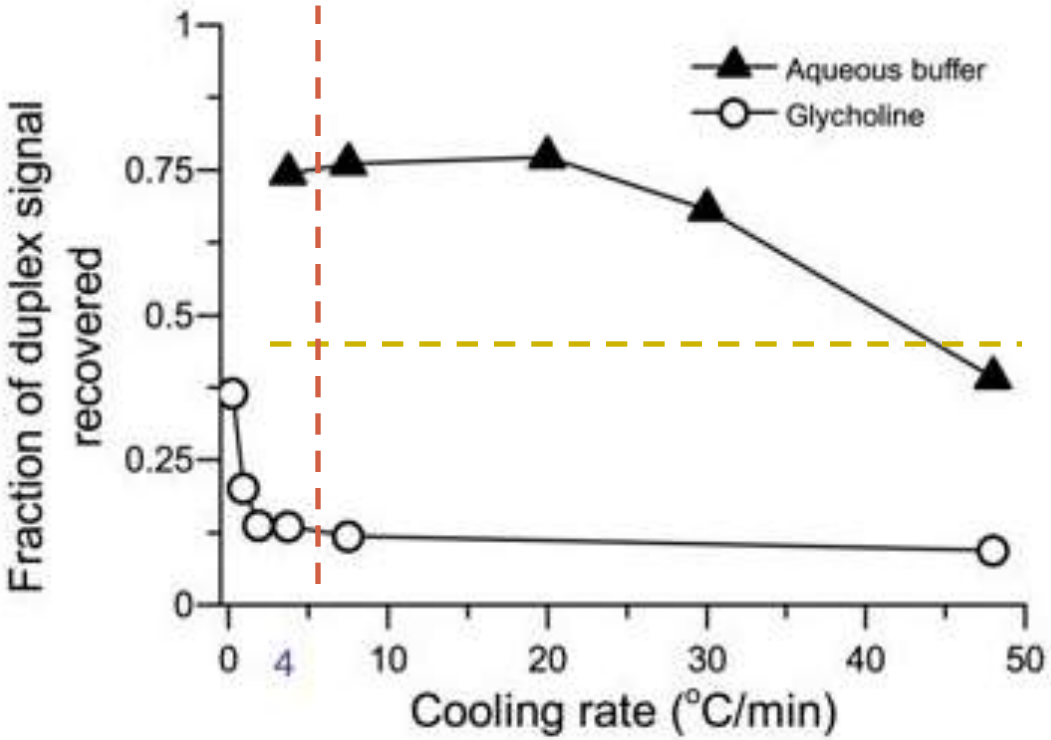
Does not depress duplex melting temperatures as severely as reline

Melting temperatures of DNA duplex species				
Species	**Base Pairs**	**%GC**	**T_m in Aqueous Buffer (°C)**	**T_m in Glycholine (°C)**
Hairpin	7 bp stem, 3 nt loop	41	79.5	44.7
32-mer duplex	32	47	72.2	49.0
3 kb duplex	2957	50	88.0	50.7

CENTER FOR CHEMICAL EVOLUTION

Tuning kinetic trapping with cooling rate





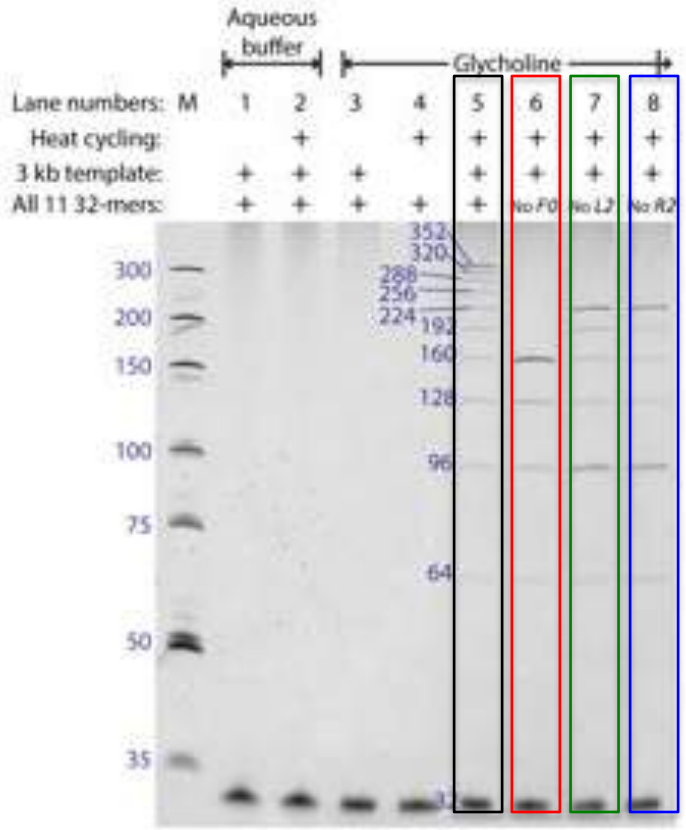
Much more gradual cooling rates can be utilized to kinetically trap a 3 kb duplex as single strands in glycholine compared to aqueous buffer.

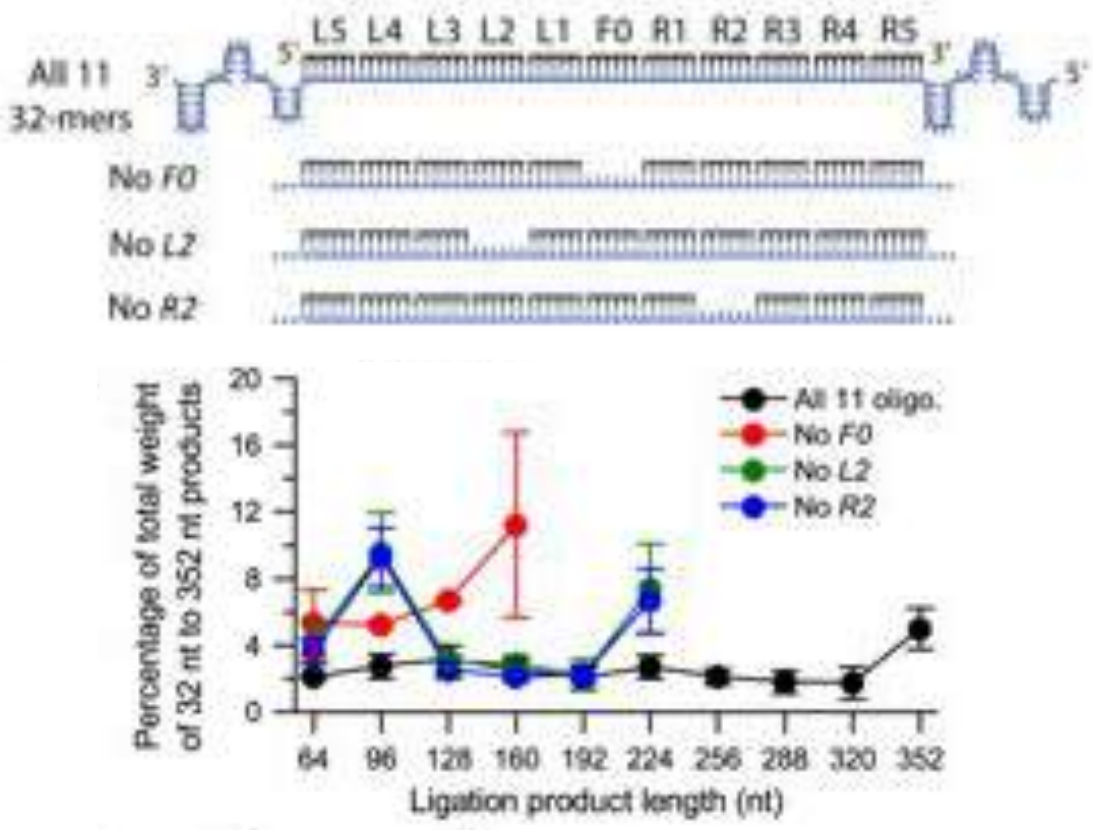
A cooling rate of 4°C/min was utilized in all subsequent experiments.

CENTER FOR CHEMICAL EVOLUTION

Ligation of assembled oligonucleotides

Ligation was carried out by T4 DNA ligase, added *after* the oligonucleotides have already been assembled on the template (overcoming strand inhibition).





Information transfer from a 3kb duplex template

He, *Nature Chemistry* (2017)

CENTER FOR CHEMICAL EVOLUTION

Progression of ideas

- Earlier idea: Use non-aqueous solvents drive condensation polymerization, solving the water problem (Mamajanov, Angewandte Chemie (2010)).
- Original idea: Add viscogens to *aqueous* buffer to preferentially slow the reannealing of the duplex.
- Version 2: Use small molecule *non-aqueous* viscous solvents instead.
- Version 3: Desired behavior observed, perhaps more due to *intramolecular folding* than preferential diffusion.
- New insight: Viscous environments may help select for long strands and folded structures.
  - Under "typical" conditions, unfolded structures are easier to replicate because they are more accessible.

Closing remarks

- The Peptide World may be more accessible than has previously been thought, challenging the dominant RNA World Hypothesis.
- In an aqueous environment, condensation polymers may have been selected in life because they could evolve.
- Viscous environments may drive selection for folding and function.
- Research serendipity may (or may not) suggest that these systems mimic some key features of prebiotic environments and chemistries.

CENTER FOR CHEMICAL EVOLUTION

Acknowledgments

- Members of the Center for Chemical Evolution
- Sheng-Sheng Yu, Christine He, Sara Walker, Irena Mamajanov
- Nick Hud, Ram Krishnamurthy, Joe Schork
- Funding from NSF and NASA (CHE-1504217)

Publications

- "A viscous solvent enables information transfer from gene-length nucleic acids in a model prebiotic replication cycle," C. Y. He, I. Gállego, B. Laughlin, M. A. Grover, N.V. Hud, *Nature Chemistry* (2017), DOI:10.1038/nchem.2628.
- "Kinetics of prebiotic depsipeptide formation from the ester–amide exchange reaction," S. Yu, R. Krishnamurthy, F. M. Fernández, N. V. Hud, F. J. Schork, M. A. Grover, *Phys. Chem. Chem.* Phys., **18**, 28441–28450 (2016), DOI: 10.1039/c6cp05527c.
- "Universal sequence replication, reversible polymerization and early functional biopolymers: A model for the initiation of prebiotic sequence evolution," S. I. Walker, M. A. Grover, N. V. Hud, *PLoSONE*, 7(4) e34166 (2012), DOI: 10.1371/journal.pone.0034166.
- "Surveying the sequence diversity of model prebiotic peptides by mass spectrometry," J. G. Forsythe, A. S. Petrov, W. C. Millar, S.-S. Yu, R. Krishnamurty, M. A. Grover, N. V. Hud, F. M. Fernandez, *PNAS* (2017), DOI: 10.1073/pnas.1711631114.

CENTER FOR CHEMICAL EVOLUTION