

A Ribonucleotide Origin for Life – Fluctuation and Near-ideal Reactions

Michael Yarus

Received: 9 November 2012 / Accepted: 3 January 2013 /

Published online: 24 January 2013

© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Oligoribonucleotides are potentially capable of Darwinian evolution – they may replicate and can express an independent chemical phenotype, as embodied in modern enzymatic cofactors. Using quantitative chemical kinetics on a sporadically fed ribonucleotide pool, unreliable supplies of unstable activated ribonucleotides A and B at low concentrations recurrently yield a replicating AB polymer with a potential chemical phenotype. Self-complementary replication in the pool occurs during a minority (here $\approx 35\%$) of synthetic episodes that exploit coincidental overlaps between 4, 5 or 6 spikes of arbitrarily arriving substrates. Such uniquely productive synthetic episodes, in which near-ideal reaction sequences recur at random, account for most AB oligonucleotide synthesis, and therefore underlie the emergence of net replication under realistic primordial conditions. Because overlapping substrate spikes are unexpectedly frequent, and in addition, complex spike sequences appear disproportionately, a sporadically fed pool can host unexpectedly complex syntheses. Thus, primordial substrate fluctuations are not necessarily a barrier to Darwinism, but instead can facilitate early evolution.

Keywords Cofactor · Origin of life · Replication · Selection · Initial Darwinian Ancestor

Introduction

It is a widely held hypothesis that the pivotal event in the origin of life was the origin of a replicating RNA molecule (Wu and Higgs 2011). However, there is as yet no “grand synthesis” that produces RNA, or a molecular congener, on the early Earth. Nonetheless, there has been substantial progress toward prebiotic synthesis of ribonucleotides, using precursors arguably credible under primitive planetary conditions. 2',3' cyclic pyrimidine nucleotides are recent examples, produced from cyanamide, cyanoacetylene, glycolaldehyde, glyceraldehyde and free phosphate (Powner et al. 2009). Biological purines have long been known to be synthesized

M. Yarus (✉)

Department of Molecular, Cellular and Developmental Biology, University of Colorado,
Boulder, CO 80309-0347, USA
e-mail: michael.yarus@Colorado.EDU

from NH₄CN (Oró and Kimball 1961; Borquez et al. 2005). Ribose is produced in low yield from HCHO, but in elevated yield from reactions containing HCHO, glycolaldehyde and minerals (Kim et al. 2011). Condensing purines with ribose to make purine nucleosides is easier than for pyrimidines, and occurs moderately efficiently upon heating dry materials with trimetaphosphate and magnesium (Fuller et al. 1972a, b). Purine nucleosides can be phosphorylated at low efficiency using unexceptional mineral sources of phosphate such as hydroxylapatite (Costanzo et al. 2007). Thus, it seems timely to ask: how much might be achieved after we generate primordial pyrimidine and purine ribonucleotides, and activate them?

In previous work (Yarus 2012), production of occasional low concentrations of a 5' phosphate-activated nucleotide (A) and a complementary, chemically reactive, otherwise normal 5' nucleotide (B), yields a kinetically plausible chemical origin for Darwinian life on Earth (in other words, an AB molecule that replicates and has a chemical phenotype), from known homogeneous chemical reactions. These assumptions are inspired by the existing example of dinucleotide enzyme cofactors (Yarus 2011a), like NAD. Below I look more deeply into the crucial events required for episodes of templated replication, which underlie Darwinian change in AB.

Methods

Reactions consisting of all of Fig. 1 (the “sporadically fed pool”) or subsets of the colored reactions (“simultaneous, stable substrates” or “no decay”) were expressed as systems of

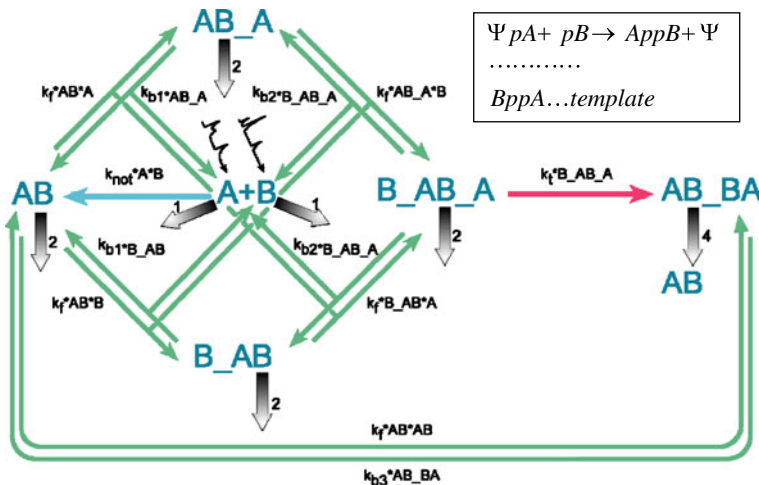


Fig. 1 Reactions of the sporadically fed pool. A and B are precursor nucleotides; A is activated for polymerization. Underscoring joins complementary base-paired reactants. A and B are present at constant concentrations or appear in spikes at uncorrelated, random times, and in amounts that are distributed as a Gaussian (sporadically fed pool mechanism; symbolized in jagged *black supply arrows*, center). *Colored arrows* represent steps which occur in both the full sporadically fed pool, and the pool with simultaneous stable substrates or no decay, used for comparison. Reaction schemes (Fig. 1) were integrated (as systems of ordinary differential equations) to yield the data shown in later figures. Direct chemical reaction of A and B can create AB dimer (*blue arrow on left*; rate constant k_{not} for notemplate). This can pair in a complementary fashion because A and B are self-complementary (*central box of green arrows*). Once completely paired, base-paired A and B paired to an AB template react to form a complementary dimer (*magenta arrow on right*, rate constant k_t , representing the rate with template). Paired dimers can dissociate to yield two AB (*green loop at bottom*), or separated AB can reassociate to basepaired dimer

ordinary differential equations and integrated by Berkeley Madonna v 8.3.18 with post-processing of kinetic array data in Microsoft Excel 2003 SP3 (Yarus 2012). Code used for simulation is available there (Yarus 2012) as a supplement.

Time is measured in mean lifetimes or average times to decay (half-life = $\ln 2 \cdot$ mean lifetime) for precursors A and B (which are assumed to be equally unstable). This ties the timescale to A and B survival, so that variations in the stability of A and B are more easily envisioned. To give a specific example, under our standard experimental conditions at 0° and pH 8, nucleotide imidazolides have mean lifetimes of about 100 days. Ribonucleotide substrates A and B arrive at the pool as randomly-timed, independent, variable but Gaussian-distributed spikes of $4 \mu\text{M} \pm 1 \mu\text{M}$ (standard deviation). Mean arrival frequency is low, 1 spike / 10 lifetimes, and the word “spikes” means that substrate arrival is linear over 0.01 lifetime. Dissociation rates are $k_{b1} = 0.2\text{E}4 \text{ lifetime}^{-1}$, $k_{b2} = 0.2\text{E}3 \text{ lifetime}^{-1}$, $k_{b3} = 0.2\text{E}2 \text{ lifetime}^{-1}$ throughout, and (templated polymerization) $k_t = 1000 \text{ lifetime}^{-1}$, (untemplated polymerization) $k_{\text{not}} = 10 \text{ M}^{-1} \text{ lifetime}^{-1}$, and (basepairing) $k_{b1} = k_{b2} = k_{b3} = 10^8 \text{ M}^{-1} \text{ lifetime}^{-1}$. These standard pool values have been rationalized elsewhere (Yarus 2012) by choosing values which are observed or slower (less favorable to replication) than published rates. All molecules in the sporadically fed pool are unstable. Gray shaded arrows represent decay in Fig. 1, and are marked with relevant mean lifetimes: 1 (for A and B), 2 (for all forms of AB) and 4 (for paired AB; which, uniquely decays to a single surviving AB). Relative lifetimes are estimated; AB and paired AB are made slightly more stable (longer mean lifetime) because increasing secondary structure and base pairing stabilize other nucleic acids (Lindahl 1993).

Results

Figure 1 shows synthesis and decay in a sporadically fed pool (Yarus 2012) which hosts replication of a small, self-complementary ribonucleotide. The mechanism includes A and B substrate supplies which are unrelated and delivered sporadically, parallel undirected chemical and templated AB synthesis, and general molecular instability appropriate for materials in cold, primordial conditions. Nucleotide substrates arrive as Gaussian-distributed, randomly timed A and B substrate spikes (jagged arrows, middle, Fig. 1), undergo unguided chemical polymerization (blue arrow), base-pairing (square of green arrows), and possibly replication (magenta arrow), with first-order decay of all molecules (gray-gradient arrows). The (green) loop at the bottom represents pairing and dissociation of the base-paired dimer (von Kiedrowski 1986), AB_BA (underscores symbolize base pairing), which is the replication product of self-complementary A and B. Colored arrows can be taken together to describe other reaction logic: for example, reliable, constant supplies of A and B, which stable synthesis is later contrasted with the sporadically fed pool (Yarus 2012).

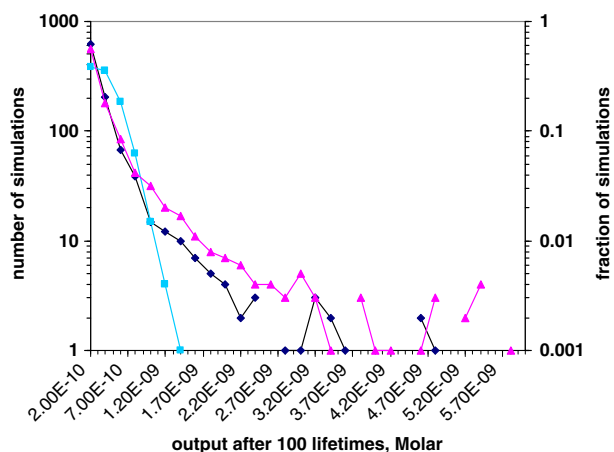
The Fig. 1 inset (upper right) describes a possible AB synthesis in more detail. Ψ is an activating group that allows polymerization, as in the nucleotide phosphorimidazolide introduced by Orgel (Sawai and Orgel 1975), and shown to have a simple abiotic synthesis by Lohrmann (Lohrmann 1977). Below the dotted line is a possible template (A and B are assumed to be complementary; (Yarus 2012)), to emphasize that AB synthesis can plausibly proceed via either untemplated (inset top only; (Kanavarioti et al. 1992)) or templated means (replication; inset top + bottom). The AB backbone is drawn 5'-5' in emulation of cofactors like NAD, which are ancient (White 1976) and conceivably combine templating and chemical activities (Yarus 2011a). However, the chemical identity of AB is not crucial to conclusions here, though it can likely be identified by a Bayesian inquiry (Yarus et al. 2005) into the existence of crucial templating reactions.

Net replication in a sporadically fed pool is explored in Fig. 2, which plots number of pools versus total AB output in 1,000 consecutive simulations run for 100 A or B lifetimes. Values employed for rates and equilibria are those of the “standard system” used previously ((Yarus 2012), Fig. 2), which was designed to emulate known RNA chemistry and be mildly replicating at 100 A or B lifetimes. The plot compares integrated direct synthesis (blue in Fig. 1), integrated templated AB synthesis (magenta) and the largest AB peak (black).

Pool histories that yield large and small AB synthesis (Fig. 2) are different in a suggestive way. The probability of direct chemical synthesis of AB (blue) declines exponentially with size of synthetic output, after a shoulder where size of the synthetic event makes little difference. By comparison, peak AB template (black) and total templated synthesis (magenta) first decline exponentially together, then flatten to a long tail where templated output exceeds peak AB template. These larger synthetic events also show increasing bias towards replication (that is, to a higher ratio of templated to direct synthesis), insofar as the statistics of 1,000 pools allows comparison (rightward in Fig. 2). In fact, the slope of Fig. 2’s peak and templated AB curves appear to decrease with larger synthesis, and there appears an unexpectedly large fraction (a few tenths of a percent) of very productive pool histories, which produce large amounts of AB via exceptionally extensive replication. Thus, the key to replication lies in atypical large synthetic episodes, where large concentrations of template AB, which are required for replication (but irrelevant to direct chemical synthesis) exist.

To clarify the connection between efficient net synthesis and replication, 250 consecutive curated AB-synthetic episodes were collected for a standard system. “Curated” means that these 250 AB syntheses were isolated (the first synthetic episodes to occur) and therefore independent of other events. “Episode” includes all events associated with AB synthesis for the lifetime of one AB population. Operationally, an episode begins with the first spike that will alter AB output (see Fig. 6 discussion below for examples), and ends when net integrated AB synthesis becomes constant to the 5th calculated significant figure. Curated episodes were individually measured; so direct and templated AB synthesis are causally associated within this set of 250 episodes, and further, each can be associated with its own instantaneous AB peak (instead of the less directly relevant) largest peak during 100 lifetimes, as for Fig. 2. Figure 3 shows 250 individual total (direct + templated) AB syntheses, plotted as a function of the number of substrate spikes in the episode.

Fig. 2 Numbers of 100-lifetime simulations with particular integrated AB output after 1000 total simulations of the sporadically fed pool. *Blue* is integrated direct AB synthesis (*blue arrow* in Fig. 1); *magenta* is integrated replication (templated synthesis; *magenta arrow* in Fig. 1), and *black* is the largest AB peak during a 100 lifetime pool simulation



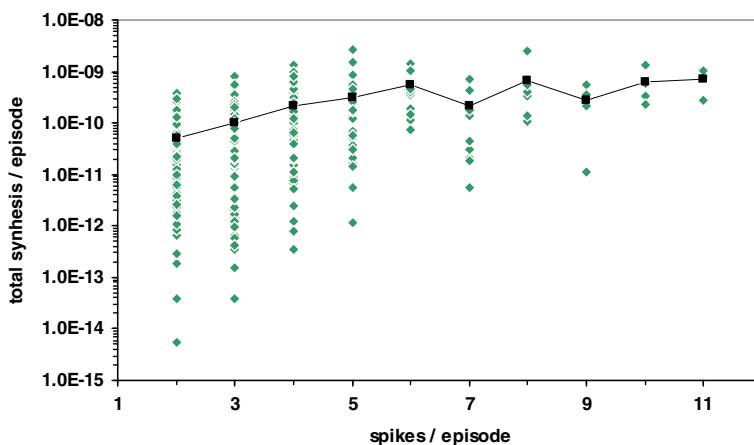


Fig. 3 Total sporadically fed pool output during 250 consecutive curated synthetic episodes. *Diamonds* – total AB synthesized in 250 individual synthetic episodes. *Squares* – mean total AB output from each type of episode (that is, with the same number of A and B spikes)

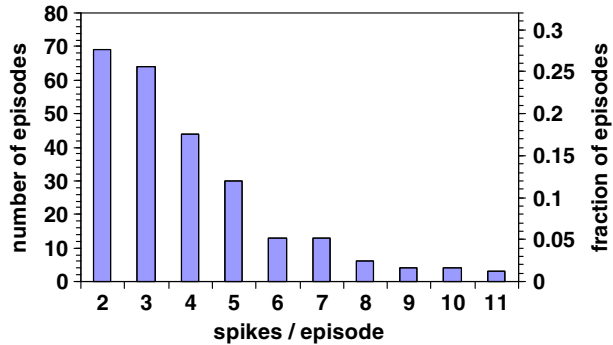
Episodic synthesis is highly varied, with AB yields ranging over about 7 orders of magnitude. Further, episodes of similar complexity vary – even the simplest synthetic episodes, with 2 intersecting spikes of substrate, give total yields of AB ranging over 5 orders of magnitude in this sample of 250. Thus many AB magnitudes are not associated with any particular history. Indeed, it would be possible to choose a range of total AB synthesis which could have occurred by intersection of 2 to 11 substrate spikes.

Nonetheless, there are clear regularities in Fig. 3. The smallest events increase in size from 2 to 6 spikes. Therefore, if one chose the smallest syntheses, these would be predominantly the simplest episodes, with 2 and 3-spikes. Similarly, large syntheses increase from 2 to 6 spikes: if one chose the largest syntheses, these would be 4, 5 and 6 spike episodes, with a definite but smaller contribution from more complex events. Mean AB yields (black) increase 11-fold from 2 to 6 spikes, and thereafter do not notably increase. The most complex events are not as well-determined because there are few of them in this sample of 250 (Fig. 3). Nevertheless, because every large event (having 7–11 spikes/episode) lies below the projection of the relation from less complex episodes (having 2–6 spikes/episode), more complex events do not have increased output. This, because mean substrate arrival is fixed at once per 10 lifetimes, may be because more complex spike trains allow more time for decay, which nearly balances the effect of their greater substrate input. These characteristics are central to the potential synthetic capacity of the sporadically fed pool (*Discussion*, below).

This distribution of spikes/episode is clarified in Fig. 4. The simplest synthetic episode, with two intersecting spikes (of different kinds, since AB synthesis must result) is narrowly the most frequent, at about 27.6 % of all episodes. However, even though A or B substrate spikes arrive at long average intervals (averaging 1 spike per 10 A or B lifetimes), it seems useful to restate the same fact by saying that a substantial majority, 72.4 % of all synthetic episodes, involve the coincidence of 3 or more substrate spikes (Fig. 4). And the tail at the right of Fig. 4 seems quite clear; more complex events are increasingly more probable than intuition might expect. For example, standard system events that engage 9, 10 or 11 substrate spikes are each a few percent of total AB synthetic episodes.

The route to net replication in this randomly-supplied pool is elucidated in Fig. 5, which shows integrated total AB output (black), AB output via unguided chemical synthesis (blue;

Fig. 4 Distribution of synthetic episodes among observed spike / episode types. *Left ordinate* – number of episodes out of 250 curated examples, using standard spikes. *Right ordinate* – fraction of episodes in each class of curated events

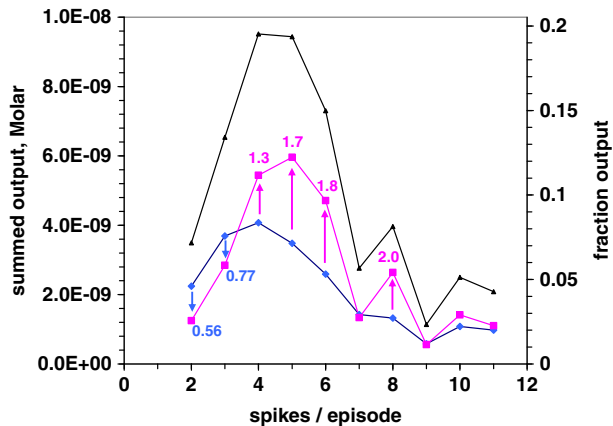


blue arrow in Fig. 1), and templated AB synthesis (magenta; magenta arrow in Fig. 1), together against the same scales. In the center of the graph, the net replication in each kind of curated synthetic episode is shown as the ratio of templated (magenta) to direct (blue) synthesis (numbers, arrows). Notably, the three largest sources of total synthesis (4, 5 and 6 spikes) coincide with the three largest sources of AB from templated synthesis (replication). In fact, two- and 3-spike episodes do not produce net replication under standard conditions (Fig. 5, blue arrows). Thus, all other considerations aside, synthetic episodes in which 4, 5 or 6 spikes contribute dominate the total synthesis of AB (54 % of total output (Fig. 5) from 35 % of episodes (Fig. 4)), thereby making the overall system one that replicates. There is also a small contribution to replication from 7 to 11 spike episodes, but this is less significant because, despite their similar size, they are less frequent (Figs. 4 and 5).

The ‘standard system’ was chosen to be one that replicated to a small degree, just ‘past the Darwinian boundary’, in order to investigate the onset of replication (Yarus 2012). If the mean replication of the curated system in Fig. 5 is calculated by summing the products (fraction output times the ratio of templated to direct synthesis) for all episodes, a system composed of these curated episodes replicates 1.36-fold, in agreement with prior overall behavior of the standard pool (Yarus 2012). Thus the 250 curated episodes quantitatively account for the mean behavior of the standard sporadically fed pool integrated over 100 lifetimes, supporting this episodic analysis.

These outcomes can be explained: replication is more complex than direct chemical synthesis of AB, because templated synthesis requires the prior synthesis of an AB template.

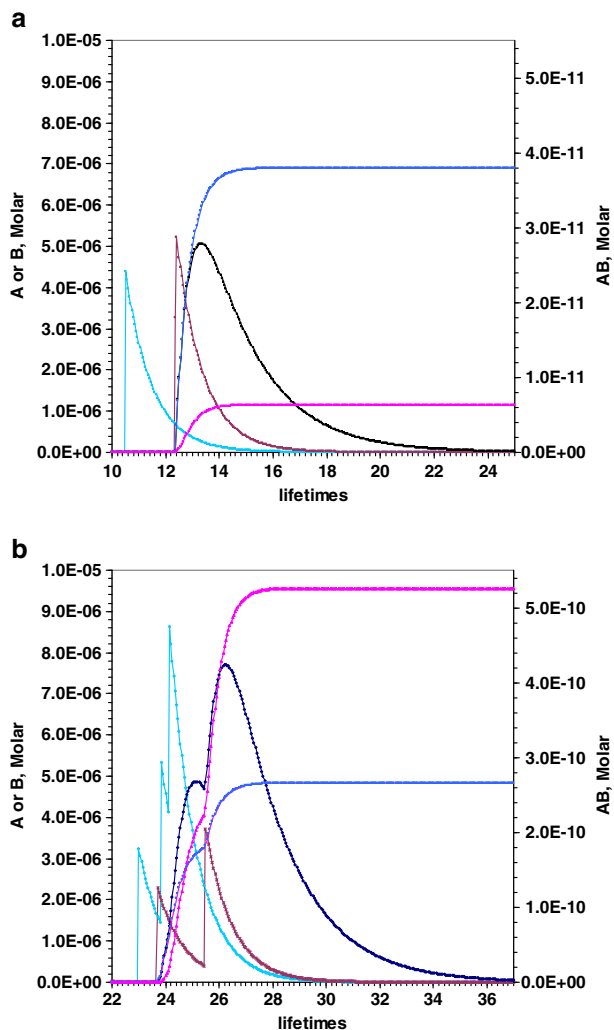
Fig. 5 Total, templated and direct output from each type of episode in the 250 curated episodes. *Black* – total AB, *Magenta* – templated AB, *Blue* – directly synthesized AB. Numbered arrows give ‘fold-replication’ for each episode class. *Left ordinate* – total output, summed. *Right ordinate* – fraction of total output, summed



Consider designing a reactor to produce AB - delivery of a spike of A and a spike of B in either order suffices for direct chemical synthesis. However, to replicate in the reactor we must ideally make AB template and then supply unstable A and B again for templated synthesis. Therefore, the ideal sequence of substrate spikes for a replication reactor has ≥ 4 spikes. Importantly, the sporadically fed pool is a reactor that utilizes near-ideal reaction sequences for replication without outside instruction, relying only on random substrate arrival to recurrently replicate AB, and thereby recurrently test the potentialities of Darwinian change.

This discussion can be made more concrete by comparing example episodes (all events significantly changing synthesis during the course of a single population of AB) from standard pool simulations. Figure 6a and b illustrate the kinetics for a typical 2-spike episode and a 5-spike episode, respectively, plotted over 15 A or B lifetimes. For clarity, only one of every 50 calculated kinetic points is shown.

Fig. 6 Simple (a) and complex (b) synthetic episodes in a complete sporadically fed system; chosen for illustration a. Two substrate spikes coincidentally overlap. Light blue is substrate A; brown is substrate B (both on left axis); blue is direct AB synthesis; magenta is templated AB, black is total AB in all forms and from all sources (all AB on right axis). **b** Five substrate spikes coincidentally overlap during the history of one AB population. Color and plotting the same as panel A, but note difference in right axis scales. For clarity, only every 50th calculated point has been plotted



Panel 6a shows the simplest kind of episode, in which single peaks of A (at 10.4 lifetimes, light blue) and B (at 12.3 lifetimes, brown) appear in the pool at accidentally overlapping times. As a result, a peak due to direct chemical synthesis of AB appears (black). A and B substrates are sufficiently stable to overlap prior untemplated AB. Thus there is (after ≈ 12.5 lifetimes) also replication (magenta) of previously chemically synthesized AB (blue). However, A and B have decayed substantially (declines on the right of A and B peaks; e^{-1} per mean lifetime) by the time replication is under way. Thus, total instantaneous AB (black) and chemically synthesized AB (blue) visibly diverge (at > 13 lifetimes). Accordingly, in panel 6a, AB template replication is limited by the availability of free A and B, yielding 16.6 % replication (magenta on right divided by blue on right).

Figure 6b shows 15 lifetimes during a more complex, rarer (Fig. 4) 5-spike episode, embracing 3 A spikes of various sizes, as well as 2 spikes of B. This episode more effectively synthesizes AB (note the larger scale for AB on the right, compared to panel 6a). Though there is only 0.1 spike of A or B per lifetime on average, by chance 3 spikes of A occur during the survival of the first one (at ≈ 23 lifetimes). This (blue) almost triples substrate A available for synthesis, to greater than double the mean spike size. Thus, random arrival of A (the first before any AB synthesis) can yield elevated total A, as well as yielding usefully sequenced and timed substrates.

Secondly, the random sequence of A and B spikes is here very productive. After total AB begins its rise (black; 23.7 lifetimes) due to the first spike of B (note that this represents direct synthesis - (blue) and total instantaneous AB (black) rise together), later spikes of A and a second spike of B enable a second peak of total AB (just past 26 lifetimes) which is mostly replication (note that templated synthesis (magenta) and total AB (black) rise together, almost identical). By contrast, total direct chemical AB synthesis (blue) is more subdued late in this episode. The result is AB mostly via replication (magenta/blue = 1.98 at 37 lifetimes, on the right). Recurrence of episodes like Fig. 6b account for the predominance of replication of the standard pool (Fig. 5). Further, Fig. 6b illustrates the extension of AB lifetime during more complex events, which underlies a realistic estimate of the capabilities of the sporadically fed pool (Discussion, below).

Discussion

Taking current calculations with prior results, known ribonucleotide solution chemistry appears sufficient to initiate Darwinian evolution on Earth. Some chemical qualities of a primordial ribonucleotide replicator may even be specified from biological examples (Yarus 2011a) or from calculations based on the likely chemical environment (Yarus 2012). Therefore, if small chemically competent oligonucleotide replicators exist, we can now outline an empirically validated scheme for the origin of Darwinian life. Accordingly, a major experimental task now is to detect such small replicators, and study possible ribonucleotide origins by examining their properties (Yarus 2012). In this work, new properties for the earliest selectable replicating system (the IDA) appear, implicit in the apparently simple chemistry of the sporadically fed pool.

Crucial Templating Events In A Sporadically Fed Pool A standard sporadically fed pool is poised just above the 'Darwinian boundary' (Yarus 2012) at which net templated replication begins. Thus the properties of the standard pool should account for this beginning.

Net replication (Fig. 1) is specifically associated with a class of efficient templated AB synthesis events (Fig. 2). Such association of template and product is a quality expected of replication, but not of direct chemical AB synthesis. Considering measurements on 250

individual synthetic episodes, elevated production of AB can be traced to a specific subset of synthetic episodes in which multiple A and B spikes-at-random intersect a single surviving population of AB templates (Fig. 3). These productive syntheses are a substantial minority of all synthetic episodes (Fig. 4). With one spike of A or B every 10 A or B lifetimes, most total AB synthesis occurs in events involving 4, 5 or 6 spikes of substrate, thereby constituting a near-ideal reactor for replication (Fig. 5). Such sporadic trains of substrate spikes are near-ideal because they both increase available nucleotide concentrations, and also ensure that A and B are available while template AB exists (Fig. 6), thereby generating net replication.

A More Precise Description Of The Darwinian Transition Previous discussion of the sporadically fed pool was conducted in terms of the requirements for net replication over time (Yarus 2012); that is, for transfer of information to descendant AB molecules during pool lifetimes, thereby permitting Darwinian evolution. Because we now know that replication near the Darwinian boundary occurs during particular substrate spike trains, prior conclusions can be restated in more explicit molecular terms.

For example, there is very strong internal selection for molecular stability in the sporadically fed pool, which, given variance in stability, will drive the pool toward replication and Darwinism (Yarus 2012). This inevitable stability selection can now be recognized as the effect of longer-surviving reactants on the assembly of effective episodes of synthesis, which necessarily require the co-survival of sparse AB, A and B.

There are other parallel clarifications, but instead of a list, I paraphrase a major earlier conclusion (in *Epistemology*; (Yarus 2012)) that includes most simpler re-statements. 'Known ribonucleotide reactions thus appear sufficient for robust progress from undirected beginnings (near zeroth order in A–B) to a system structured by Darwinian change (near first order in A–B; (Szathmáry and Maynard Smith 1997)) where efficient episodes of templating account for AB synthesis and potentially, evolution'.

A Sporadically Fed Pool Can Do More Than Asked Here The confinement of effective synthesis to a small intermediate set of templating episodes (as in Fig. 5) is informative. Two- and three-spike episodes cannot constitute ideal conditions for replication, so one expects increase in output when more spikes contribute. Large numbers of spikes in one episode are improbable because they require coincidence of a greater number of elementary events (Fig. 4), and they do not proportionately elevate AB output (Fig. 3) in any case. One thus expects a decline in total AB output for complex episodes, and therefore a peak for intermediate numbers of spikes, as observed.

The above reasoning has implications for the ultimate constructive capacity of the sporadically fed pool. Suppose it were necessary, in order to emulate an IDA, to make an RNA more complex than a self-complementary ribonucleotide. Because of the increasing probability of further substrate spikes as an episode enlarges, complex many-spike episodes are more abundant than intuition suggests. This is embodied in the long tail of increasingly complex intersections with substrate that appears rightward in Fig. 4. To take a specific example, Fig. 4 shows that if one needed ≥ 8 spikes (rather than ≥ 4 , as here) for a particular synthesis, the standard pool still might accommodate this more complex construction in $\approx 7\%$ of episodes. Above, the ≥ 4 -spike episodes that are near-ideal for AB templating occur 35% of the time. Thus, productive spike trains of ≥ 8 would be $\approx 20\%$ the frequency of $4-6$, which are optimal for AB synthesis.

There is a second factor which assists more complicated pool synthesis. In Fig. 3 (despite poorer sampling for complex episodes), the variance of AB output clearly shrinks in going

from less to more complex episodes. Thus, output from complex episodes is more reliable. If one required a certain level of product for a useful biochemical effect, for intermediate AB levels it will occur in a greater fraction of episodes ≥ 8 , than in episodes of 4-6 spikes. Complex episodes with large templated outputs (Fig. 2) are therefore even more visible to selection than their 20 % relative abundance suggests. The sporadically fed pool therefore seems readily capable of more intricate products, even perhaps hosting simple catalytic activities (Illangasekare and Yarus 2012; Yarus 2011b).

A Triumph For The Replicators In the initial description of a primordial oligonucleotide replicator (Yarus 2011a), continuity of the proposed origin of life sequence requires the emergence of functional replicators from a profoundly heterogeneous background of spontaneous oligonucleotide synthesis. Now that we have a quantitative account of such replicator emergence (Figs. 5 and 6), the original evolutionary sequence is made more probable. In particular, natural variability in the supply of precursors should not now be counted an insuperable obstacle.

The Cost Of Disorganized Conditions Figure 5 exhibits an unanticipated result: it shows that, under plausible conditions, overall output occurs mostly via a minority of near-ideal, high-yielding episodes of templated replication (compare Figs. 2, 3 and 6). These elevated yields are supported by above-average substrate concentrations and also effective templating, possible when substrate recurs in uncorrelated multi-spike trains (e.g., Fig. 6b). This striking ability of a sporadically feed pool to replicate by exploiting the 35 % of spike trains that are potentially near-ideal raises the question of the true cost of unreliable substrate supplies. Unreliable substrates are likely unavoidable under primordial conditions; what penalty does this impose?

The question has no unique quantitative answer; but I assume that the pool's role will be to supply a chemically-competent replicator (or a set of them) for the next phase of evolution. Therefore the minimal time required for this event may provide a useful index. Comparison can be phrased in terms of the time required for net replication (T_{Darwin} , in the spirit of (Yarus 2012)). A standard sporadically fed pool presented with simultaneous, constant, completely stable influxes of substrates (constant A, B, colored processes, Fig. 1) begins net replication at 0.425 lifetimes, when templated AB synthesis first exceeds direct synthesis. If A and B are not constant, but instead consumed by oligomer syntheses, T_{Darwin} is unchanged because replication occurs before consumption of significant A and B. Neither of these calculations represent a realistic primitive condition, but they serve as standards for the argument. If usual molecular decays (Fig. 1, legend) are introduced to a pool given simultaneous A and B, T_{Darwin} becomes 1.41 lifetimes, longer because substrates and reactants decay instead of engaging in replication. Thus far, times are determinate, but the sporadically fed pool is stochastic. If we take the median for T_{Darwin} of the stochastic pool (allowing now for sporadic substrate supply spikes as well as their decay), time to net templating is 166 lifetimes (median of 100 pool simulations).

Thus, using one spike of unstable substrate at random every 10 lifetimes, replication and potential selection (the Darwinian era) are delayed ≈ 400 fold with respect to synchronized, completely stable substrates. If one asks about sporadic A and B supply only (allowing decay), T_{Darwin} is delayed ≈ 120 fold in the sporadically fed pool (Fig. 1). The cost of unpredictable chemical supplies is therefore apparent, and mostly attributable to sporadic substrate arrival, but not an insuperable bar, given time.

Particularity And Generalization Quantitative kinetics of the sporadically fed pool suggest that fluctuating natural precursors unexpectedly may support the appearance of the IDA.

Fluctuation therefore unexpectedly ameliorates its own effect in making substrates unavailable. There are two points of support, the first quantitative. There are more opportunities for reaction than might be evident: a variety of sporadic events allow frequent overlap between unstable spikes and templated synthesis (Fig. 2 and 3), and disproportionate numbers of complex spike trains favor such synthesis (Fig. 4). Secondly, and qualitatively: fluctuation in arrival of uncontrolled precursors actually promotes complex reactions (Figs. 5 and 6) by allowing the recurrence of a required complicated sequence of substrate supplies, even though, by hypothesis, no particular train of spikes can be favored.

Because the initial kinetic argument (Yarus 2012) is framed in terms of stability and rates of nucleotide reactions, it is tempting to tie these conclusions to a particular set of molecules and conditions. There is some point to this; for example, the distribution of synthetic magnitudes in Fig. 2 is a result bound to the standard pool's A and B spike frequency, magnitudes and lifetimes.

However, such a limitation is not general. Even in Fig. 2, the prominence of large, rare templated events (pools that replicate) is more broadly valid. More generally, the interesting properties of the sporadically fed pool are founded on unpredictable, unstable biomolecules, and the necessity for complex reaction sequences to make complex products. Because all of these are persistent obstacles to any plausible primordial process, parallels to present conclusions seem likely to recur. George Wald's dictum about the origin of life is relevant here: "Time is the hero of the plot...One has only to wait: time itself performs the miracles." (Wald 1954). Viewed from the sporadically fed pool, time is useful, but not because it makes possible intrinsically slow reactions. Very slow reactions require very long-lived reactants, which is not an apt description of activated nucleotides. But long times can still mediate a ribonucleotide pool origin because time hosts vast numbers of trials, only one of which must succeed.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Borquez E, Cleaves HJ, Lazcano A, Miller SL (2005) An investigation of prebiotic purine synthesis from the hydrolysis of HCN polymers. *Orig Life Evol Biosph* 35(2):79–90
- Costanzo G, Saladino R, Crestini C, Ciciriello F, Di Mauro E (2007) Nucleoside phosphorylation by phosphate minerals. *J Biol Chem* 282(23):16729–16735
- Fuller WD, Sanchez RA, Orgel LE (1972a) Studies in prebiotic synthesis. VI. Synthesis of purine nucleosides. *J Mol Biol* 67(1):25–33
- Fuller WD, Sanchez RA, Orgel LE (1972b) Studies in prebiotic synthesis: VII. Solid-state synthesis of purine nucleosides. *J Mol Evol* 1(3):249–257
- Illangasekare M, Yarus M (2012) Small aminoacyl transfer centers at GU within a larger RNA. *RNA Biol* 9(1):59–66
- Kanavarioti A, Rosenbach MT, Hurley TB (1992) Nucleotides as nucleophiles: reactions of nucleotides with phosphoimidazolide activated guanosine. *Orig Life Evol Biosph* 21:199–217
- Kim HJ, Ricardo A, Illangkoon HI, Kim MJ, Carrigan MA, Frye F, Benner SA (2011) Synthesis of carbohydrates in mineral-guided prebiotic cycles. *J Am Chem Soc* 133(24):9457–9468
- Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362:709–715
- Lohrmann R (1977) Formation of Nucleoside 5'-Phosphoramidates under Potentially Pre-Biological Conditions. *J Mol Evol* 10(2):137–154
- Oro J, Kimball AP (1961) Synthesis of purines under possible primitive earth conditions. I. Adenine from hydrogen cyanide. *Arch Biochem Biophys* 94:217–227
- Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459(7244):239–242

- Sawai H, Orgel LE (1975) Oligonucleotide synthesis catalyzed by the zinc ion. *J Am Chem Soc* 97(12):3532–3533
- Szathmáry E, Maynard Smith J (1997) From replicators to reproducers: the first major transitions leading to life. *J Theor Biol* 187:555–571
- von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. *Angew Chem Int Ed Engl* 25(10):932–935
- Wald G (1954) The origin of life. *Sci Am* 191(August):44–53
- White HB III (1976) Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* 7:101–104
- Wu M, Higgs PG (2011) Comparison of the roles of nucleotide synthesis, polymerization, and recombination in the origin of autocatalytic sets of RNAs. *Astrobiology* 11(9):895–906
- Yarus M (2011a) Getting Past the RNA World: the Initial Darwinian Ancestor. In: Atkins JF, Cech TR, Gesteland RF (eds) *RNA worlds: From Life's origins to diversity in gene regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 43–50
- Yarus M (2011b) The meaning of a minuscule ribozyme. *Philos Trans R Soc Lond B Biol Sci* 366(1580):2902–2909
- Yarus M (2012) Darwinian behavior in a cold, sporadically fed pool of ribonucleotides. *Astrobiology* 12(9):870–883
- Yarus M, Caporaso JG, Knight R (2005) Origins of the genetic code: the escaped triplet theory. *Annu Rev Biochem* 74:179–198