Dynamics of prebiotic RNA reproduction illuminated by chemical game theory

Jessica A. M. Yeates[a], Christian Hilbe[b], Martin Zwick[c], Martin A. Nowak[b], and Niles Lehman[a,1]

[a]Department of Chemistry, Portland State University, Portland, OR 97207; [b]Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138; and [c]Systems Science Graduate Program, Portland State University, Portland, OR 97207

Edited by Peter Schuster, University of Vienna, Vienna, Austria, and approved March 15, 2016 (received for review December 22, 2015)

Many origins-of-life scenarios depict a situation in which there are common and potentially scarce resources needed by molecules that compete for survival and reproduction. The dynamics of RNA assembly in a complex mixture of sequences is a frequency-dependent process and mimics such scenarios. By synthesizing *Azoarcus* ribozyme genotypes that differ in their single-nucleotide interactions with other genotypes, we can create molecules that interact among each other to reproduce. Pairwise interplays between RNAs involve both cooperation and selfishness, quantifiable in a 2 × 2 payoff matrix. We show that a simple model of differential equations based on chemical kinetics accurately predicts the outcomes of these molecular competitions using simple rate inputs into these matrices. In some cases, we find that mixtures of different RNAs reproduce much better than each RNA type alone, reflecting a molecular form of reciprocal cooperation. We also demonstrate that three RNA genotypes can stably coexist in a rock–paper–scissors analog. Our experiments suggest a new type of evolutionary game dynamics, called prelife game dynamics or chemical game dynamics. These operate without template-directed replication, illustrating how small networks of RNAs could have developed and evolved in an RNA world.

RNA | prebiotic chemistry | origins of life | game theory | ribozyme

A plausible description of a sequence of events that could have led to the origins of life on the Earth from a purely chemical milieu has long been desirable, yet remains elusive. The RNA world hypothesis has helped sharpen our focus on what could have taken place 4 Gya, in that RNA serves as a powerful model for a self-sustaining chemical system capable of evolutionary change (1–6). Although this hypothesis has engendered much debate, both in its general applicability and in the details of its implementation (7–9), there are some clear emerging trends. Among the recent advances in prebiotic RNA studies is the concept of an evolving "network" of RNAs being required to kick-start life, rather than a single selfish entity. This idea dates back to the formative studies of Eigen and Schuster in the 1970s (10, 11). However, it can be sharply seen in the 20+-y effort aimed at developing a generalized RNA replicase ribozyme in the laboratory: new successes have taken advantage of a fragmentation of the best such artificial ribozyme and invoke a network of reactions to provide for its assembly (12). Our own laboratories have focused on a variety of "prelife" (13, 14) and cooperative network (15, 16) approaches to understand how evolving RNA systems could have arisen from abiotic sources of nucleotides and short oligomers. Many others have also stressed the need for distributed functionality at the onset of life, both chemically (17) and in space and time (18, 19).

To advance a network approach to the "single biomolecule problem" in the RNA world, what is needed now is an understanding of how prebiotic networks could have evolved. Auspiciously, the mechanisms of network evolution are beginning to be unraveled (20–23). For example, Aguirre et al. (23) have recently provided a framework for studying how networks can actually compete with one another. To apply this type of thinking to prebiotic RNA networks, we first need to understand how pairs and small numbers of RNAs could influence the appearance and reproduction of others. In short, we need to understand the frequency-dependent dynamics of small clusters of RNAs before we can begin to decompose the mechanisms by which complex networks of RNAs could have evolved.

In this work, we provide an empirical demonstration of frequency-dependent dynamics that take place for small (one to three) numbers of catalytic RNA genotypes that interact while reproducing. Using the covalently self-assembling *Azoarcus* ribozyme system that we had previously elucidated (15, 16, 24), and in which a complex network ecology is possible (16), we quantify and model the growth rates of single genotypes as they compete with others for reproduction using RNA source fragments. We focus on interactions among pairs and in one triplet of RNAs to ask: which chemical behaviors engender the greatest numerical payoffs to various genotypes when mixed with others? We show that the dynamics of small networks can be studied in the laboratory, realizing the line of investigation first imagined by Eigen (10).

Moreover, we demonstrate that the resulting dynamics among RNA molecules can be interpreted, and in fact predicted, using concepts from evolutionary game theory. It has been noted before that both prebiotic evolution and the evolution of biological systems may follow similar equations (10, 11, 25). Using our empirical chemical system, we make this connection explicit. The game-theoretic framework provides an additional perspective on chemical kinetics. It allows us to summarize the dynamics between different genotypes in a single payoff matrix, whose values can easily be interpreted. Using only this matrix, we can calculate the final genotypic equilibria in two- or three-molecule interactions.

Results

The Chemical System. For a prebiotic system, we used the covalently self-assembling *Azoarcus* $tRNA^{Ile}$ intron described previously (15, 24, 26). This ~200-nt ribozyme (Fig. 1*A*) can be broken into two, three, or four pieces that can spontaneously reassemble into the covalently contiguous ribozyme when incubated in a warm (48 °C) $MgCl_2$ solution (24). The assembly process is initiated through a 3-nt base-pairing interaction between two RNA fragments, and, importantly, changing these nucleotide triplets can alter the specificity

Significance

The origins of life required a means for information-containing molecules to compete with one another for survival and reproduction. Using an analysis based on game theory, we can predict the situations in which cooperation, selfishness, or a mixture of the two is beneficial to the future evolutionary success of RNAs.

Author contributions: N.L. designed research; J.A.M.Y. performed research; J.A.M.Y., C.H., M.Z., M.A.N., and N.L. analyzed data; and N.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: niles@pdx.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1525273113/-/DCSupplemental.

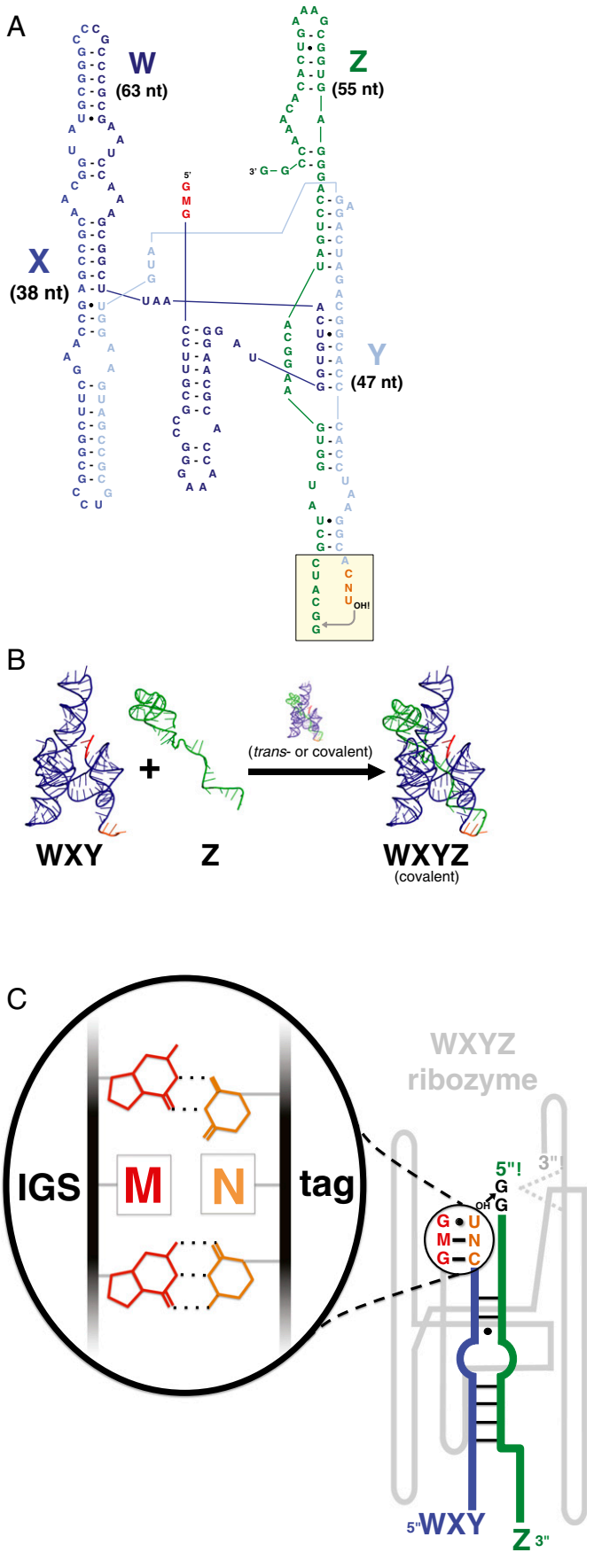


Fig. 1. Self-reproducing ribozyme system. (*A*) The *Azoarcus* ribozyme. The 148-nt **WXY** portion (blue) has an internal guide sequence (IGS) (GMG; red) on the 5′ end, and a 3-nt "tag" sequence (CNU; orange) on the 3′ end. The 55-nt **Z** portion is shown in green. Shaded box shows the *trans*-esterification reaction that occurs at the **Y**–**Z** junction. (*B*) The **WXY** + **Z** → **WXYZ** reaction. (*C*) The IGS-tag interaction determines assembly rates in the *Azoarcus* ribozyme broken into two pieces. The catalytically active ribozyme (gray) can be either a single covalently contiguous **WXYZ** molecule, or a noncovalent *trans* complex (24). Either catalyzes the formation of a covalent bond between **WXY** (blue) and **Z** (green) RNAs, guided by H bonding between the IGS (red) on the ribozyme and a tag (orange) on the **WXY** substrate.

of which RNAs react with one another (26). For simplicity, we focused on the two-piece assembly reaction, which can be symbolized as **WXY** + **Z** → **WXYZ**, where **W**, **X**, **Y**, and **Z** represent roughly 50-nt sections of the *Azoarcus* ribozyme (Fig. 1 *A* and *B*). Various genotypes of **WXY** molecules can be created by altering one of the first (5′) three nucleotides in the **W** region, corresponding to the ribozyme's internal guide sequence (IGS), and one of the last (3′) three nucleotides, corresponding to its "tag" that is recognized by a catalyst ribozyme to form a covalent bond with a **Z** fragment, creating a **WXYZ** molecule (Fig. 1*C*). We allowed fourfold variation in the middle nucleotide of both the IGS and the tag (M and N, respectively) to allow 16 possible molecular genotypes. For example, 1 of the 16 possible genotypes would be $_{GGG}$**WXY**$_{C\Delta U}$, which can be abbreviated with just the middle nucleotides: GA in this case. These genotypes could be pitted against and among each other to form various small networks in which the shared resource **Z** molecule is required to create full-length, covalently contiguous **WXYZ** molecules.

Self-Assembly. To dissect the dynamics of intragenotype and intergenotype interactions, we first compared the abilities of self-assembly among the 16 genotypes in isolation. We did this by measuring the autocatalytic rate constants (k_a) (cf. ref. 15) in **WXY** + **Z** → **WXYZ** reactions (*SI Appendix*, Fig. S1). As expected, when M and N are Watson–Crick pairs, much higher rates of self-assembly occur, but all possible pairings allow some degree of assembly (*SI Appendix*, Fig. S1). The autocatalytic rate constant is a measurement of the contribution of autocatalytic feedback to the overall self-assembly reaction (26). By doping various amounts of the fully formed autocatalyst **WXYZ**, we have previously measured values of k_a in similar *Azoarcus* ribozyme reactions (15), and here we used the same doping method (*SI Appendix*, Fig. S1) to measure it in these reactions. The efficiency of growth in self-reproducing systems (when autocatalysis is critical; e.g., prebiotic ones) is best reflected in the k_a parameter (27, 28), and thus we used this measure for all of our analyses below (*SI Appendix*).

Cross-Assembly. In an uncompartmentalized milieu, akin to a "warm little pond" scenario but extendable to other prebiotic scenarios, continuously interacting genotypes may be receiving assembly benefits from others as well as from like genotypes. Thus, our next step was to measure rates of cross-assembly. Cross-assembly has been studied in catalytic RNAs before, for example, in the case of two possible genotypes in a self-ligating ribozyme system (29). In our study, there are 120 possible pairwise interactions among dissimilar genotypes, and the reaction is by *trans*-esterification (i.e., recombination) rather than by ligation. We measured assembly rate constants when one genotype interacts with a different genotype in the same tube for 0–30 min. These rate constants are the dynamical variables in a setting when two genotypes compete for the shared resource **Z**. To do this, we tracked the amounts and proportions of each **WXYZ** genotype over time using differential ^{32}P labeling of the 5′ ends of the **W**-containing fragments (Fig. 2 *A* and *B*). By combining results from self- and cross-assemblies, we could now compile the four types of intramolecular and intermolecular events that could occur when two genotypes interact. These can be displayed in a 2 × 2 matrix that identifies the components of molecular "fitness" in a prebiotic competition. Although we did not measure all possible pairwise two-genotype interactions, we chose a

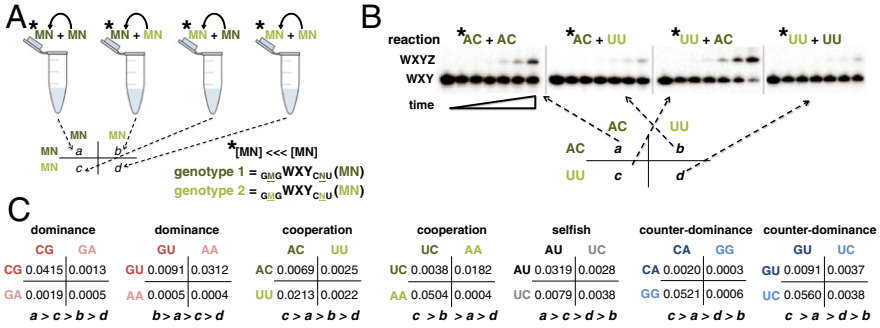


Fig. 2. Single-round competitions between two **WXY** genotypes. (*A*) Differential ^{32}P-labeling method to separately obtain *a*, *b*, *c*, and *d* values (autocatalytic rate constants: k_a, in units of minutes^{-1}) in 2 × 2 matrices. The 5′-^{32}P•**WXY** RNA is a small (<<0.1%) dopant in 1 μM unlabeled **WXY** RNA plus 1 μM **Z**. Values *a* and *d* were obtained as in *SI Appendix*, Fig. S1, whereas *b* and *c* were obtained by doping genotype 1 into genotype 2. Asterisks (*) denote ^{32}P-labeled RNAs. (*B*) Example gel used for raw data. (*C*) Empirical matrices compiled from k_a values for seven selected competitions.

few that would include competitions between both rapid and slow self-assembling RNAs. The data for seven representative matrices are given in Fig. 2*C*.

Serial Dilution Experiments. Having now both self- and cross-assembly rates for a single 30-min bout of competition for reproduction, we could compare their values and thus predict what would transpire when two genotypes of differing prelife fitnesses were allowed to compete iteratively over time in an evolutionary setting (30). We also hoped to be able to devise an analogous method with the potential to predict results from the myriad three-genotype interactions, and so on (see below). For the two-genotype experiments, we designed a serial-dilution technique in which a pair of **WXY** genotypes are mixed at some ratio, typically 1:1, provided **Z**, and then reacted for a brief period (5 min). At this time, when RNA production is still in exponential growth, we transferred a small fraction (10%) to a new reaction vessel in which new raw materials were present (Fig. 3*A*). In the receiving tube, we provided more unreacted **WXY** of each genotype, plus fresh **Z** and buffer. This technique was pioneered by Sol Spiegelman and coworkers (31) and has been used in many in vitro molecular evolution experiments with RNA (16, 32). We tracked the amounts and proportions of each **WXYZ** genotype over eight transfers using differential ^{32}P labeling. This allowed us to quantify the chemical equivalent of evolutionary success across generations ("bursts" of RNA assembly).

We pitted the seven pairs of **WXY** RNAs studied above against each other in two-genotype contests (Fig. 3*B*). Among these seven cases, we observed situations where one genotype clearly dominates, and cases in which coexistence of the two genotypes is attained after three to four bursts (Fig. 3*C*). In at least one case of the latter situation (AU vs. UC), we varied the genotype ratio across a broad range of values but always observed similar final steady-state frequencies reached by the two genotypes (*SI Appendix*, Fig. S2). Note that "extinction" is not possible in such serial dilution scenarios because fresh material is added each burst (31). However, the serial dilution format is prebiotically relevant in that it simulates a periodically replenished pool, as in wet–dry cycles.

Modeling Chemical Dynamics. In parallel with the experimental results, we created ordinary differential equation (ODE) models of this system to visualize more clearly the dynamics of the genotypic assembly. We first developed a simple model in which the frequencies of two competing RNA types were tracked in a flow reactor setting that is a continuous analog of the serial dilution experiments. In this model, the frequency changes of the two strategies over time ($\dot{x}$ and $\dot{y}$) are described by the following:

$$\dot{x} = ax + by - \phi x \; ; \; \dot{y} = cx + dy - \phi y. \qquad [1]$$

Here, *a*, *b*, *c*, and *d* are the rate constants of self-assembly (*a* and *d*) and cross-assembly (*b* and *c*), as visualized in a 2 × 2 matrix of possibilities when two genotypes interact (Fig. 2*C*). The death (or dilution) term, $\phi = (a+c)x + (b+d)y$, guarantees that $\dot{x} + \dot{y} = 0$ and $x + y = 1$. This parameterization is appropriate because the reaction rate is a linear function of RNA abundances, and because we maintained RNA assembly in its exponential growth phase across transfers (*SI Appendix*, Fig. S1). The unique equilibrium values, $\hat{x}$ and $\hat{y}$, for each competition are given by the following:

$$\hat{x} = \frac{a - 2b - d + \sqrt{(a-d)^2 + 4bc}}{2(a + c - b - d)}, \qquad [2]$$

and $\hat{y} = 1 - \hat{x}$ (*SI Appendix*). This model closely predicted the qualitative outcomes of the serial-dilution competitions (Fig. 3). With the *a*, *b*, *c*, and *d* values obtained empirically from Fig. 2 entered into the model, experimental data and model outcomes match in all cases, both qualitatively and quantitatively (compare Fig. 3 *C* and *D*). We could also predict the cross-assembly values from only the self-assembly values, and Fig. 3*E* shows that this technique still gives agreement between data and model. We explore this more in *SI Appendix*, Fig. S3, and in *Discussion*.

Game-Theoretic Treatment. The ODE model based on chemical kinetics suggests a new type of evolutionary game theory. Game theory is a field that was first developed to study strategic and economic decisions among humans (33, 34). It later found its way into biology in the form of evolutionary game theory (35, 36). There, fitness depends on the frequency of different strategies (or phenotypes) in the population. The classical equation of evolutionary game theory is the so-called replicator equation (e.g., ref. 29): $\dot{x}_i = x_i[f_i(\vec{x}) - \phi(\vec{x})]$, where x_i is the frequency of genotype *i*, $f_i(\vec{x})$ is the fitness of this genotype, and $\phi(\vec{x})$ is the average fitness of all genotypes. This describes a frequency-dependent replication rate. In contrast, in our system, there is no replication but rather frequency-dependent assembly.

To extend a game-theoretic treatment to an abiotic situation, we realized a parallel between the 2 × 2 matrix that exists to describe components of fitness (Fig. 2*A*) and a game-theoretic payoff matrix. In the latter, each matrix entry is the payoff to the row genotype when interacting with the column genotype. Importantly, the evolving entities need not be rational agents for a game-theoretic analysis to have explanatory power (30), and thus could be applied to a molecular system. In fact, there have been at least two recent predictions that game theory could be useful in the interpretation of biochemical behavior (37, 38), and the *Azoarcus* system in particular was singled out as a good candidate (37). Game theory has been proposed to be manifest at the chemical level (39–41), but this has never been shown empirically. We thus sought a practical demonstration that this could be the case, reasoning that game theory could augment our ODE analysis by offering a simple fitness-based explanation of how selection could choose, say, molecular cooperation.

Because there are four values in a 2 × 2 payoff matrix (Fig. 2), and, with the assumption that at the chemical level no two of these could be exactly the same, there are 24 possible strict ordinal rankings of these values (e.g., $a > b > c > d$) (30). Additionally, we can assume that $a > d$ without loss of generality (otherwise, one only needs to relabel the genotypes), lowering the number of possible outcomes that could result from an iterative

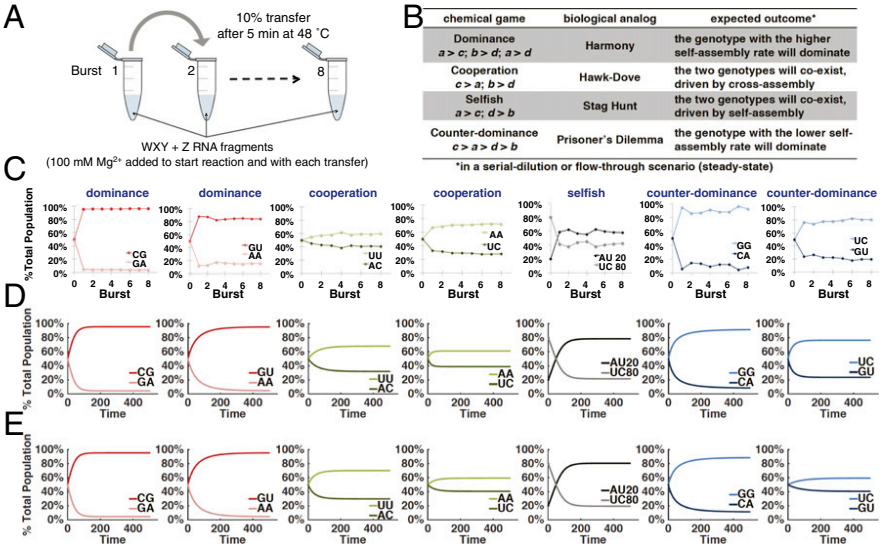


Fig. 3. Serial-dilution experiments for two-genotype competitions. (*A*) Schematic of a serial-dilution experiment. (*B*) Classes of two-genotype (two-strategy) interactions in game theory. (*C*) Plots of relative frequencies of **WXYZ** genotypes as a function of time (bursts) in the serial-dilution format for the same seven competitions described in Fig. 2*C*. For the AU vs. UC competition, results from using skewed (AU:UC::20:80) genotype frequencies are shown; other ratios converge to the same qualitative result with AU > UC (*SI Appendix*, Fig. S2). (*D*) Predicted dynamics of the genotypes in these same competitions based on a simple ODE model in a flow reactor scenario using measured cross-assembly rates (Fig. 2*C*). (*E*) Modeling results using estimated cross-assembly rates in the 2 × 2 matrix (see text).

two-genotype interaction to 12. Based on groups of payoff matrix inequalities, we divided these outcomes into four categories (Fig. 3*B*) that will have evolutionary significance based on analogies with biological systems (30, 42), and we assigned names to various scenarios of experimental outcomes that we observed (Figs. 2 and 3). In the "Dominance" scenario, given by $a > c$ and $b > d$, one genotype is expected eventually to dominate in frequency, in this case the genotype with the higher self-assembly rate (k_a). In the "Cooperation" scenario, $c > a$ and $b > d$, such that cross-assembly will always exceed self-assembly, and hence the population will adopt a mixture of the two genotypes. In the "Selfish" scenario, $a > c$ and $d > b$, such that self-assembly will always exceed cross-assembly, meaning that a coexistence mixture will also result, but for the opposite mechanism than in the Cooperation scenario. Finally, in the "Counter-dominance" scenario, $c > a > d > b$, the genotype with the lower self-assembly rate is, counterintuitively, expected to dominate in frequency. These four outcomes have rough parallels in the biological games (Fig. 3*B*). For example, a game with the payoffs of the Counter-dominance scenario can be interpreted as a prisoner's dilemma (PD). In evolutionary biology, the PD is often taken as the baseline model for situations in which a group-beneficial trait—expressed by the $a > d$ inequality—is selected against at the individual level because $a < c$ and $d > b$. Similarly, the payoff configuration of the Selfish scenario corresponds to the so-called stag hunt game, in which a trait is only successful if it is common (as for example, when members of a group need to decide whether to join a stag hunt). Similar biological interpretations can be given for the other two classes (Fig. 3*B*) (30, 35, 36).

Among our two-genotype RNA competitions, we observed examples of all four of the categories described above (Figs. 2*C* and 3*C*). When CG is pitted against GA for example, CG dominates because it can assemble itself far better than can GA (*SI Appendix*, Fig. S1). The interactions between these two molecules are weak: the middle nucleotide of the IGS of one genotype does not pair well with the middle nucleotide of the other genotype's tag. Self-assembly is the major determinant in this competition, leading to a Dominance outcome, because $a >> c$ and $b >> d$. However, when we pitted CA against GG, the latter (GG) dominates despite its more than threefold worse self-assembly rate constant. This matchup is thus an example of the Counter-dominance scenario, in which a nonintuitive result emerges: in isolation, CA self-assembles far more robustly than GG (*SI Appendix*, Fig. S1), but when in competition with GG, this CA genotype is greatly outperformed for assembly. In this case, the interaction of CA with GG is very strong. The middle nucleotide of the IGS in CA forms a Watson–Crick interaction with the middle nucleotide of the tag of GG. Thus, the distinction between the "cooperator" (CA) and the parasitic genotype, or "defector" (GG), becomes clear, as in a classical PD. The PD has been biologically demonstrated in viruses (43) and yeast (44, 45). However, to our knowledge, our data are the first example of it manifesting at the pure molecular level, in which a genotype with a lower self-assembly rate can become predominant. Similar phenomena may explain the evolution of other biochemical functions such as single-turnover ("suicide") enzymes, e.g., methyltransferases used in DNA repair.

We also observed examples of the other two categories of two-genotype competitions, those that lead to coexistence (Fig. 3*C*). When we pitted AC against UU, both genotypes persisted at high frequency (>40%) stably over time in a Cooperation outcome. Here, both self-assembly rates are expected to be moderate, along with one of the cross-assembly rates (UU → AC), whereas the other cross-assembly rate is strong (AC → UU). This leads to a situation where cross-assembly is generally more effective than self-assembly, with the consequence that each genotype predominantly assembles the other: a chemical analog to simultaneous reciprocal altruism. The result is that both genotypes are assembled to substantial frequencies. When we pitted AU against UC, genotypes with the same aggregate nucleotides as the AC vs. UU competition, again coexistence of both genotypes eventually resulted (Fig. 3*C*). However, the route to this result differed from that in the Cooperation scenario. In AU vs. UC, self-assembly is generally more effective than cross-assembly, such that the major dynamical determinant is each genotype doing the same, selfish, action of self-assembly. Thus, this contest is an example of a Selfish outcome. However, unlike the biological stag hunt game, where one expects bistability depending on starting ratios, different initial frequencies of the two chemical genotypes led to the same general outcome (*SI Appendix*, Fig. S2), highlighting a distinction between the biological replicator dynamics (35), and the chemical dynamics of our system. In a biological setting, a stag hunt scenario leads to the extinction of one the two strategies depending on the initial frequencies. However, in our chemical setting, we would expect a mixed population to result because the continual replenishment of genotypes in the serial dilution protocol prevents extinction, and we observed similar final frequencies when we varied genotype ratios in a Selfish scenario (*SI Appendix*, Fig. S2). Using only self-assembly data to estimate all four values in the payoff matrix (*SI Appendix*, Fig. S3), our ODE model could forecast what would result for many possible two-genotype contests (*SI Appendix*, Fig. S4). Generally, in a two-genotype contest, both genotypes will reach similar frequencies if $a + b$ and $c + d$ are approximately equal. This is less likely to happen in, for example, a Dominance scenario than in a Selfish one.

Rock–Paper–Scissors Competition. We were also able to manifest with RNA a well-known scenario with three genotypes (or "strategies";

Discussion), namely rock–paper–scissors (RPS) (Fig. 4). Inspired by the children's game of the same name, RPS describes scenarios with a cyclical arrangement of dominance relationships; these have previously been identified in nature (46, 47). A well-known example involves the bacterium *Escherichia coli*, for which there is an evolutionary sequence from the wild type, to mutants that produce toxin together with an immunity protein, to mutants that only produce the immunity protein, and back to the wild type again (47). We were interested to see whether a similar scenario could also occur among molecular genotypes. RPS is one of ~50 qualitatively different three-strategy game outcomes (48). It is a contest in which three genotypes will all attain substantial frequencies jointly despite the fact that the equilibrium frequency of any one would be far lower than another without the presence of the third. Based on our expectations from the results of other two-genotype contests, we chose three **WXY** genotypes—AA, UC, and GU—that we anticipated could generate an RPS game. In isolation, we would predict that AA beats UC, UC beats GU, and GU beats AA. When we pitted these genotypes against each other two-at-a-time in a serial-dilution format, we indeed saw that one genotype reaches at least 70% superiority in each game (Fig. 3*C*). However, when we pitted all three against each other in the same reaction vessel in a serial-dilution format, their joint frequencies quickly attained values near 30–40% and remained there (Fig. 4*A*). Notably the steady-state frequencies of each genotype in the RPS scenario were distinctly higher than their "losing" values in two-genotype games (Fig. 3*C* vs. Fig. 4*A*), and they appeared to converge on an internal point in a simplex plot. Again, using the simple ODE model described above, we were able to predict correctly these outcomes. To do this, we created a 3 × 3 payoff matrix (*SI Appendix*, Fig. S6) derived from the results of the three two-genotype payoff matrices (Fig. 2*C*). From this matrix, we calculated that there should be a stable internal equilibrium point (cf. ref. 36) consisting of 39% UC, 35% AA, and 26% GU, that agreed qualitatively with the empirical data in the three-strategy serial-dilution experiment (Fig. 4*B*).

Discussion

We have characterized the dynamics that occur when two, and in one case three, RNA genotypes compete for reproduction using a common resource. By representing possible interactions in a 2 × 2 matrix, it is possible to predict rather accurately, based solely on a few lower-level data points, what would happen in various competition scenarios. This holds up qualitatively to the three-genotype interaction level and may extend beyond that. Likely, the complexity of the system and unpredictable interactions such as nonproductive binding events (16) will begin to play more important roles in larger networks. Prior efforts have focused on either simulation models of similar RNA network dynamics (e.g., refs. 11, 13, 37, and 49–51), or on broad-scale experimental data (e.g., refs. 12, 16, 29, 52, and 53). Here, we demonstrate that experiment and modeling can agree. Future work is now possible to understand how small networks can evolve into larger ones, and to extend these methods to other, simpler RNA systems.

This matrix approach has predictive power. To highlight this fact, we obtained 16 pieces of empirical data, which correspond to the diagonal values in the 2 × 2 matrix, i.e., the rate enhancement an RNA gets when interacting with its own genotype: *a* or *d*. For forecasting purposes, we can use these rate constants to estimate the off-diagonal terms in the matrix: *b* and *c*. Specifically, we used the corresponding nucleotide–nucleotide pairs from self-assembly to predict what would happen in cross-assembly (*SI Appendix*, Fig. S3). Then we input these four values into the ODE model to compare its outputs with the results from the serial dilution experiments. In Fig. 3*E*, the modeled dynamics are shown given the estimated values of *b* and *c* using this strategy (*SI Appendix*, Fig. S3). The outcomes using either measured or estimated *b* and *c* values both emulate the experimental results (Fig. 3*C*) almost perfectly. Such strong agreement of model and data underscores the utility of the 2 × 2 matrix approach to forecast complex dynamics from lower-dimensional data—the four values *a*, *b*, *c*, and *d*—and portend the ability to extend this type of analysis to three (or even more) interacting members of a prebiotic RNA network, as in the RPS scenario (Fig. 4). Thus, we demonstrate that self-assembly data alone can allow estimates of steady-state genotype frequencies; measurements of cross-assembly rates help refine these estimates but are not necessary for a qualitative prediction of dynamics. The static game theory tables organize the parameters that determine the dynamics into diagonal and off-diagonal terms that have clear meanings (autocatalytic and cross-catalytic).

Our experimental system is particularly amenable to a game-theoretic interpretation. The assembly rates of **WXYZ** molecules can be considered "payoffs" because **WXYZ**s are covalently contiguous RNAs that would represent successful genotypes in an evolutionary contest, their rates of production being dependent on the current environment of other **WXYZ** molecules. New **WXYZ** molecules being produced are analogous to progeny in a biological setting. "Strategies" in a contest at the chemical level—molecular reproduction strategies—would be the phenotypes displayed by the genotypes: in our case, their abilities to catalyze the assembly of **WXYZ** RNAs. When provided with the **Z** molecule resource, these genotypes can assemble into **WXYZ** molecules, which then can catalyze even faster production of themselves because the covalently contiguous **WXYZ** has a roughly twofold higher catalytic activity than a noncovalent (*trans*) complex **WXY**-**Z** (16). Such phenotypes would be a combination of self-assembly, where a **WXY** drives the assembly of a **WXYZ** molecule of like genotype, and cross-assembly of other genotypes. In a molecular system, the genotype directly determines the phenotype, and hence the strategy, whose payoff is its fitness.

Our analyses lead to some subtle but important distinctions between what we describe here and classical evolutionary game theory (35, 36). Biological systems reproduce via template-directed polymerization, characterized by the replicator equation. In our chemical system, there is no replication per se. It is more de novo synthesis (reproduction rather than replication) that could be described as prelife (13, 14) in which the replicator equation does not apply. Here, Eq. **1** applies instead. The *Azoarcus* ribozyme system reproduces through a recombination of fragments rather than a polymerization of nucleotides (24). Specifically, the information that differentiates "self" from "nonself" is embodied in the thermodynamics of only a single nucleotide pair (Fig. 1*C*). Thus, there are only a few chemical moieties on the base-pairing surface of these nucleotides that influence the alternative strategies in a game-theoretic sense, rather than a large continuum of genotypes that would be available in a true biological system. However, a key facet of the dynamics of frequency changes of molecular genotypes is autocatalytic feedback, meaning that the strategies—such as intergenotype cooperation (54)—used by

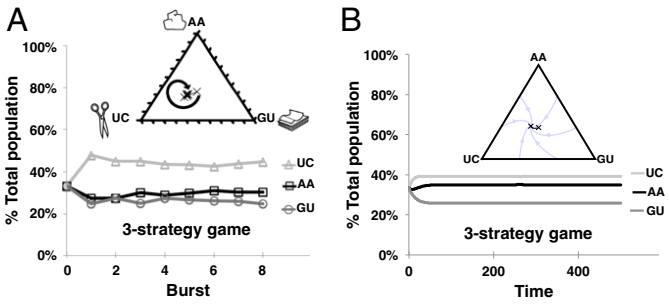


Fig. 4. A RPS competition among three **WXY** genotypes. (*A*) Empirical data from a serial-dilution experiment. The experiment was performed in the same fashion as in Fig. 3, except here three genotypes (AA, UC, and GU) were separately tracked. Each pair of two-genotype competitions is expected to give a clear winner (Fig. 3*C*); but here coexistence results. (*B*) Predictions from the ODE model. The simplex plots show (× symbols) the joint frequencies starting from the center and approaching a stable internal equilibrium point.

molecules in creating more molecules are optimized by kinetic selective forces (55). Explicit differences in kinetics lead to cases (e.g., the Selfish scenario) where chemical and biological game theory lead to qualitatively distinct outcomes. This can also be seen in the RPS scenario, where, although the results predicted from the replicator equation and from our chemical dynamics lead to qualitatively similar results (a stable interior equilibrium point), a small deviation in the quantitative results can be detected (*SI Appendix*, Fig. S7). Such variance reveals that, in prelife game dynamics, the specific outcomes of intergenotypic competitions can be predicted in a fashion parallel to, but not identical to, those used to calculate Nash equilibria (34) in classical games (*SI Appendix*). Thus, a game-theoretic approach gives us an appreciation of the chemical ecology of how reproduction (production from the environment) could evolve before biological replication. Knowing the mechanics behind the interactions among two and three genotypes, it should now be possible to predict how larger RNA networks could have evolved.

Materials and Methods

RNA Self-Assembly. Reactions, containing **WXY** (1 μM), **Z** (1 μM), **WXYZ** (0–2 μM), and/or ^{32}P-labeled **WXY** (≤0.003 μM), were initiated with the addition of buffer (100 mM $MgCl_2$ and 30 mM EPPS, pH 7.5). Time point samples were drawn and immediately quenched at 0.5–30 min, and the **WXY** and **WXYZ** RNAs were separated by 8% (mass/vol) polyacrylamide/8 M urea gel electrophoresis. For serial-transfer experiments, after 5 min, 10% of each reaction was transferred to a new tube with fresh reagents, over eight transfers.

Autocatalytic Rate Constants. Initial rates were calculated from slopes of the linear portion of plots of the product ratio versus time; for fast reactions this was ≤5 min and for slower reactions ≤10 min. The rate constant (k_a) was calculated from the slope of the initial **WXYZ** concentration versus the initial rate of the reaction (*SI Appendix*, Fig. S1) (15). For cross-assembly reactions, the **WXY** (1 μM) and **WXYZ** (0–2 μM) were of one genotype and ^{32}P-labeled **WXY** (≤0.003 μM) was a second. Values are averages of three separate trials. Each rank order of *a*, *b*, *c*, and *d* values was significant ($P < 0.05$; *SI Appendix*, Fig. S8).

Mathematical Modeling. A dynamical ODE model was constructed assuming a flow-reactor scenario in which the frequency changes of two competing strategies were described by Eq. **1** above. This was extended to a competition among three or more strategies in a similar fashion (*SI Appendix*).

ACKNOWLEDGMENTS. We thank I. Chen, J. Fletcher, E. Hayden, and P. Nghe for discussions. Funding was from NASA Grant NNX14-AK21G (to N.L.).

1. Rich A (1962) On the problems of evolution and biochemical information transfer. *Horizons in Biochemistry*, eds Kasha M, Pullman B (Academic, New York, NY), pp 103–126.
2. Woese CR (1967) *The Genetic Code: The Molecular Basis for Genetic Expression* (Harper and Row, New York).
3. Crick FHC (1968) The origin of the genetic code. *J Mol Biol* 38(3):367–379.
4. Orgel LE (1968) Evolution of the genetic apparatus. *J Mol Biol* 38(3):381–393.
5. Joyce GF (2002) The antiquity of RNA-based evolution. *Nature* 418(6894):214–221.
6. Neveu M, Kim H-J, Benner SA (2013) The "strong" RNA world hypothesis: Fifty years old. *Astrobiology* 13(4):391–403.
7. Shapiro R (1987) *Origins: A Skeptic's Guide to the Creation of Life on Earth* (Bantam New Age, New York).
8. Carter CW, Jr (2014) Urzymology: Experimental access to a key transition in the appearance of enzymes. *J Biol Chem* 289(44):30213–30220.
9. Wächtershäuser G (2014) The place of RNA in the origin and early evolution of the genetic machinery. *Life (Basel)* 4(4):1050–1091.
10. Eigen M (1971) Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58(10):465–523.
11. Eigen M, Schuster P (1977) The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften* 64(11):541–565.
12. Mutschler H, Wochner A, Holliger P (2015) Freeze-thaw cycles as drivers of complex ribozyme assembly. *Nat Chem* 7(6):502–508.
13. Nowak MA, Ohtsuki H (2008) Prevolutionary dynamics and the origin of evolution. *Proc Natl Acad Sci USA* 105(39):14924–14927.
14. Chen IA, Nowak MA (2012) From prelife to life: How chemical kinetics become evolutionary dynamics. *Acc Chem Res* 45(12):2088–2096.
15. Hayden EJ, von Kiedrowski G, Lehman N (2008) Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. *Angew Chem Int Ed Engl* 47(44):8424–8428.
16. Vaidya N, et al. (2012) Spontaneous network formation among cooperative RNA replicators. *Nature* 491(7422):72–77.
17. Cafferty BJ, et al. (2013) Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J Am Chem Soc* 135(7):2447–2450.
18. Kauffman SA (1993) *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford Univ Press, Oxford, UK).
19. Benner SA, Kim H-J, Carrigan MA (2012) Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. *Acc Chem Res* 45(12):2025–2034.
20. Nghe P, et al. (2015) Prebiotic network evolution: Six key parameters. *Mol Biosyst* 11(12):3206–3217.
21. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512.
22. Albert R, Jeong H, Barabási A-L (2000) Error and attack tolerance of complex networks. *Nature* 406(6794):378–382.
23. Aguirre J, Papo D, Buldú JM (2013) Successful strategies for competing networks. *Nat Phys* 9(4):230–234.
24. Hayden EJ, Lehman N (2006) Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem Biol* 13(8):909–918.
25. Hofbauer J, Schuster P, Sigmund K (1979) A note on evolutionary stable strategies and game dynamics. *J Theor Biol* 81(3):609–612.
26. Draper WE, Hayden EJ, Lehman N (2008) Mechanisms of covalent self-assembly of the *Azoarcus* ribozyme from four fragment oligonucleotides. *Nucleic Acids Res* 36(2):520–531.
27. von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. *Angew Chem Int Ed Engl* 25(10):932–935.
28. von Kiedrowski G, Wlotzka B, Helbing J, Matzen M, Jordan S (1991) Parabolic growth of a self-replicating hexadeoxynucleotide bearing a 3′-5′-phosphoamidate linkage. *Angew Chem Int Ed Engl* 30(4):423–426.
29. Kim D-E, Joyce GF (2004) Cross-catalytic replication of an RNA ligase ribozyme. *Chem Biol* 11(11):1505–1512.
30. Nowak MA (2006) *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press of Harvard, Cambridge, MA).
31. Mills DR, Peterson RL, Spiegelman S (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc Natl Acad Sci USA* 58(1):217–224.
32. Wright MC, Joyce GF (1997) Continuous in vitro evolution of catalytic function. *Science* 276(5312):614–617.
33. von Neumann J, Morgenstern O (1944) *Theory of Games and Economic Behavior* (Princeton Univ Press, Princeton).
34. Nash JF (1950) Equilibrium points in n-person games. *Proc Natl Acad Sci USA* 36(1): 48–49.
35. Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics* (Cambridge Univ Press, Cambridge, UK).
36. Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. *Science* 303(5659):793–799.
37. Bohl K, et al. (2014) Evolutionary game theory: Molecules as players. *Mol Biosyst* 10(12):3066–3074.
38. Liao D, Tlsty TD (2014) Evolutionary game theory for physical and biological scientists. I. Training and validating population dynamics equations. *Interface Focus* 4(4):20140037.
39. Bartel H-G (1984) Zur Anwendung der Spieltheorie in der Physikalischen Chemie. *Z Phys Chem* 265:1186–1192.
40. Hart S (2005) An interview with Robert Aumann. *Macroecon Dyn* 9:673–740.
41. Shuster S, Kreft J-U, Schroeter A, Pfeiffer T (2008) Use of game-theoretic models in biochemistry and biophysics. *J Biophys* 34:1–17.
42. Maynard Smith J (1982) *Evolution and the Theory of Games* (Cambridge Univ Press, Cambridge, UK).
43. Turner PE, Chao L (1999) Prisoner's dilemma in an RNA virus. *Nature* 398(6726):441–443.
44. Frick T, Schuster S (2003) An example of the prisoner's dilemma in biochemistry. *Naturwissenschaften* 90(7):327–331.
45. Van Dyken JD, Müller MJI, Mack KML, Desai MM (2013) Spatial population expansion promotes the evolution of cooperation in an experimental Prisoner's dilemma. *Curr Biol* 23(10):919–923.
46. Sinervo B, Lively CM (1996) The rock–paper–scissors game and the evolution of alternative male strategies. *Nature* 340:240–243.
47. Kerr B, Riley MA, Feldman MW, Bohannan BJM (2002) Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature* 418(6894):171–174.
48. Bomze IM (1995) Lotka-Volterra equation and replicator dynamics: New issues in classification. *Biol Cybern* 72:447–453.
49. Wu M, Higgs PG (2009) Origin of self-replicating biopolymers: Autocatalytic feedback can jump-start the RNA world. *J Mol Evol* 69(5):541–554.
50. Vasas V, Fernando C, Santos M, Kauffman S, Szathmáry E (2012) Evolution before genes. *Biol Direct* 7:1, discussion 1.
51. Czárán T, Könnyű B, Szathmáry E (2015) Metabolically coupled replicator systems: Overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J Theor Biol* 381:39–54.
52. Striggles JC, Martin MB, Schmidt FJ (2006) Frequency of RNA-RNA interaction in a model of the RNA world. *RNA* 12(3):353–359.
53. Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. *Science* 323(5918):1229–1232.
54. Higgs PG, Lehman N (2015) The RNA world: Molecular cooperation at the origins of life. *Nat Rev Genet* 16(1):7–17.
55. Pross A (2011) Toward a general theory of evolution: Extending Darwinian theory to inanimate matter. *J Syst Chem* 2:1.

Supporting Information

Dynamics of prebiotic RNA reproduction illuminated by chemical game theory

J.A.M. Yeates, C. Hilbe, M. Zwick, M.A. Nowak, and N. Lehman

Contents:

Full Methods

Experimental Studies

Mathematical Modeling

Discussion of the utility of game theory at the chemical level

Supplementary Figure 1 – Calculation of self-assembly autocatalytic rate constants

Supplementary Figure 2 – Varying intial frequencies in the AU *vs*. UC contest

Supplementary Figure 3 – Use of IGS and tag nucleotides to predict 2-strategy outcomes

Supplementary Figure 4 – All 72 2-strategy results from predicted self-assembly matrices

Supplementary Figure 5 – Exclusion of contests with equal values in the payoff matrix

Supplementary Figure 6 – The predicted 3x3 payoff matrix for the RPS scenario

Supplementary Figure 7 – Comparison of biological and chemical game dynamics

Supplementary Figure 8 – Statistical tests for ordinal rankings

References for Supporting Information

Full Methods

Experimental Studies

RNA preparation

The **WXY** and **WXYZ** molecules, which are portions of the self-splicing group I intron from the isoleucine pre-tRNA in the purple bactierium *Azoarcus* (1), were prepared by *in-vitro* transcription from DNA plasmid templates. The **Z** fragment was purchased from TriLink Biotechnologies (San Diego, CA) and was gel purified prior to use. All RNAs were resuspended in 1–10 μM solutions in 0.1 mM EDTA. For quantification, < 0.1 μM of the **WXY**-fragment was 5′-labeled with γ[^{32}P]•ATP using OptiKinase (USB, Cleveland, OH).

Self-assembly kinetics (Quadrants* a *and* d *of the payoff matrix)

See Fig. 2 for a schematic of this process. Reaction mixtures containing **WXY** (1 μM), **Z** (1 μM), **WXYZ** (0 – 2 μM), and ^{32}P-labeled **WXY** (≤ 0.003 μM), all of the same IGS and tag genotype, were heated to 80 ˚C for 2 minutes then cooled to 48 ˚C. Time "zero" aliquots were drawn and quenched with equivolume quench solution (125 mM EDTA and 2X loading dye containing formamide and bromophenol blue). Reactions were initiated with the addition of reaction buffer (100 mM $MgCl_2$ and 30 mM EPPS, pH 7.5). Time point samples were drawn and immediately quenched with quench solution at 0.5, 1.0, 2.0, 5.0, 10, and 30 minutes. Samples were loaded on an 8% polyacrylamide/ 8M urea gel and **WXY** and **WXYZ** bands were separated. Visualization and quantification was possible *via* phosphorimaging on a Typhoon Trio+ variable mode phosphorimager (GE Healthcare) and accompanying ImageQuant software (GE Healthcare). A product ratio was calculated by comparing the RNA in the product **WXYZ** band to the unreacted **WXY** band (% reacted = [reacted / (reacted + unreacted)]*100%). Kinetic values were calculated as previously described (2). Briefly, initial rates were calculated from the slope of the linear portion of the reaction curve from a plot of the product ratio versus time (total of $n = 3$ trials for each concentration). For fast reactions this was ≤ 5 minutes and for slower reactions ≤ 10 minutes. The rate constant (k_a) was calculated from the slope of the initial **WXYZ** concentration versus the initial rate of the reaction (Fig. S1).

Kinetic parameter justification

The kinetic parameters of a ribozyme form of the *Azoarcus* group I intron have been studied by Kuo *et al.*, who demonstrated that the chemical step of *trans*-esterification was the rate-limiting process in this ribozyme (3). Previously we demonstrated that the full-length ribozyme **WXYZ**, can be broken into two (or three or four) fragments that could spontaneously self-assemble (2, 4) Self-assembly occurs when fragments hybridize through base-pairing (2˚) and tertiary (3˚) interactions to form non-covalent "*trans*" complexes of the ribozyme. Once formed, *trans* complexes can catalyze recombination reactions on other hybridized fragments (*i.e.*, **WXY** + **Z**) to synthesize covalent versions of the ribozyme. Recognition between a catalytic complex, either a *trans* complex or a fully covalently-contiguous ribozyme, is driven by the strength of the IGS-tag base pairing, as shown in Table 1 in Fig. S1 below. For example, the wildtype IGS in the *Azoarcus* ribozyme is 5′–GUG–3′, which *in vivo* matches with the pseudocomplement 5′–CAU–3′ as a consequence of a requisite G-U wobble preceding the splice site. A catalytic event creates a covalent closure of the stem-loop, often with the release of one or two G nucleotides from the 5′ end of the **Z** molecule GGCAU (ref. 24); thus this is a recombination reaction (5). In a population of molecules, this reaction is autocatalytic because the product (**WXYZ**) is a ribozyme that has an approximately 2-fold higher k_{cat} than the *trans* complex (2, 6).

In our system, RNA genotypes assemble one another from background material *via* kinetics that are driven by first order (or pseudo-first order) reactions; the units of the autocatalytic rate constants are per minute. In cases where molecules are being covalently formed from their own fragments, the initial rate has both autocatalytic (k_a) and non-autocatalytic (k_b) contributions: $(d[\mathbf{WXYZ}]/dt)_i = k_a[\mathbf{WXYZ}]^p + k_b$, where p represents a variable reaction order that must be experimentally determined (7)). We have previously shown that these self-assembly reactions display a high degree of autocatalytic rate enhancements, with autocatalytic efficiencies (the ratio of the slope to the y-intercept in the plots in Fig. S1 below) near the high end of such reported values (2). Although the exact value of the order (p) of the autocatalytic reaction is not straightforward, our modeling in ref. 2 suggests that a first-order ($p \sim 1$) fits the kinetic data very well, leading to units of min^{-1} in Table 1 below (Fig S1).

2-strategy kinetics (Quadrants* b *and* c *of the payoff matrix)
See Figs. 2*A*, 2*B*, and 2*C* for a schematic of this process. Reactions, visualization and quantitation were performed same as the self-assembly kinetics with the exception of the fragment genotypes. For the *b* quadrant, the initial mixture contained 1 μM of the player 2 genotype and a trace amount of ^{32}P-labeled **WXY** (≤ 0.003 μM) of the player 1 genotype. For the *c* quadrant, the initial mixture contained 1 μM of the player 1 genotype and a trace amount of ^{32}P-labeled **WXY** (≤ 0.003 μM) of the player 2 genotype.

Serial dilutions
A master mix reaction mixture was formed containing equimolar **WXY** genotype 1 (0.5 μM) and **WXY** genotype 2 (0.5 μM) and **Z** (1.0 μM). The mixture was then divided in two equal volumes. One part was doped with ^{32}P-labeled **WXY** genotype 1 and the other part with ^{32}P-labeled **WXY** genotype 2. The two reaction mixtures were then aliquoted into eight tubes each (one for each burst). (In the case of AU *vs*. UC, the original master mix made for the eight bursts was made at a 20:80, 50:50, or 05:95 ratio, and then divided into eight portions, and then this was used as above. See Fig. S2.) All tubes were heated up to 80 ˚C for 2 minutes and then cooled to 48 ˚C. The reaction in the first tube was initiated with the addition of reaction buffer (100 mM $MgCl_2$ and 30 mM EPPS, pH 7.5). At 5 minutes, 10% of the solution volume from tube #1 was transferred to tube #2, and tube #1 was placed on ice. Reaction buffer was immediately added to tube #2 while tube #1 was subsequently quenched with equal volume of quench solution. The transfer protocol was repeated through eight bursts. The two-part master mix containing ^{32}P-labeled **WXY** was used as a negative control for the assay. Gel separation, visualization and quantitation was performed same as kinetic assays. Three-strategy serial dilutions (Fig. 4*A*) were performed using the same protocol as above with the addition of a third genotype.

Mathematical modeling

Derivation of the kinetic equation

To describe the dynamics of the serial dilution experiment, we derive a simple ODE model. We consider a contest with two strategies, *A* and *B*. The payoffs are given by the matrix:

	A	*B*
A	*a*	*b*
B	*c*	*d*

In our context all entries are positive: $a, b, c, d > 0$.

We use the following variables: x_0 is the concentration of the *A* precursor (**WXY**), y_0 is the concentration of the *B* precursor (**WXY**), z denotes the concentration of the **Z** molecule, x_1 is the complex formed between the A precursor and Z, y_1 is the complex formed between the B precursor and **Z**. The variables x and y denote the concentration of the *A* and *B* molecule (**WXYZ**), respectively.

In the dilution experiment, the two precursor molecules and the **Z** molecule are provided at constant level, and the complex is formed in a reversible chemical reaction

$x_0 + z \rightleftharpoons x_1$
$y_0 + z \rightleftharpoons y_1$

Thus, x_1 and y_1 are also provided at constant level, and they give rise to the respective **WXYZ** molecule according to the catalyzed reactions

$x_1 \overset{ax}{\rightarrow} x$
$x_1 \overset{by}{\rightarrow} x$
$y_1 \overset{cx}{\rightarrow} y$
$y_1 \overset{dy}{\rightarrow} y$

Therefore, the kinetic equation (that would occur in a flow reactor) is

[3]
$$\dot{x} = (ax + by)x_1 - \phi x$$
$$\dot{y} = (cx + dy)y_1 - \phi y,$$

Here ϕ is a parameter chosen such that the concentration of *A* and *B* is constant, $x + y = C$ and $\dot{x} + \dot{y} = 0$. Without loss of generality, we can set C=1 (we only need to replace the variables x and y by the transformed variables x/C and y/C). In that case, ϕ can be calculated as

$\phi = (ax + by)x_1 + (cx + dy)y_1.$

Moreover, since x_1 and y_1 are provided at equal concentrations, we may set $x_1 = y_1 = 1$. This may lead to a change of the time scale, but it leaves the trajectories of Eq. **[3]** unchanged. Thus, the dynamical equation simplifies to

[1]
$$\dot{x} = ax + by - \phi x$$
$$\dot{y} = cx + dy - \phi y,$$

with $\phi = (a + c)x + (b + d)y$.

Properties of the dynamical equation

In the following, we list a few interesting properties of the dynamical equation.

Property 1.
The dynamical equation **[1]** has a unique and globally stable equilibrium $(\hat{x}, \hat{y})$ with $0 < \hat{x}, \hat{y} < 1$ and $\hat{x} + \hat{y} = 1$. In the generic case that $a + c \neq b + d$, the equilibrium frequency of $\hat{x}$ is given by

[2] $$\hat{x} = \frac{a-2b-d+\sqrt{(a-d)^2+4bc}}{2(a+c-b-d)}.$$

Otherwise, if $a + c = b + d$, the equilibrium frequency is $\hat{x} = b/(b + c)$.

Proof. Since $x + y = 1$, the first equation in **[1]** can be written as

$$\dot{x} = ax + b(1 - x) - \phi x = b + (a - 2b - d)x - (a + c - b - d)x^2 =: f(x).$$

The function $f(x)$ has the unique zero $\hat{x}$ in the unit interval (0,1). Since $f(0) = b > 0$, it follows that $f(x) > 0$ when $0 \leq x < \hat{x}$, whereas $f(x) < 0$ when $\hat{x} < x \leq 1$. Therefore, for any given initial frequency x, orbits converge towards $\hat{x}$. □

Let us next explore how the position of the equilibrium is affected by the entries of the payoff matrix. The following results follow directly from Eq. **[2]**.

Property 2.
1. The equilibrium frequency $\hat{x}$ is strictly increasing in a and b, and strictly decreasing in c and d.
2. All other parameters unchanged, $a \to \infty$ or $b \to \infty$ implies $\hat{x} \to 1$, whereas $c \to \infty$ or $d \to \infty$ implies $\hat{x} \to 0$.
3. The equilibrium frequency satisfies $\hat{x} = 1/2$ if and only if $a + b = c + d$. Similarly, it satisfies $\hat{x} > 1/2$ if and only if $a + b > c + d$.

As a consequence of the previous result, we can also draw the following connection between the equilibrium frequency and the type of game considered.

Property 3.
1. In a Dominance or Counter-Dominance scenario, $\hat{x} > 1/2$ if and only if it is A that dominates B.
2. In a Selfish scenario, $\hat{x} > 1/2$ if and only if A risk-dominates B.
3. In a Cooperation scenario, $\hat{x} > 1/2$ if and only if A is also played with higher frequency in the symmetric Nash equilibrium.

Dynamical equation for $n \times n$ contests

To describe the dynamics of the rock-paper-scissors contest, we generalize the previous dynamical equation to arbitrary $n \times n$ contests. Let $M = (m_{ij})$ be the payoff matrix of such a contest, and let $x = (x_1, \cdots, x_n)^T$ be the vector that gives the frequency of each **WXYZ** molecule, such that $x_1 + \cdots + x_n = 1$. Then the n-strategy analogue of Eq. **[1]** is

[4] $$\dot{x} = Mx - \phi x,$$

with $\phi = \sum_{i,j} m_{ij} x_j$. Equation **[4]** can be considered as a slightly generalized version of the quasi-species equation. It has a unique fixed point in the interior of the state space, which is globally stable. The fixed point can be found by solving the eigenvector problem $Mx = \lambda x$, where λ is the largest eigenvalue of M. The theorem of Perron and Frobenius for positive matrices guarantees that the corresponding normalized eigenvector x is unique, and that all entries of x are positive.

In the following, let us summarize a few simple properties of the dynamical equation **[4].**

Property 4.

1. The unit simplex $\Delta = \{x \in \mathbb{R}^n \mid x \geq 0, x_1 + \cdots + x_n = 1\}$ is invariant under the dynamical equation **[4]**; that is, if the initial state $x(0) \in \Delta$ then $x(t) \in \Delta$ for all times *t*.
2. The edges of the unit simplex are not invariant under the dynamics in **[4]**; if $x_i = 0$ then $\dot{x}_i > 0$.

Proof. The sum of all entries of x does not change over time, due to our choice of ϕ. Moreover, if $x_i = 0$, then it follows from Eq. **[4]** that $\dot{x}_i = (Mx)_i - 0 > 0$.

The previous result points to an important difference between the dynamical equation **[4]** and replicator dynamics (see also Fig. S7). Replicator dynamics is *non-innovative* – if a strategy is initially absent, then the evolutionary dynamics does not introduce this strategy at some later time point. In contrast, the kinetic dynamics described in **[4]** predicts that absent **WXYZ** strategies are introduced immediately, due to the catalytic effect of the other **WXYZ** molecules (provided that the required precursor **WXY** for the absent strategy is available).

Next, let us describe the relationship between the Nash equilibria of a contest with payoff matrix *M*, and the unique equilibrium of the kinetic equation **[4]**.

Property 5.
Let $M = (m_{ij})$ be the payoff matrix of an *n*-strategy contest, and let $x = (x_1, \cdots, x_n)^T$ be the unique equilibrium of Eq. **[4]**. Then the following are equivalent:

1. x is a Nash equilibrium.
2. The equilibrium is in the center of the simplex, $x = (1/n, \cdots, 1/n)^T$.
3. The row sums of *M* coincide, $m_{i1} + \cdots + m_{in} = m_{j1} + \cdots + m_{jn}$ for all $1 \leq i, j \leq n$.

Proof.

$1 \Rightarrow 2$. As x is a Nash equilibrium in the interior of the state space, all strategies yield the same expected payoff,

$$(Mx)_i = (Mx)_j \text{ for all } 1 \leq i, j \leq n.$$

Moreover, since x is the fixed point of Eq. **[4]**, and hence the eigenvector of M corresponding to some real eigenvalue $\lambda > 0$, it follows that

$$\lambda \cdot x_i = (Mx)_i = (Mx)_j = \lambda \cdot x_j.$$

In particular, $x_i = x_j$ for all $1 \leq i, j \leq n$, and thus $x = (1/n, \cdots, 1/n)^T$.

$2 \Rightarrow 3$. As $Mx = \lambda x$ for some $\lambda > 0$ and for $x = (1/n, \cdots, 1/n)^T$, it follows that

$$\frac{1}{n}(m_{i1} + \cdots + m_{in}) = \lambda/n \text{ for all } 1 \leq i \leq n.$$

Since the right hand side does not depend on *i*, neither does the left hand side. Therefore,

$$m_{i1} + \cdots + m_{in} = m_{j1} + \cdots + m_{jn} \text{ for all } 1 \leq i, j \leq n.$$

$3 \Rightarrow 1$. If the row sums of M coincide, it is easy to check that $x = (1/n, \cdots, 1/n)^T$ is the unique fixed point of Eq. **[4]**. In this fixed point, $(Mx)_i = (Mx)_j$ for all $1 \leq i, j \leq n$, and hence x is a Nash equilibrium. □

The previous result shows that in general, the kinetic equilibrium of Eq. **[4]** is not a Nash equilibrium – the only exception occurs when all rows of the payoff matrix sum up to the same value.

Discussion of the utility of game theory at the chemical level

Given the concordance between our experimental results and the ODE models that we constructed, the game-theoretic analysis appears to be a natural consequence of chemical kinetics, but kinetics viewed in an entirely new way. Our formalism allows us to summarize the dynamics between two genotypes in a single matrix, whose values can easily be interpreted. One only needs to know this matrix to calculate genotype equilibria. The dynamics can be interpreted using frequency-dependent selection. Each step along this analytical process is performed exactly as in evolutionary game theory, although the applied equilibrium concepts differ.

Put another way, game theory adds to the ability to assess rapidly the evolutionary outcomes of contests among pre-biotic genotypes such as RNA. Specifically it can allow one to understand how two *general* outcomes arise (dominance *vs*. co-existence), but more importantly, which *specific* mechanism is operational in any given situation: auto- and/or cross-catalysis. The Counter-dominance situation is a good example of this. An ODE analysis on its own would tell you that one genotype would rise to a high frequency at the expense of the other, but the specific mechanism for this – the receipt of greater cross-assembly benefits than self-assembly benefits – may be obscured. The 2x2 payoff matrix reveals this dynamic quickly and allows insight into the precise molecular events that underlie the evolutionary dynamics. The Prisoner's Dilemma has been biologically demonstrated in viruses (8) and yeast (9, 10), and we now demonstrate it at the raw (bio)chemical level.

In many environments envisaged for plausible conditions for the origin of life, there is a steady-state flux of resources or energy. Examples include, but are not limited to, thermal gradients leading to thermophoresis (11, 12), streams flowing downhill absorbing leaching chemical precursors (13, 14), and sporadically-fed aqueous pools (15). In these situations, nascent reproducing molecules would have had to compete for common resources, and our molecular experiments and model were designed to capture these characteristics, targeting a "pre-Darwinian" description of molecular evolution (*e.g*., 16, 17). The chemical game theoretic treatment allows an extention of the benefits of traditional evolutionary game theory down to a simpler, pre-life level.

With these parallels, we propose that game theory is applicable to the events leading to the chemical origins of life. Intermolecular interactions, driven mainly by non-covalent bonding strengths, can be hypothesized as giving rise to a build-up of network complexity among prebiotic polymers as has been discussed for other RNAs (18), proteins (19), and lipids (20). Here we have provided an empirical example of how game-theoretic patterns can be manifest in a chemical system, and one that has prebiotic relevance. While the dynamics of chemical reproducers can be described in terms of classical ODE equations, the game-theoretic construct gives insight as to what happens when one goes from a small number of nodes in a network to a larger number of nodes.

Although one needs to be wary of over-interpretation, game theory at the chemical level allows for a clear evolutionary perspective of prebiotic dynamics. From a practical standpoint, what we have shown is that from knowing the influence on catalysis of a single nucleotide-pair interaction – here the middle nucleotide of the IGS-tag triplet, akin to the middle position in a codon-anticodon pairing – on catalysis, one could predict general outcomes from contests among a small number of competitors vying for a shared resource. Note that in the predicted Cooperation competitions – those leading to significant steady-state frequencies of both genotypes – none of the participating genotypes has a Watson-Crick pair between their M and N nucleotides. The presence of a non-canonical nucleotide pair in the IGS-tag recognition process leads to the situation where the off-diagonal terms in the payoff matrix (*b* and *c*) are greater than the diagonal terms (*a* and *d*), thereby promoting a type of molecular cooperation (21), in that molecules are forced kinetically to forego some of their reproductive potential to assemble others (22). In the converse case, the Selfish scenario, the diagonal terms in the matrix exceed the off-diagonal terms, leading also to co-existence, but one achieved in the opposite manner. (The CG *vs*. AU game is perhaps the most extreme example of a Selfish scenario, and while the CG self-assembly rate constant exceeds that of AU by only 30%, we predict that the steady-state frequency of CG should reach 96% even with a continual supply of equimolar AU; see Fig. S4.)

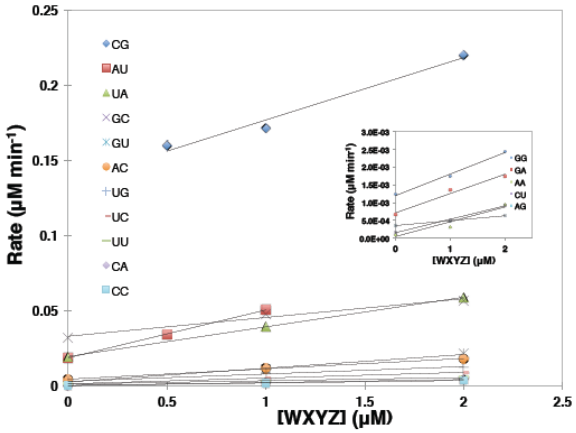
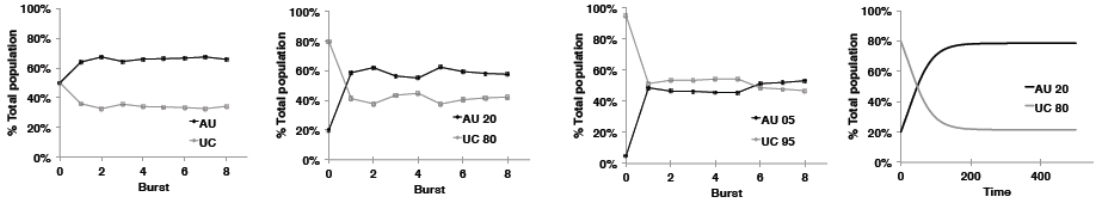


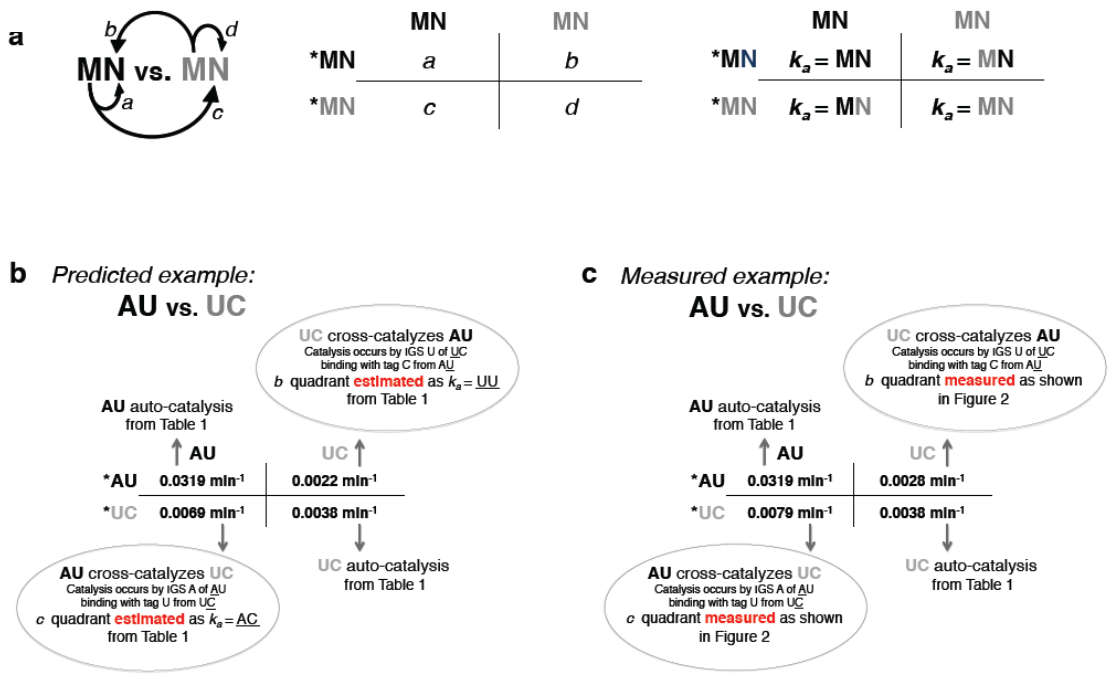
Table 1. Rate constants, k_a (min^{-1}), for the 16 genotype variations of WXY.

Genotype	k_a (min^{-1})	Std. error	r^2
CG	0.0415	0.0066	0.98
AU	0.0319	0.0011	1.00
UA	0.0197	0.0004	1.00
GC	0.0125	0.0021	0.97
GU	0.0091	0.0007	0.99
AC	0.0069	0.0002	1.00
UG	0.0049	0.0004	0.99
UC	0.0038	0.0002	1.00
UU	0.0022	0.0001	1.00
CA	0.0020	0.0000	1.00
CC	0.0016	0.0001	1.00
GG	0.0006	0.0001	0.99
GA	0.0005	0.0001	0.98
AA	0.0004	0.0001	0.92
CU	0.0004	0.0000	1.00
AG	0.0001	0.0000	0.99

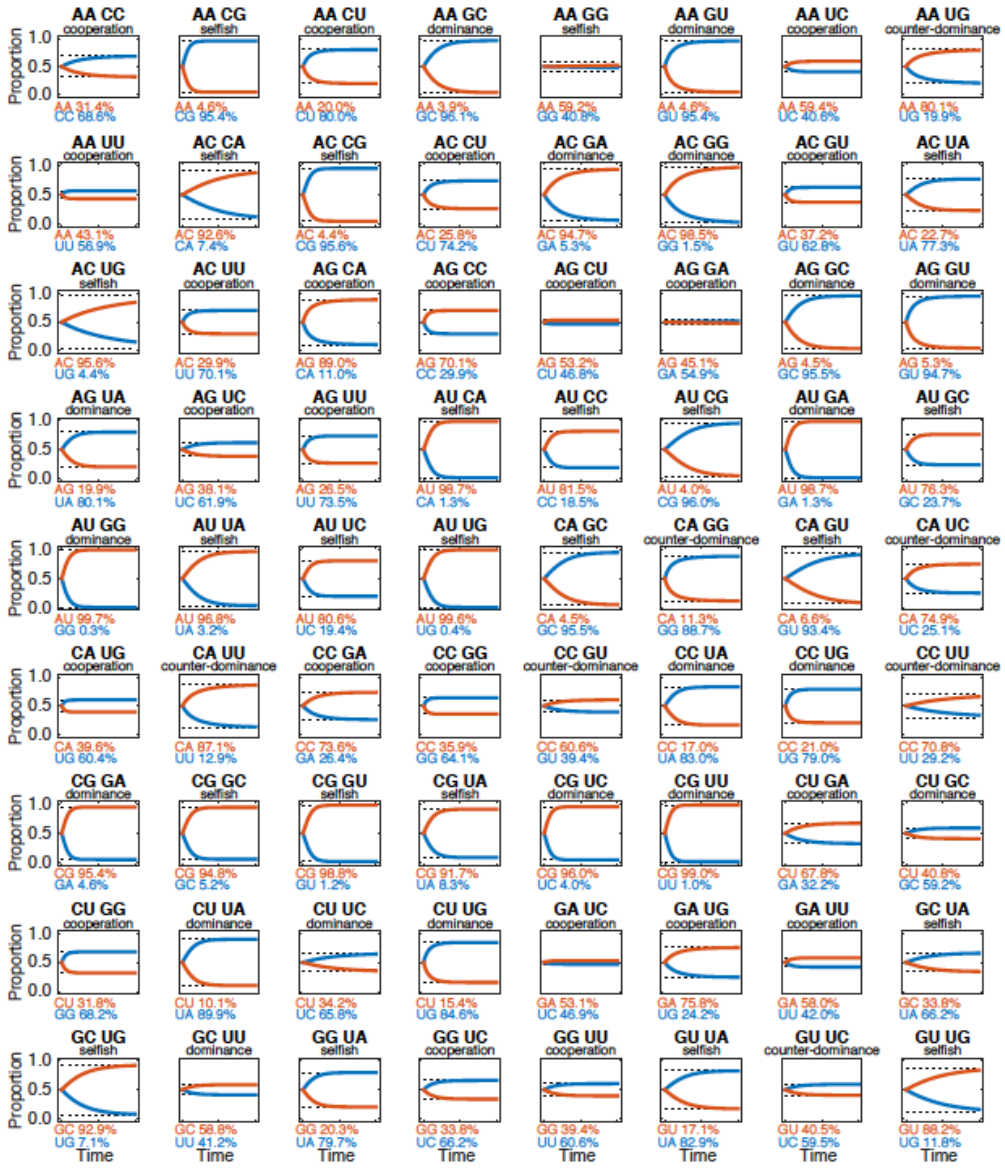
Supplementary Figure 1. Calculation of self-assembly autocatalytic rate constants. Left: raw data for the rates of self-assembly for the 16 genotypes $_{GMG}$**WXY**$_{CNU}$, where M and N are free to vary. Rates were measured by doping in 0 μM, 0.5 μM, 1 μM, or 2 μM full-length **WXYZ** into reactions containing 1 μM $_{GMG}$**WXY**$_{CNU}$ and 1 μM **Z** as described in ref. (2) and the Methods above. Each point represents the average of three independent trials. Main plot: data for the 11 fastest self-assembling genotypes, where rates were measured for reaction times of 5 minutes or less. Inset: data for the five slowest self-assembling genotypes, where rates were measured for reaction times of 10 minutes or less. Right: Table (with r^2 linear regression values included) that depicts the computed autocatalytic rate constants (k_a) via the means described above (2), based on the method of von Kiedrowski (7).



Supplementary Figure 2. Variation of the initial frequencies in the AU *vs*. UC contest. From left to right, the initial **WXY** molar frequencies are AU:UC:50:50 (experimental), AU:UC:20:80 (experimental), AU:UC:5:95 (experimental), and AU:UC:20:80 (model). All experiments converge on similar final equilibrium concentrations, with AU > UC (but not exclusionary), as predicted by the model. Note, in this 2-strategy contest, the chemical outcome (Selfish) is an analog of the Stag Hunt biological game scenario. In biological evolutionary competitions, the Stag Hunt scenario is a bi-stable one, in which the final outcome is heavily dependent on the initial conditions, with a definable tipping point. Such bi-stability is not observed in these experiments, highlighting a key difference between biological games based on replication and chemical games based on reproduction (assembly). See main text for more discussion.



Supplementary Figure 3. Use of IGS and tag nucleotides to predict 2-strategy outcomes. Here the logic for the construction of the predicted 2x2 payoff matrices is explained. (**a**) A depiction of how the MN notation translates into the *a*, *b*, *c*, and *d* values in the payoff matrix. The *a* and *d* values (diagonal terms) are data taken from the empirical self-assembly experiments tabulated in the Table in Fig S1. (**b**) and (**c**) An example game of AU *vs*. UC is shown. The *a* value is the self-assembly autocatalytic rate constant that results when $_{GAG}$**WXY**$_{CUU}$ is incubated with **Z**. This reaction is determined by the strength of a A-U nucleotide pair within the interaction between an IGS triplet in one RNA fragment and the "tag" in another (Fig. 1*C*). Likewise the *d* value is the self-assembly autocatalytic rate constant that results when $_{GUG}$**WXY**$_{CCU}$ is incubated with **Z**. On the other hand, the *b* and *c* values (off-diagonal terms) can be derived in two different manners. They can either be *estimated* using the appropriate nucleotide-pair value from the Table in Fig. S1 (panel **c**), or they can be *measured* using an experimental competition between two **WXY** genotypes as shown in Fig. 2 (panel **b**). Using *estimated* values of *b* and *c*, we have predicted the outcomes of all 72 contests in which no values are predicted to be equal (Fig. S4). For example, the AU *vs*. UC contest is predicted to give a Selfish outcome by this method, as shown. In all panels, the asterisk (*) here denotes that in a 2-genotype interaction, the payoff is to the row player in competition with the column player, and using the differential ^{32}P-labeling technique (Fig. 2A), we can track the rate of assembly of single genotypes in a mixture.



Supplementary Figure 4. Plots of outcomes of all 72 strict 2-strategy contests from predicted self-assembly matrices. By eliminating cases with equal payoffs (*e.g.*, AU *vs*. AC; Fig. S5), there are 72 possible such contests. The predictions are that 26% of the scenarios should lead to the Dominance, 32% to the Cooperation, 32% to the Selfish, and 10% to the Counter-dominance outcomes (Fig. S4). It is important to realize that the Dominance and Counter-dominance scenarios occasionally lead to fairly high frequencies of the "losing" genotype (*e.g.*, 30% UU, in CC *vs*. UU; Fig. S4). Conversely the two co-existence scenarios can sometimes lead to the near fixation of one genotype (*e.g.*, CG in CG *vs*. GU). These cases are the exceptions to the rule however, and are a consequence of the steady-state nature of both the experiments and modeling where the Z resource is continually replenished. Each panel shows the expected 2-strategy dynamics, for all 72 payoff matrices that can be derived from the self-assembly data in the Table in Fig. S1. For details on how these matrices were obtained, see Fig. 2*B* and Fig. S3. These matrices were then used as our input to calculate the expected dynamics according to the kinetic equation **[1]** above. As this figure suggests, different 2-strategy contests can have remarkably different dynamics, even if the contests are taken from the same game class. The predicted equilibrium frequencies according to equation **[2]** above are depicted by the dotted black lines. Exact equilibrium frequencies are provided at the bottom of each panel.

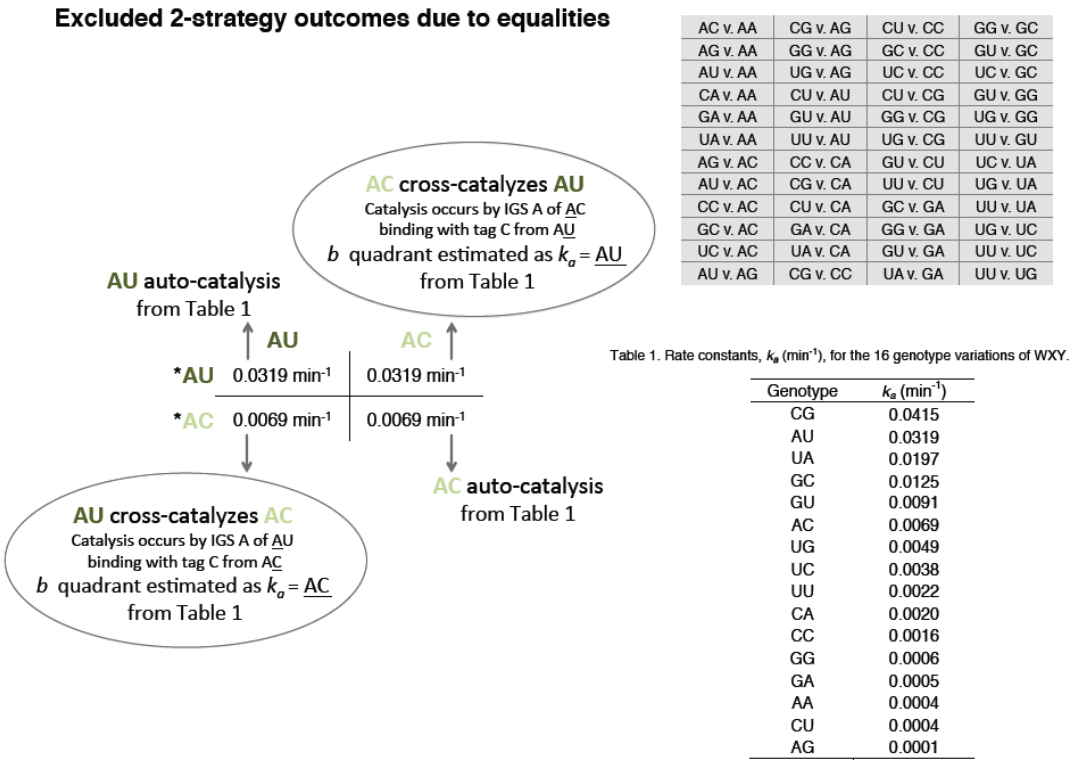


Table 1. Rate constants, k_a (min^{-1}), for the 16 genotype variations of WXY.

Genotype	k_a (min^{-1})
CG	0.0415
AU	0.0319
UA	0.0197
GC	0.0125
GU	0.0091
AC	0.0069
UG	0.0049
UC	0.0038
UU	0.0022
CA	0.0020
CC	0.0016
GG	0.0006
GA	0.0005
AA	0.0004
CU	0.0004
AG	0.0001

Supplementary Figure 5. Tabluation of the 48 2-genotype contests that generate predicted equal values in their 2x2 payoff matrices. There are 16(15)/2 = 120 distinct 2-strategy contests from 16 **WXY** genotypes, but 48 of these will generate values in the predicted payoff matrix that are *equal*, using the logic shown here. When two genotypes compete with the same M or N nucleotide, then the self-assembly values (Table) will predict two identical values in the matrix. For example, when AU competes *vs*. AC as shown here, the *a* and *b* values in the matrix both would derive from the AU self-assembly rate constant value in the Table, which is 0.0319 min^{-1}. The *a* value represents an A-U pairing during the catalytic self-assembly of ${}_{GAG}\mathbf{WXY}_{CUU}$ and **Z**, while the *b* value represents an A-U pairing during the catalytic cross-assembly of ${}_{GAG}\mathbf{WXY}_{CUU}$ and **Z** by a ${}_{GAG}\mathbf{WXY}_{CCU}$-containing ribozyme. Likewise in this particular contest, the *c* and *d* values would be predicted to be equal. Equal values in the payoff matrix do not lead to strict equilibria. Thus for the purposes of forecasting outcomes from all possible 2-strategy contests using only self-assembly data, these 48 cases are excluded, leaving 72 possible contests, as shown in Fig. S4. Again, the asterisk (*) here denotes that in a 2-genotype interaction, the payoff is to the row player in competition with the column player, and using the differential ^{32}P-labeling technique (Fig. 2A), we can track the rate of assembly of single genotypes in a mixture.

a **3 x 3 experimental matrix**

	UC	AA	GU
UC	0.0038	0.0182	0.0560
AA	0.0504	0.0004	0.0005
GU	0.0037	0.0312	0.0091

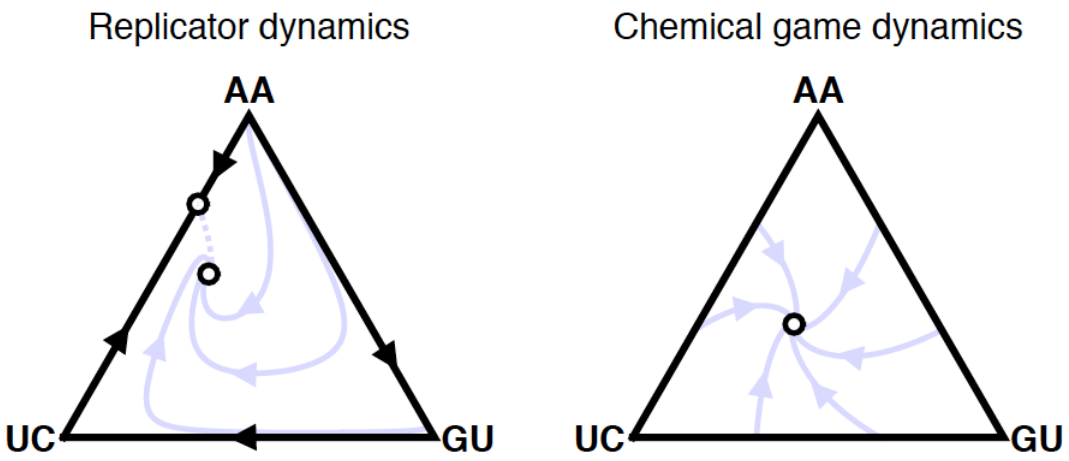
b **Three 2 x 2 experimental matrices**

	UC	AA	
UC	0.0038	0.0182	***c > b > a > d***
AA	0.0504	0.0004	**cooperation**

	GU	AA	
GU	0.0091	0.0312	***b > a > c > d***
AA	0.0005	0.0004	**dominance**

	GU	UC	
GU	0.0091	0.0037	***c > a > d > b***
UC	0.0560	0.0038	**counter-dominance**

Supplementary Figure 6. Predicted 3x3 payoff matrix in the 3-strategy contest between UC, AA, and GU (Fig. 4). **(a)** The 3x3 matrix, compiled from two sources (in analogy to Fig. 2*B*). The diagonal terms (0.0038, 0.0004, and 0.0091) were self-assembly autocatalytic rate constants from the Table in Fig. S1. The off-diagonal terms were collected from the appropriate off-diagonal terms measured in individual 2-strategy contests shown in panel (**b**). For example, the value 0.0504 (first column, second row) in the 3x3 matrix derives from the measured value of *c* in the 2x2 matrix for the UC *vs*. AA contest. A stable interior equilibrium point is predicted to exist (23, 24). This was in fact seen in the experiment (Fig. 4*A*), and the scenario is thus a chemical form of Rock-Paper-Scissors (RPS). Because one pairwise interaction (UC *vs*. AA) in this context results not in pure dominance, but Cooperation, the present contest is technically a weak form of a RPS, in the sense that every strategy can invade the previous strategy, and it can be invaded by the next strategy (see Fig. S7). In our system with only 16 possible genotypes that are roughly binary in their interactions with other genotypes (Watson-Crick base pairing in the M-N interaction or not), it is not possible to construct a strictly strong RPS scenario, because at least one pairwise interaction will not be a dominance scenario. With this matrix analysis, one can see a more general connection to network evolution. The values in the payoff matrix are equivalent to the weights of the connections in a network, and the sum of the weighted paths in a network gives an approximation of the dominant eigenvector x_i (the steady-state frequencies of each of *i* species in a network). See reference (25) for more discussion of this point.



Supplementary Figure 7. Comparison of biological (replicator) dynamics with the chemical game dynamics. Based on the 3x3 experimental matrix displayed in Fig. S6, the two graphs show the resulting replicator dynamics (left) and the dynamics according to the kinetic equation **[4]** above (right). Replicator dynamics predicts cyclical behavior: AA can be invaded by GU, GU can be invaded by UC, and UC in turn can be invaded by AA. The fixed point on the edge between UC and AA is unstable, and all orbits spiral towards the unique Nash equilibrium in the interior of the state space (with equilibrium proportions UC 35.8%, AA 50.4%, GU 13.8%). However, in the kinetic equilibrium (right), AA is considerably less abundant than predicted by the Nash equilibrium (UC 39.1%, AA 35.0%, GU 25.8%). Moreover, in the right graph, the edges of the Simplex are no longer invariant because absent strategies are introduced continually (as described in detail in the *Mathematical Modeling* section above).

Kruskal-Wallis test

Game	H	Critical value	P value
dominance CG vs. GA	7.821	6.897436	0.046
dominance GU vs. AA	9.359	6.897436	0.025
cooperation AC vs. UU	8.504	8.435897	0.009
cooperation UC vs. AA	10.385	6.897436	0.016
selfish AU vs. UC	10.385	8.435897	0.009
counter-dominance CA vs. GG	9.667	8.435897	0.009
counter-dominance GU vs. UC	8.197	6.897436	0.042

Supplementary Figure 8. Statistical tests for ordinal rankings. The Kruskal-Wallis non-parametric test was used to test the statistical significance of the rank order of the values in the 2x2 payoff matrices. The seven 2-strategy contests depicted in Fig. 2 were analyzed in this fashion, using the variation in empirical data provided by the three independent replicates of the experiments. The test statistic *H* was computed as per Sokal and Rohlf (26) and compared to the critical values provided by Meyer and Seaman (27) from the exact probability distribution. All rank-order values for these games were statistically significant at the $P < 0.05$ value or better.

References for Supplementary Information

1. Reinhold-Hurek B, Shub DA (1992) Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* 357:173–176.
2. Hayden EJ, von Kiedrowski G, Lehman N (2008) Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. *Angew Chem Int Ed* 47:8424–8428.
3. Kuo LY, Davidson LA, Pico S (1999) Characterization of the *Azoarcus* ribozyme: tight binding to guanosine and substrate by an unusually small group I ribozyme *Biochim Biophys Acta* 1489:281–292.
4. Hayden EJ, Lehman N (2006) Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem Biol* 13:909–918.
5. Riley CA, Lehman N (2003) Generalized RNA-directed recombination of RNA. *Chem Biol* 10:1233–1243.
6. Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N (2012) Spontaneous network formation among cooperative RNA replicators. *Nature* 491:72–77.
7. von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. *Angew Chem Int Ed Engl* 25:932–935.
8. Turner PE, Chao L (1999) Prisoner's dilemma in an RNA virus. *Nature* 398:441–443.
9. Frick T, Schuster S (2003) An example of the prisoner's dilemma in biochemistry. *Naturwissenshaften* 90:327–331.
10. Van Dyken JD, Müller MJI, Mack KML, Desai MM (2013) Spatial population expansion promotes the evolution of cooperation in an experimental Prisoner's Dilemma. *Curr Biol* 23:919–923.
11. Baaske P, Weinert FM, Duhr S, Lemke KH, Russell MJ, Braun D (2007) Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc Natl Acad Sci USA* 104:9346–9351.
12. Mast CB, Schink S, Gerland U, Braun D (2013) Escalation of polymerization in a thermal gradient. *Proc Natl Acad Sci USA* 110:8030–8035.
13. Benner SA, Kim H-J, Carrigan MA (2012) Asphalt, water, and the prebiotic synthesis of ribose, ribonucleotides, and RNA. *Acc Chem Res* 45:2025–2034.
14. Patel BH, Percivalle C, Ritson DJ, Duffy CD, Sutherland JD (2015) Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chemistry* 7:301–307.
15. Yarus M (2012) Darwinian behavior in a cold, sporadically fed pool of ribonucleotides. *Astrobiology* 12:870–883.
16. Nowak MA, Ohtsuki H (2008) Prevolutionary dynamics and the origin of evolution. *Proc Natl Acad Sci USA* 105:14924–14927.
17. Arnoldt H, Strogatz SH, Timme M (2015) Switching between distributed species state in a simple model of early life. *Phys Rev E* 92:052909.
18. Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. *Science* 323:1229–1232.
19. Ashkenasy G, Jagasia R, Yadav M, Ghadiri MR (2004) Design of a directed molecular network. *Proc Natl Acad Sci USA* 101:10872–10877.
20. Segré D, Ben-Eli D, Lancet D (2000) Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. *Proc Natl Acad Sci USA* 97:4112–4117.
21. Higgs PG, Lehman N (2015) The RNA world: Cooperation and conflict at the origins of life. *Nature Rev Genet* 16:7–17.
22. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563.
23. Nowak MA (2006) *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press of Harvard, Cambridge, MA).
24. Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological game. *Science* 303:793–799.
25. Nghe P, Hordijk W, Kauffman SA, Walker SI, Schmidt FJ, Kemble H, Yeates JAM, Lehman N (2015) Prebiotic network evolution: six key parameters. *Mol BioSyst* 11:3206–3217.
26. Sokal RR, Rohlf FJ (1981) *Biometry: The Principles and Practice of Statistics in Biological Research.* (W. H. Freeman & Co., New York, NY).
27. Meyer JP, Seaman MA *Expanded Kruskal-Wallis Tables*, downloaded on 1July2015 from http://faculty.virginia.edu/kruskal-wallis/.