Supplementary Information for

Mineral surfaces select for longer RNA molecules

Ryo Mizuuchi, Alex Blokhuis, Lena Vincent, Philippe Nghe, Niles Lehman, David Baum

Ryo Mizuuchi
Email: mizuuchi@pdx.edu

This PDF file includes:

Materials and Methods

Mineral preparation. Mineral samples were obtained from Ward's Science (Rochester, NY, USA; magnetite: #470025-672, pyrite: #470206-112, pyrrhotite: #470025-750, calcite: #470025-512, apatite: #470226-354). Pyrite, pyrrhotite, and magnetite were ground mechanically using a jaw crusher and disk mill sieved to exclude grains smaller than 75 μm and greater than 150 μm and acid-washed as described in.[1] Briefly, the samples were sonicated in 95% ethanol 8‑10 times, soaked in 0.1 M nitric acid for 1 minute, and then rinsed thoroughly with nanopure water with a final 95% ethanol wash to prevent re-oxidation. The grains were then allowed to dry completely in an anaerobic chamber. Removal of the oxidized layers was verified by measuring the amount of sulfate released using ion chromatography. Calcite and apatite were ground by hand in a pestle and mortar (calcite, apatite) and sieved to remove particles under 75 μm.

RNA preparation. All the random RNAs and the **X**, **Y**, and **Z** RNA fragments were purchased from Tri-Link Biotechnologies (San Diego, CA, USA). The W fragment and the **WXYZ** product (size marker) were prepared through run-off *in vitro* transcription as described previously.[2] For quantification, random RNAs and the **W** fragment were 5´-labeled with γ^{32}P•[ATP] using T4 polynucleotide Kinase (NEB).

RNA adsorption onto mineral surfaces. Typically, 0.6 μM of each length-class of random RNA and their ^{32}P-labeled fraction (≤0.005 μM each) were incubated in 10 μL buffer (100 mM $MgCl_2$, 30 mM EPPS, pH 7.0) in the presence or absence of 0.2 mg of one type of mineral particle at 22 ºC (room temperature) for 2 h. In a subset of experiments, a 10-fold higher concentration (6 μM) was used instead. The RNAs were heated to 65 ºC for 2 min and cooled to room temperature just before use. The incubation time (2 h) was more than sufficient for the RNA length distribution on mineral surfaces to equilibrate (Fig. S7). After incubation, an aliquot of reaction solution was centrifuged (6,000 rpm, 30 sec) to remove as much supernatant as possible. An aliquot of the supernatant was put in nine volumes of stripping buffer (50% formamide, 8 M urea, 50 mM EDTA, pH 8.0, 0.025% bromophenol blue). Mineral particles were quickly washed with >300-fold wash solution (100 mM $MgCl_2$, 30 mM EPPS, pH 7.0) and centrifuged (6,000 rpm, 30 sec) to remove unbound RNAs. After the removal of wash solution, mineral particles were mixed with 50 μL of the stripping buffer per 0.2 mg minerals and incubated for 30 min at 22 ºC and 3 min at 80 ºC. The supernatant and stripped RNA solutions were subjected to polyacrylamide gel electrophoresis. Samples were heated to 80 ºC for 2 min, cooled on ice for 1 min, loaded onto a 15% polyacrylamide / 8 M urea denaturing gel in 1X TBE buffer, and electrophoresed at 900–1000 V for approximately 2 h. Gels were visualized by phosphorimaging on a Typhoon Trio+ imager (GE Healthcare), and band intensities were quantified using the ImageQuant™ TL software (GE Healthcare).

Self-assembly reaction with hydroxyapatite. The standard reaction mixture (20 μL) contained 2 μM **W**, **X**, **Y**, and **Z** fragments, 32P-labeled W fragment (≤0.02 μM each), 100 mM $MgCl_2$, 30 mM EPPS (pH 7.5), with or without 0.2 mg hydroxyapatite. The RNAs were heated to 65 ºC for 2 min and cooled to 22 ºC before use. The mixture was incubated at 48 ºC for 2 h. After the incubation, supernatant was put in the stripping

buffer, and RNAs that were adsorbed onto mineral surfaces were stripped off as described above. The RNA products were subjected to 8% polyacrylamide / 8M urea denaturing gel electrophoresis in 1X TBE buffer (800 V, approximately 2 h) and analyzed as described above.

Full details of mathematical models.

We first derive a model for the simultaneous adsorption of different oligomers on a 1D surface, to obtain exact expressions for the surface fraction covered by each oligomer in a low-dimensional case. We then consider some extensions of the model to higher dimensions for stiff and flexible oligomers. These approaches are derived from the works of Ramirez-Pastor et al.,[3,4] where they have been applied to the case of a single adsorbent.

I Adsorption of oligomers on a 1D lattice

We start by considering a large solution of RNA oligomers, each of which maintained at a fixed dimensionless concentration $\bar{c}$. In addition, the solution contains a mineral, with an exposed surface on which oligomers can adsorb. We consider the exposed surface to have Madsorption sites, with a size comparable to a single monomer. Correspondingly, to fully adsorb a k-mer, k adsorption sites need to be occupied. For our purposes, the RNA solution contains 8-,12-,16-,20- and 24-mers, and whenever we take a sum (e.g. $\sum$) it will denote a sum over these values.

Let us denote with N_i the number of adsorbed RNA oligomers of length i. In total, these oligomers occupy $i\,N_i$ mineral sites. Consequently, we find that the number of unoccupied mineral sites $N_\emptyset$ can be written as

$$N_\emptyset = M - \sum_i i\,N_i. \tag{1}$$

We denote by W the number of empty mineral sites plus the number of adsorbed species

$$W = N_\emptyset + \sum_i N_i. \tag{2}$$

A surface state is completely described by the exact sequence in which the surface bound molecules and empty sites appear. The number of states is consequently given by all their possible permutations

$$\Omega(\{N_i\}, M) = \binom{W}{N_8, N_{12}, N_{16}, N_{20}, N_{24}, N_\emptyset} = \frac{W!}{N_8!\,N_{12}!\,N_{16}!\,N_{20}!\,N_{24}!\,N_\emptyset!}. \tag{3}$$

We introduce a standard free energy for a k-mer on a mineral surface, $\mu^\circ_{k,min}$, which we model as an affine function of k

$$\mu^\circ_{k,min} = a_0 + a_1\,k\,, \tag{4}$$

where we suppose that $a_1 < 0$. It follows that $\Omega(\{N_i\}, M)$ is the partition function for a microcanonical ensemble, which can be related to a canonical ensemble $Q(\{N_i\}, M, T)$ via

$$Q(\{N_i\}, M, T) = \Omega(\{N_i\}, M) \exp(-\beta \sum_i N_i \, \mu^\circ_{i,min}), \tag{5}$$

where $\beta = \frac{1}{k_b T}$, with Boltzmann's constant k_b and the absolute temperature *T*. We can then extract the Helmholtz free energy F by

$$\beta F(\{N_i\}, M, T) = -\ln Q(\{N_i\}, M, T) = -\ln \Omega(\{N_i\}, M) + \beta \sum_i N_i \, \mu^\circ_{i,min}, \tag{6}$$

We can rewrite $\ln \Omega(\{N_i\}, M)$ as a mixing entropy, by performing a Stirling approximation $\ln N! = N \ln N - N + O(\ln N)$:

$$\ln \Omega(\{N_i\}, M) = W \ln W - \sum_i N_i \ln N_i - N_\emptyset \ln N_\emptyset, \tag{7}$$

which can be rewritten as

$$\ln \Omega(\{N_i\}, M) = -\sum_i N_i \ln \frac{N_i}{W} - N_\emptyset \ln \frac{N_\emptyset}{W}. \tag{8}$$

This has the functional form of a mixing entropy (but taken relative to W instead of M), which gives a clear interpretation of this object. However, we prefer to work in terms of surface coverages $\theta_i = i\frac{N_i}{M}$, $\theta_0 = \frac{N_\emptyset}{M}$, $\theta_W = \frac{W}{M}$. Let us therefore write:

$$\begin{aligned} \ln \Omega(\{N_i\}, M) &= W \ln \frac{W}{M} - \sum_i N_i \ln \frac{N_i}{M} - N_\emptyset \ln \frac{N_\emptyset}{M} \\ &= M \left[\theta_W \ln \theta_W - \sum_i \frac{\theta_i}{i} \ln \frac{\theta_i}{i} - \theta_0 \ln \theta_0 \right], \end{aligned} \tag{9}$$

where we have introduced $\theta_k = k\frac{N_k}{M}$ (fraction of sites covered by k-mers), $\theta_0 = \frac{N_\emptyset}{M}$ (fraction of empty sites), $\theta_W = \frac{W}{M}$ (fraction of empty sites and oligomers). The $\theta_W \ln \theta_W$ term in Eq. (9) is a consequence of multisite adsorption. We can now extract the chemical potential of adsorbed monomers of length k, using

$$\mu_{k,min} = \left(\frac{\partial F}{\partial N_k}\right)_{T,M,N_i \neq N_k}, \tag{10}$$

which after taking the appropriate derivatives affords the expression

$$\mu_{k,min} = \mu^\circ_{k,min} + k_b T \left[\ln \frac{\theta_k}{k} + (k-1) \ln \frac{\theta_W}{\theta_0} - \ln \theta_0 \right]. \tag{11}$$

If our mixture would contain only a single type of oligomer of length k, we recover the isotherm in [3].

Let us now put our system in contact with a large solution of oligomers, each of which maintained at a dimensionless concentration $\bar{c} = c_k/c^\circ$. Where c° is a standard concentration (1 M). The oligomers in solution have a chemical potential

$$\mu_k = \mu_k^\circ + k_b T \ln \bar{c}, \tag{12}$$

where μ_k° is a standard free energy of formation at concentration c°. We consider μ_k° to be an affine function of k:

$$\mu_k^\circ = b_0 + b_1\, k. \tag{13}$$

A reversible adsorption process will lead to chemical equilibrium, at which $\mu_{k,min} = \mu_k$. Let us now substitute Eq. (12) in Eq. (11), and write

$$0 = \Delta\mu_{k,ads}^* + k_b T \left[\ln \frac{\theta_k}{k} + (k-1) \ln \frac{\theta_W}{\theta_0} - \ln\, \theta_0 \right]. \tag{14}$$

From Eq. (4),(12), and (13) it follows that $\Delta\mu_{k,ads}^*$ is again an affine function. We write

$$\Delta\mu_{k,ads}^* = \epsilon + \delta\, k, \tag{15}$$

where $\epsilon = a_0 - b_0 - k_b T \ln \bar{c}$ and $\delta = a_1 - b_1$. Eq. (14) then gives the equilibrium surface coverage for k-mers

$$\theta_k = k \exp\big(-\beta(\epsilon + \delta k)\big) \left(\frac{\theta_0}{\theta_W}\right)^{k-1} \theta_0. \tag{16}$$

Let us define $r_k = 5 \frac{\theta_k}{k} / \sum \frac{\theta_i}{i}$ as the relative concentration of a k-mer. The ratio of k-mers to j-mers is then

$$\frac{r_k}{r_j} = \frac{j\, \theta_k}{k\, \theta_j} = \left(\frac{\theta_0}{\theta_W}\right)^{k-j} \exp\big(-\beta\, \delta\, (k-j)\big). \tag{17}$$

Because $W \geq N_\emptyset$, $\theta_0/\theta_W \leq 1$. If we consider $k > j$, we see that the entropic term $\left(\frac{\theta_0}{\theta_W}\right)^{k-j}$ favors shorter oligomers. This is to be expected, as smaller oligomers allow for more possible surface configurations. Because $\delta < 0$, the factor $\exp\big(-\beta\, \delta\, (k-j)\big)$ favors longer oligomers. From Eq. (17), we deduce that the relative concentrations of adsorbed oligomers follow an exponential trend. We will now extend the model by relaxing the 1D assumption. Interestingly, this can largely be taken account by simply shifting the constants ϵ and δ. Consequently, we can proceed using Eq. (16).

IIa 2D: Connectivity ansatz

As shown in [3], an effective way to describe stiff oligomers on a 2D lattice is by introducing a connectivity ansatz. Let c be the number of connections of a lattice point (for a 2D square lattice: 4, on the line: 2). By supposing Ω scales with dimension as the one for the Flory model, it can be shown that

$$\frac{\Omega(M,N,c)}{\Omega(M,N,c')} = \left[\frac{c-1}{c'-1}\right]^{\sum_i N_i\,(i-1)}. \tag{18}$$

Perfoming our previous calculation and setting $c' = 2$, we find for arbitrary c

$$\mu_{k,min,c} = \mu_{k,min,2} - kT\,(k-1)\ln(c-1). \tag{19}$$

This means we can incorporate it in $\Delta\mu^*_{ads}$ by defining $\epsilon' = \epsilon + k_bT\ln(c-1)$ and $\delta' = \delta - k_bT\ln(c-1)$.

IIb 2D: Dilute lattice placements

Another extension, put forward in,[4] is to study a dilute limit and consider the number of ways an oligomer can be placed on a lattice. For stiff oligomers, which can only be placed on square lattices, we can then consider every 1D placement, and place them along all $c/2$ directions. The microcanonical partition function then grows as:

$$\frac{\Omega(M,N,c)}{\Omega(M,N,c')} = \left[\frac{c}{c'}\right]^{\sum_i N_i}. \tag{20}$$

Which yields a constant correction to μ_k

$$\mu_{k,min,c} = \mu_{k,min,2} - k_bT\,\ln(c/2)\,. \tag{21}$$

It can be absorbed in the expression for $\Delta\mu^*_{ads}$, by defining $\epsilon' = \epsilon + k_bT\ln(c/2)$. For dilute flexible oligomers, the number of single-oligomer configurations $\gamma(c,k)$ is the number of self-avoiding random walks of length k. On a 2D square lattice, this quantity behaves as $\gamma(c,k) = u^k k^v$, with $u_{2d} \approx 2.62$, $v_{2d} = \frac{11}{32}$. On a 3D lattice, we have $v_{3d} \approx 0.16$. The correction for dilute systems is then

$$\mu_{k,min,4} = \mu_{k,min,2} - k_bT[k\ln u_{2d}/u_{3d} + (v_{2d} - v_{3d})\ \ln k]. \tag{22}$$

Just as with the connectivity ansatz, we can absorb a contribution proportional to k, because we can write $\delta' = \delta - k_bT\ln u$. The $\ln k$ contribution gives a new factor $\left(\frac{k}{j}\right)^{v_{2d}-v_{3d}}$. However, because $v_{2d} - v_{3d} \approx 0.18$, this contribution is relatively small, and we will neglect it in our further derivation. Overall, we see that the extension of the model to 2D for stiff and flexible polymers can be accounted for by shifting the parameters in the adsorption energy. In the subsequent sections, we will solve the model for Eq. (16).

III Solving for θ_i

Since the solutions are expressed in terms of θ_0/θ_W, we do not have a full solution. Expressed in terms of θ_i, we find

$$\theta_W = \sum_i \frac{\theta_i}{i} + \theta_0, \tag{23}$$

and

$$\theta_0 = 1 - \sum_i \theta_i. \tag{24}$$

We can then write:

$$\theta_W = \sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\ (\theta_0/\theta_W)^{i-1}\theta_0 + \theta_0. \tag{25}$$

Let us define $\zeta = \frac{\theta_0}{\theta_W}$. From Eq. (23), we then find

$$1 = \sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\ \zeta^i + \zeta, \tag{26}$$

which is a nonlinear polynomial equation from which we need a real root $\zeta < 1$.
We can then express θ_0 as

$$\theta_0 = \frac{1}{1 + \sum_i i\ \exp\left(-\beta(\epsilon + \delta i)\right)\ \zeta^{i-1}}, \tag{27}$$

and thus we can numerically solve the system of equations by finding ζ.

IV Temperature dependence

It is observed that the relative abundance of longer RNAs increases at higher temperature. To investigate this effect, let us again consider the quantity $\frac{r_k}{r_j}$ in Eq. (17), where $k > j$, and study its derivative with regard to temperature T.

$$\frac{d\left(\frac{r_k}{r_j}\right)}{dT} = \left(\frac{r_k}{r_j}\right)\left[\frac{\delta - T\frac{\mathrm{d}\delta}{dT}}{kT^2}(k-j) + (k-j)\frac{1}{\zeta}\frac{d\zeta}{dT}\right]. \tag{28}$$

Since δ and ϵ correspond to a Gibb's free energy change, we can write them as enthalpies $\Delta h_\delta, \Delta h_\epsilon$ and entropies $\Delta s_\delta, \Delta s_\epsilon$.

$$\delta = \Delta h_\delta - T\Delta s_\delta, \tag{29}$$

$$\epsilon = \Delta h_\epsilon - T\Delta s_\epsilon, \tag{30}$$

And thus

$$\delta - T\frac{\mathrm{d}\delta}{dT} = \Delta h_\delta. \tag{31}$$

Taking the derivative with regard to T of Eq. (26) , we find that

$$\left(\sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\ i\ \zeta^{i-1} + 1\right)\frac{d\zeta}{dT} = -\sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\ \left(\Delta h_\epsilon + i\ \Delta h_\delta\right)/kT^2\ \zeta^i, \tag{32}$$

which after rewriting gives

$$\frac{d\zeta}{dT} = \frac{\frac{\Delta h_\delta}{kT^2}\zeta - \frac{\Delta h_\epsilon}{kT^2}\sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\zeta^i}{\sum_i \exp\left(-\beta(\epsilon + \delta i)\right)\ i\ \zeta^{i-1} + 1} - \frac{\Delta h_\delta}{kT^2}\zeta. \tag{33}$$

Plugging this back in Eq. (28), we then have

$$\frac{d\left(\frac{r_k}{r_j}\right)}{dT} = \left(\frac{r_k}{r_j}\right)(k-j)\left[\frac{\frac{\Delta h_\delta}{kT^2} - \frac{\Delta h_\epsilon}{kT^2}\sum_i \exp\left(-\beta(\epsilon+\delta i)\right)\zeta^{i-1}}{\sum_i \exp\left(-\beta(\epsilon+\delta i)\right)\, i\,\zeta^{i-1} + 1}\right]. \tag{34}$$

It follows that selectivity can increase with temperature, provided that the enthalpic contributions obey

$$\Delta h_\delta - \Delta h_\epsilon \sum_i \exp\left(-\beta(\epsilon+\delta i)\right)\zeta^{i-1} > 0. \tag{35}$$

Example: consider Fig. 4D, for which $\epsilon = 3.4\ k_b T^*$, $\delta = -0.8\ k_b T^*$ and $\left(\frac{r_{24}}{r_8}\right) = 2.90$, with $T^* = 293$K a reference temperature. Numerically, we find that selectivity would increase with T in this case if

$$-0.884\ \Delta h_\epsilon > -\Delta h_\delta. \tag{36}$$

As an illustration, let us choose $\Delta h_\epsilon = -11.5\ k_b T^*$, $\Delta h_\delta = -2.5\ k_b T^*$ From Eq. (29) and Eq. (30) it follows that $\Delta s_\epsilon = 14.9\ k_b$, $\Delta s_\delta = 1.7\ k_b$. Augmenting the temperature with ΔT =15 K, we would then have for a 24-mer versus an 8-mer

$$\frac{d\left(\frac{r_{24}}{r_8}\right)}{dT}\Delta T \approx 0.072 \cdot 15 = 1.08, \tag{37}$$

which corresponds well with the order of magnitude observed in the experiment.

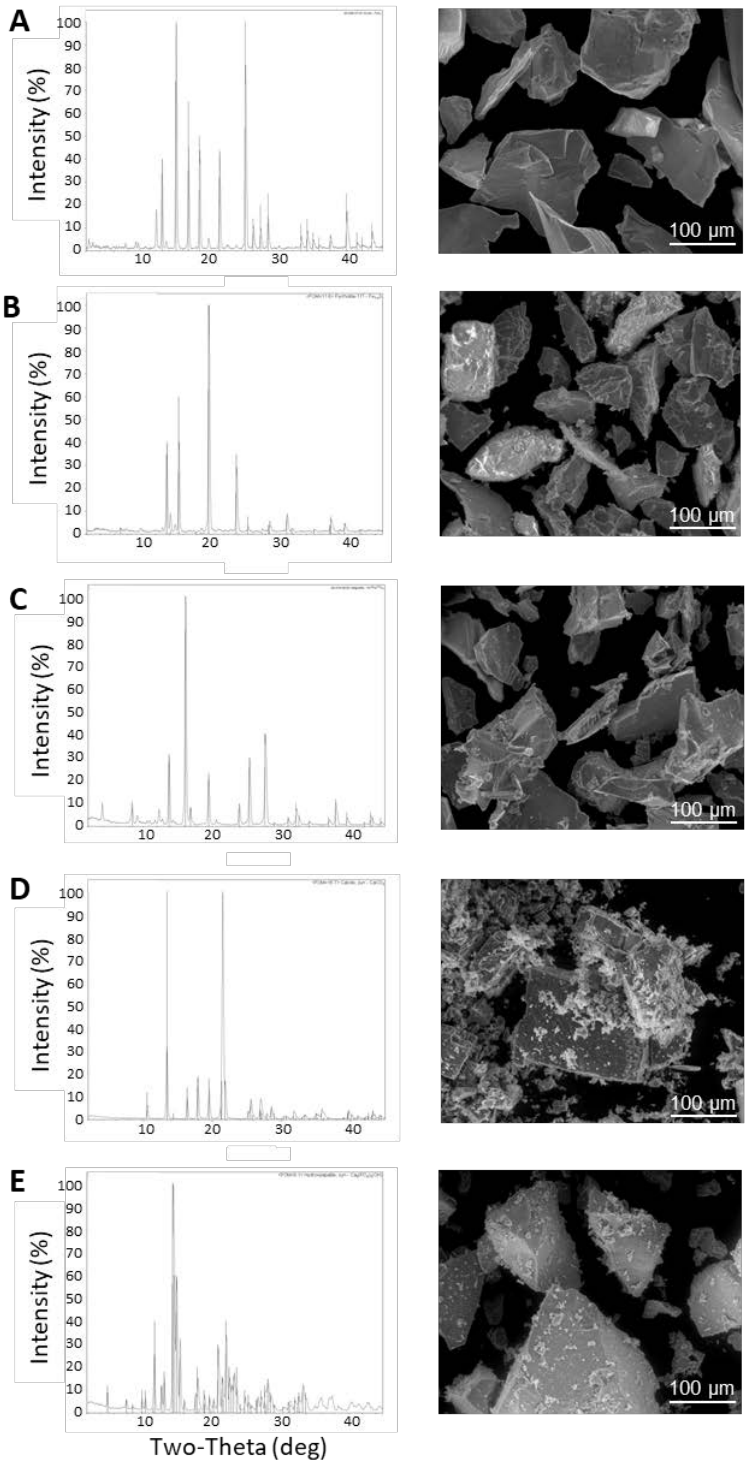

Fig. S1. Analysis of (A) pyrite, (B) pyrrhotite, (C) magnetite, (D) calcite, and (E) hydroxyapatite grains by X-ray diffraction (*left*) and scanning electron microscopy (*right*). Diffractograms were obtained on a RIGAKU D/Max Rapid II instrument (50 kV; 50 mA; 30-60 min exposure) and the resulting diffraction patterns were identified using Jade software and Powder Diffraction Files from the International Centre for Diffraction Data (ICDD). The grains were imaged on a FEI Quanta 200 in low-vaccum mode.

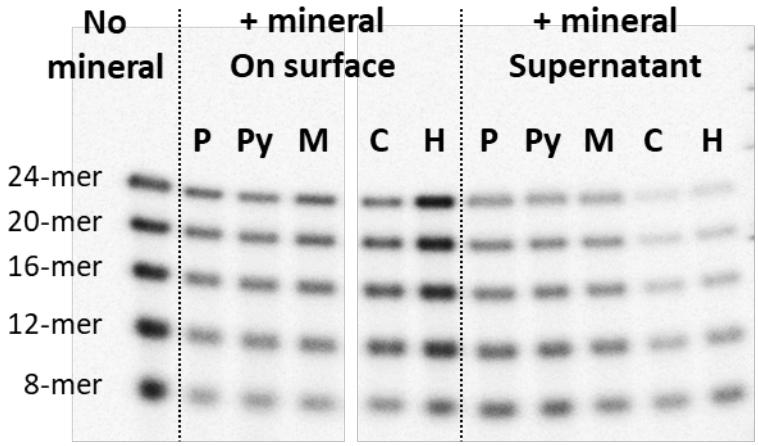

Fig. S2. An example of a polyacrylamide gel from the adsorption experiment. The RNA samples (no mineral control reaction, adsorbed on surface, in supernatant) were loaded onto a 15% polyacrylamide / 8 M urea denaturing gel. The type of minerals used in the experiments were: P, pyrite; Py, pyrrhotite; M, magnetite; C, calcite; H, hydroxyapatite. The two presented images were on the same gel (same contrast).

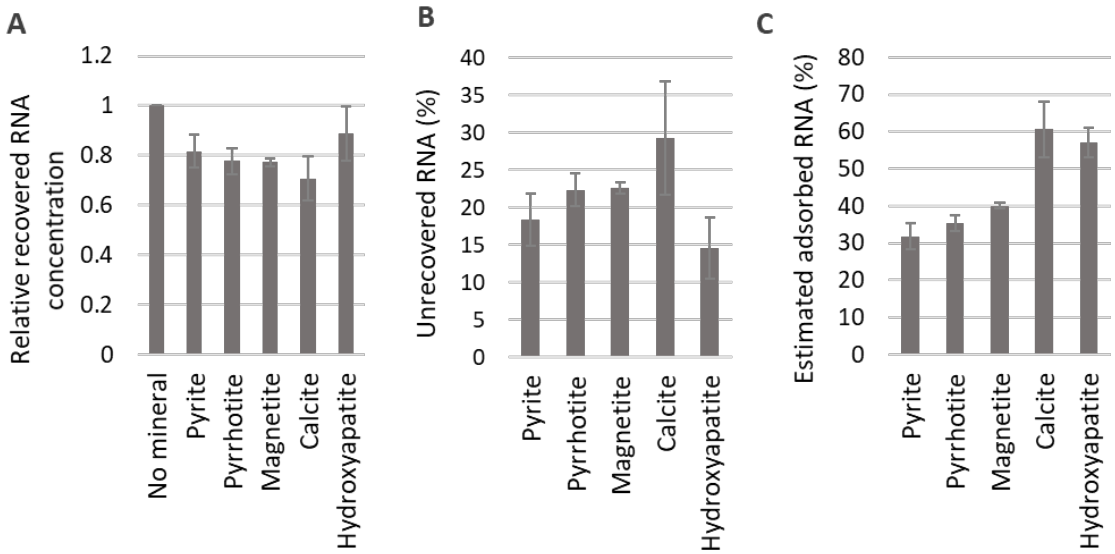

Fig. S3. Estimated concentration of RNAs in Fig. 1A, B. (A) Total concentration of recovered RNAs in each reaction, calculated by summing RNAs stripped from minerals and RNAs in supernatant. The concentration was normalized to the level of the control reaction performed in the absence of minerals. (B) Estimated percentage of unrecovered RNAs, whose length distribution could not be determined. This may represent RNAs unstripped off the minerals, degraded or precipitated in the presence of minerals, or removed during the process of washing minerals due to weak bindings, which were unspecified. (C) The percent of total concentration of RNAs adsorbed onto minerals, estimated by dividing the concentration of RNAs stripped from minerals by the total concentrations of recovered RNAs. In all panels, the error bars indicate standard errors (N = 3).

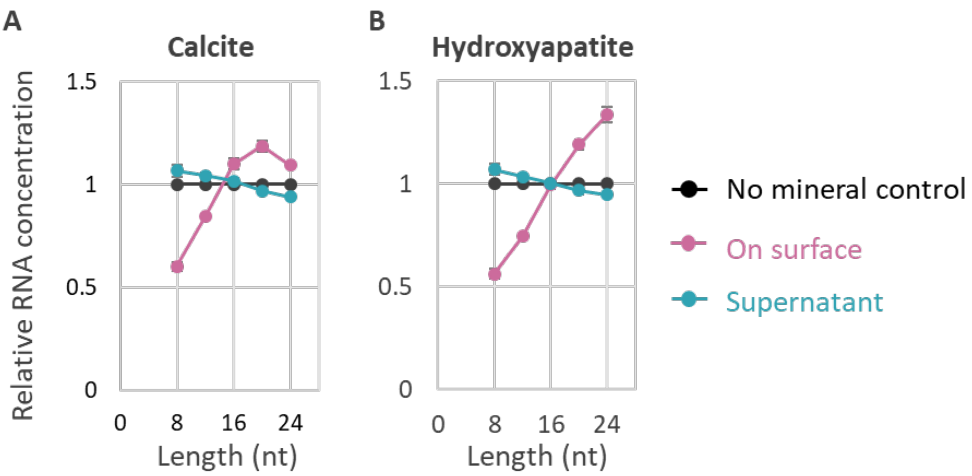

Fig. S4. The adsorption of RNA on calcite and hydroxyapatite by mineral surfaces with higher RNA concentrations. The adsorption experiments were performed with (A) calcite or (B) hydroxyapatite using the same method as the experiments in Fig. 1A, B except for a 10-fold higher concentration of the RNAs (6 μM of each length RNA). Estimated percentage of adsorbed RNA was 13% for both calcite and apatite. The error bars indicate standard deviations (N = 3).

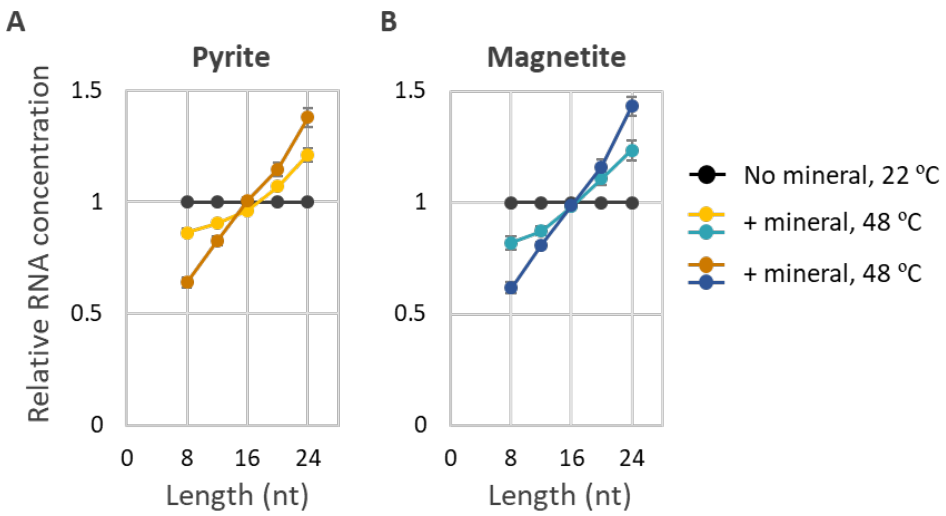

Fig. S5. The effect of a high temperature for RNA adsorption on pyrite and magnetite. The adsorption experiments were performed with (A) pyrite or (B) magnetite at 22 ºC or 48 ºC and RNA concentrations were determined in the same method as the experiments in Fig. 1. The error bars indicate standard deviations (N = 3).

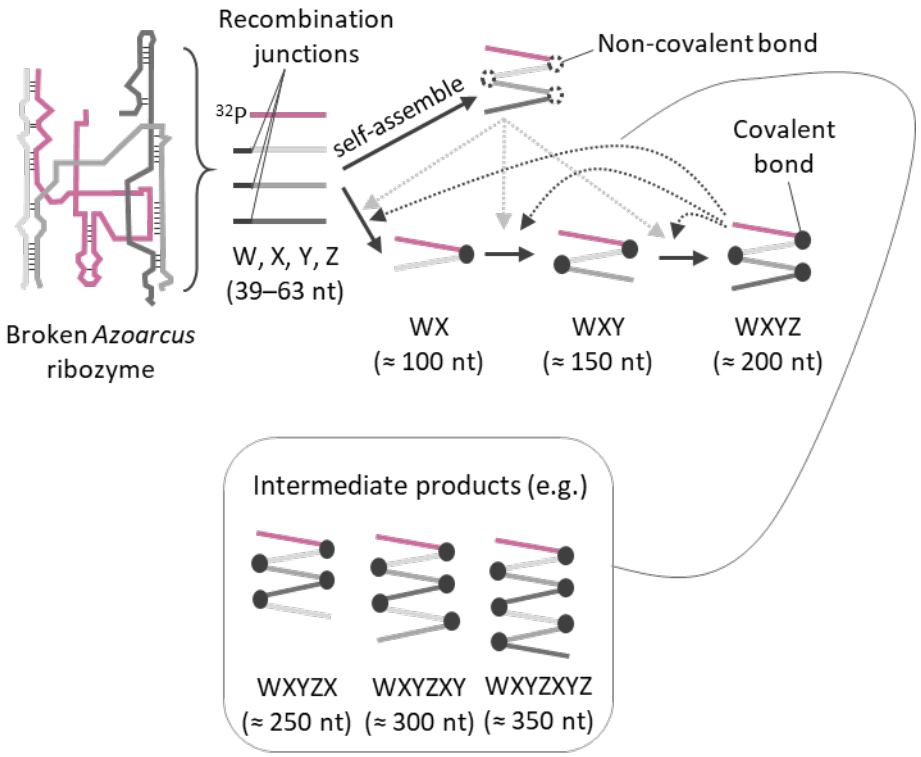

Fig. S6. Design of the *Azoarcus* self-assembly reaction. A part of **W** fragment (pink) was ^{32}P-labeled. Dotted arrows indicate weak (grey) or strong (black) catalytic reactions. There are reaction intermediates (250–350 nt products). The exact intermediate products remain unspecified.

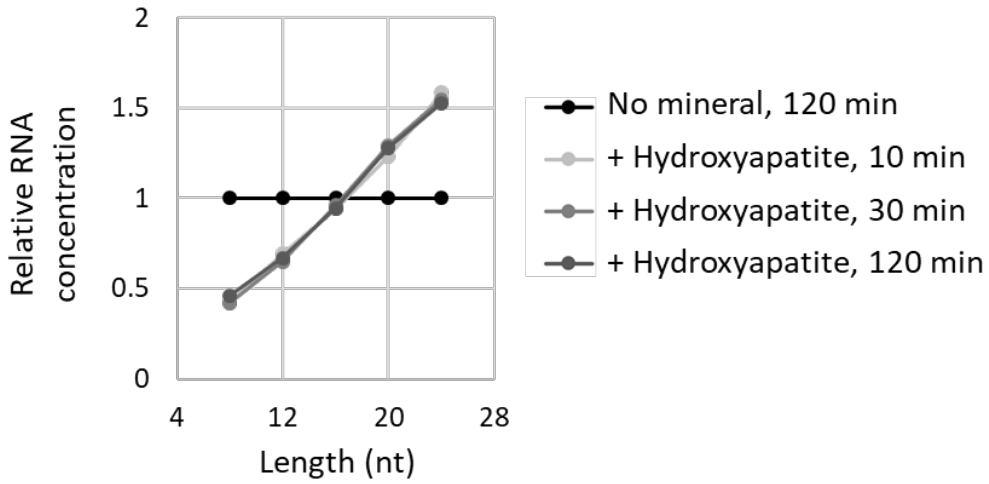

Fig. S7. Time course of the change of RNA concentrations on mineral surfaces. Adsorption experiment was performed by incubating a mixture of 8-, 12-, 16-, 20-, 24-mer fully random RNAs (0.6 μM each) and 0.2 mg hydroxyapatite in 10 μl at 22 ºC for 10 min, 30 min, or 120 min. The concentrations were determined by radioactivity of 32P-labeled RNA and normalized to the levels of the control reaction (120 min) performed without minerals.

References

1 M. A. Mckibben and H. L. Barnh, *Geochim. Cosmochim. Acta*, 1986, **50**, 1509–1520.

2 E. J. Hayden and N. Lehman, *Chem. Biol.*, 2006, **13**, 909–918.

3 A. J. Ramirez-Pastor, T. P. Eggarter, V. D. Pereyra and J. L. Riccardo, *Phys. Rev. B*, 1999, **59**, 11027–11036.

4 F. Romá, J. L. Riccardo and A. J. Ramirez-Pastor, *Ind. Eng. Chem. Res.*, 2006, **45**, 2046–2053.