Supplementary Information (SI) for Physical Chemistry Chemical Physics.

Supporting Information

Stringent Selection on Kinetics of Condensation Reactions: Early Steps in Chemical Evolution

Pau Capera-Aragones[a,b,c], Kavita Matange[a,b], Vahab Rajaei[a,b], Yuval Pinter[d], Anton S Petrov[a,b], Loren Dean Williams[a,b*], and Moran Frenkel-Pinter[b,c,e*]

[a] NASA Center for Integration of the Origins of Life, Georgia Institute of Technology, Atlanta, GA (USA)

[b] School of Chemistry and Biochemistry Georgia Institute of Technology, 315 Ferst Drive NW, Atlanta, GA 30332-0400

[c] Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 9190401, Israel

[d] Department of Computer Science, Ben-Gurion University of the Negev, Israel

[e] Blue Marble Space Institute of Science, Seattle, WA 98104, USA

*Corresponding authors:

Professor Moran Frenkel-Pinter
The Center for Nanoscience and Nanotechnology
Institute of Chemistry
The Hebrew University of Jerusalem
Edmond J. Safra Campus
Jerusalem 9190401, Israel
Ph: (+972)-2-6584171
moran.fp@mail.huji.ac.il

Professor Loren Dean Williams
School of Chemistry and Biochemistry
Georgia Institute of Technology
315 Ferst Drive NW
Atlanta, GA 30332-0400
Ph: (+1) (404) 385-6258
loren.williams@chemistry.gatech.edu

Table of Contents

Methods 3
Thermodynamic and kinetic principles. 3
Approximations 4
Dimerization reactions 4
Ordinary Differential Equations. 5
Supplementary Figures 6
Figure S1. Normalized number of types of monomers plus dimers versus T. 6
Figure S2. Normalized number of types of monomers and dimers versus T or t. 7
Figure S3. Normalized number of types of monomers plus dimers versus T or t. 8
Figure S4. Changes in the normalized concentrations of monomers and dimers over reaction times. 9
Figure S5. Robustness of combinatorial compression to alternative rate-constant parameterizations 10
Figure S6. Schematic representation of free energy landscapes. 11
References 12

Methods

Here we provide an explanation and technical rationale of our simulations including: (i) general thermodynamic and kinetic principles, (ii) the generic chemical species and reactions, and (iii) the modelling approach.

Thermodynamic and kinetic principles.

The maximum amount of non-expansion work that can be performed by a thermodynamically closed chemical reaction at constant temperature and pressure is the Gibbs free energy change of the reaction ($^{\Delta G_{rxn}}$).[1] Under standard conditions, with

reactants and products in their thermodynamic standard states of concentration, temperature and pressure, this is represented by the standard Gibbs free energy change (ΔG^{o}_{rxn}).

The defining equation for ΔG^{o}_{rxn} is:

(eq 1) $$\Delta G^{o}_{rxn} = \Delta H^{o}_{rxn} - TS^{o}_{rxn}$$

where ΔH^{o}_{rxn} is the standard enthalpy change, $S^{o}_{rxn}$is the standard entropy change and T is the absolute temperature. The magnitude of ΔG_{rxn} is minimized at zero when a system is at equilibrium:

(eq 2) $$\Delta G_{rxn} = \Delta G^{o}_{rxn} - RTlnQ$$

(where $Q =$ the reaction quotient). ΔG_{rxn} reflects the position of a reaction mixture relative to the equilibrium condition and describes the thermodynamic favorability of a reaction. A negative ΔG_{rxn} ($\Delta G_{rxn} < 0$) indicates a reaction is spontaneous and exergonic. Figure S6 graphically summarizes the energies involved on an endergonic and on an exergonic reaction. A chemical reaction is exothermic if $\Delta H_{rxn} < 0$ and endothermic if $\Delta H_{rxn} > 0$.

Rate constants can be approximated using the Eyring Transition State Theory:

(eq 3) $$k = \frac{\kappa_B T}{h} e^{-\Delta G^{\ddagger}/RT}$$

where κ_B is the Boltzmann's constant, T is the absolute temperature, h is the Plank's constant, R is the universal gas constant, and $\Delta G^{\ddagger}$ is the standard activation Free Energy. The rate of a reaction is determined by $\Delta G^{\ddagger}$, concentrations and temperature.

In sum, ΔG_{rxn} governs the thermodynamics of a reaction, and $\Delta G^{\ddagger}$ governs the kinetics. In describing our work, generally we report inputs of ΔH^{o}_{rxn}, ΔS^{o}_{rxn}, and $\Delta G^{\ddagger}$ and used those parameters with various combinations of T, time and initial concentrations to compute ΔG_{rxn}, reaction rates, and reactant and product concentrations. With ΔH^{o}_{rxn} and ΔS^{o}_{rxn} we compute ΔG^{o}_{rxn} using (eq. 1), and with ΔG^{o}_{rxn}, $\Delta G^{\ddagger}_{forward,}$ and (eq. 3) we compute forward and reverse reaction rates using $\Delta G^{\ddagger}_{forward}$=$\Delta G^{\ddagger}_{reverse} + \Delta G^{o}_{rxn}$ (see Figure S6).

Approximations

There are several approximations used in our approach. First, we assume that H^{o}_{rxn}, ΔS^{o}_{rxn}, and $\Delta G^{\ddagger}$ are independent of T. Second, we use the Eyring equation [Eq. (3)] which may not be fully accurate for molecules with complex shapes. However, we empirically confirmed our results are not qualitatively impacted by Eyring approximations. Figure S5 shows that even in the hypothetical case of rate constant not given by the Eyring equation, our results would still qualitatively hold, and combinatorial compression is maintained in the simulation.

Dimerization reactions.

In this work, we focus on reversible dimerization reactions. In some cases simulated systems consist of 11 different monomers and their 121 potential dimer products resulting from linking monomers, where the order of monomers in the dimer does matter (11^2). In other cases systems consist of 12 different monomers, in which one monomer is a kinetic compressor. These systems give 144 potential dimer products resulting from linking the monomers (12^2). Our simulations have some analogy to the systems we have characterized previously by experiment.[2]

The chemical reactions are of type:

$$M_i + M_j \rightleftharpoons D_{ij} + W \quad \text{(eq 4)}$$

where M_i and M_j are monomers, D_{ij} are dimers, and W is water. Note that as $[W]$ decreases, the equilibrium condition of the system shifts towards higher [D_{ij}] and lower [M_i] and [M_j], whereas when $[W]$ increases the equilibrium condition shifts towards lower [D_{ij}] and higher [M_i] and [M_j]. In our simulations here, the “dry” state is defined as [W] = 0.5 M, which is not anhydrous. If we assume that reaction quotient of activities is determined by concentrations, we have a reaction quotient equal to $Q = [W][D_{ij}]/[M_i M_j]$.

Simulations initiate with [$D_{i,j}$] = 0. Net dimerization is expected and is observed. Combinatorial explosion is the default expectation, because all possible dimers can form.

Ordinary Differential Equations.

In this work we performed computer simulations using systems of ordinary differential equations (ODE's). The system of equations is solved numerically using Python. Nonlinearity precludes an analytic solution. The code is available online at GitHub under the link: https://github.com/HUJI-MFP/Compression/tree/main

For the reversible dimerization reactions (eq. 4), the system of ordinary differential equations is:

$$\frac{dM_i}{dt} = -k_{M_iM_j}M_iM_j + k_{D_{ij}}D_{ij}W \quad \text{(eq 5)}$$

$$\frac{dP_{ij}}{dt} = -k_{P_{ij}}D_{ij}W + k_{M_iM_j}M_iM_j \quad \text{(eq 6)}$$

where $k_{M_iM_j}$ is the rate constant for the condensation of M_i and M_j, $k_{D_{ij}}$ is the rate constant for the hydrolysis of $D_{i,j}$. For condensation, $k_{M_iM_j}$ is the forward rate constant and $k_{D_{ij}}$ is the reverse rate constant. Rate constants were approximated from the Eyring equation (eq. 3) and are given in Table 1 of the main manuscript. N is the number of different types of M_i, typically 9, 11 or 12. The system of differential equations contains N equations of the type in (eq. 5) and N^2 equations of the type in (eq. 6). For $N = 12$, this corresponds to a total of $N + N^2 = 12 + 12^2 =$ 156 coupled ODE's.

Supplementary Figures

Figure S1. Normalized number of types of monomers plus dimers versus T.

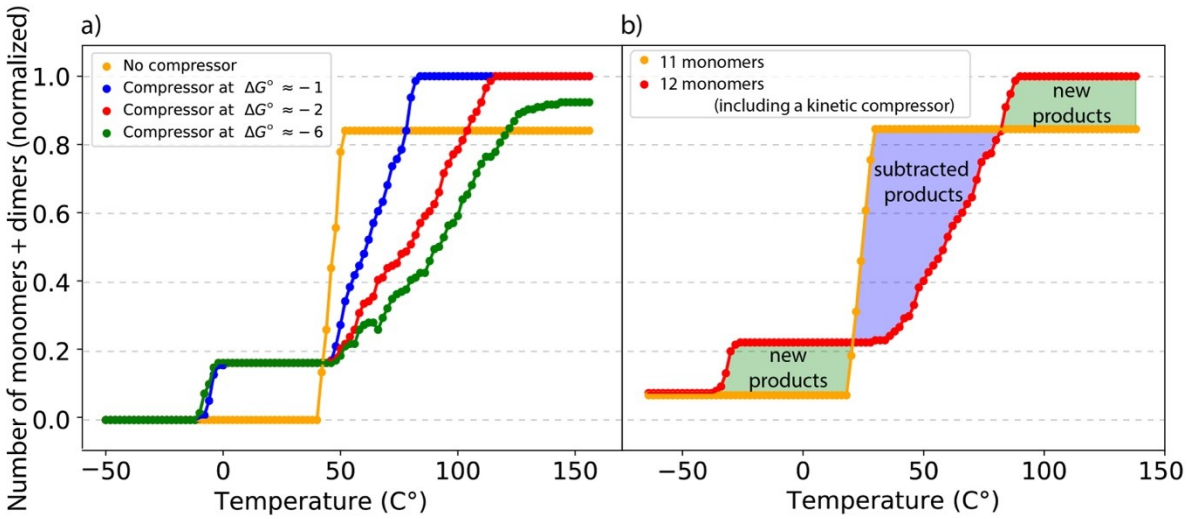

Figure S1. Normalized number of types of monomers plus dimers versus T. a) The number of product dimers is indicated in the absence of a kinetic compressor and in the presence of a kinetic compressor for ΔG^{o}_{rxn} = -1, -2, or -6 kcal/mol. The kinetic compressor has a relatively small $\Delta G^{\ddagger}$ of 25 kcal/mole and is governed by fast kinetics but can have a variety of ΔG^{o}_{rxn}. b) The number of dimer products depends upon the presence of the compressor and on T. The number of product dimers in the presence and absence of a kinetic compressor with $\Delta G^{o}_{rxn} = -2$ kcal/mol. Simulations are for a mixture of 11 different reactant monomers with 121 possible product dimers (orange lines), or 12 reactant monomers, including a kinetic compressor and 144 possible product dimers. Initial concentrations are 0.05 M for monomers and 0 M for dimers. The kinetic compressor is fed to maintain a constant concentration of 0.05 M. ΔH^{o}_{rxn} is 4, 8, or 9, kcal/mol with a variance of 0.5 kcal/mol, meaning that a random value between -0.5 and 0.5 kcal/mol is randomly added to each ΔH^{o}_{rxn}. ΔS^{o}_{rxn} is 0.033 kcal/mol. $\Delta G^{\ddagger}$ = 30 ± 0.5 kcal/mol, meaning that a random value between -0.5 and 0.5 kcal/mol is randomly added to each $\Delta G^{\ddagger}$. No variance is applied to the kinetic compressor, for which $\Delta G^{\ddagger}$ = 25 kcal/mol. Reaction time is 3 days.

Figure S2. Normalized number of types of monomers and dimers versus T or t.

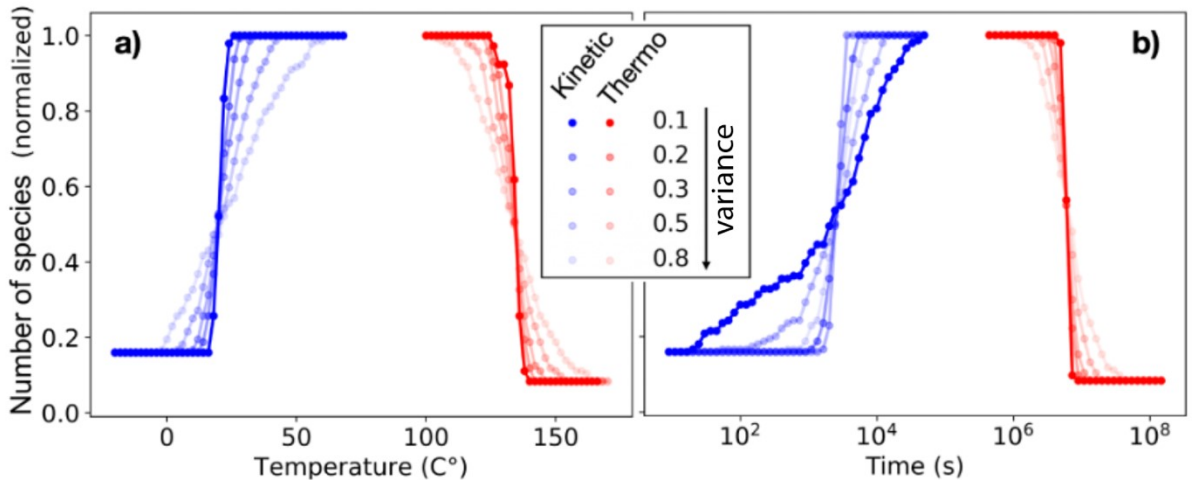

Figure S2. Normalized number of types of monomers and dimers versus temperature (T) or reaction time (t). The size of the random variance influences the sharpness of the compression/explosion transition. a) The number of monomers and dimers varies with T at constant t = 3 hours. b) The number of monomers and dimers varies with t at constant T = 40°C. The results for reactions with a kinetic compressor are indicated by blue lines. The results for reactions with a thermodynamic compressor are indicated by red lines. Transitions from kinetic compression to explosion and from explosion to thermodynamic compression are observed. These simulations used 9 monomers and 81 possible dimers. ΔH^o is 8 ± except for reactions of the thermodynamic compressor for which ΔH^o_{rxn} is 0 ± 0.5 kcal/mol. ΔS^o_{rxn} is 0.033 kcal/mol. $\Delta G^{\ddagger}$ is 24 ± 0.5 kcal/mol for all reactions except those of the kinetic compressor, for which $\Delta G^{\ddagger}$ is 19 ± 0.5 kcal/mol. As in Figure S1,the initial concentrations of monomers are 0.05 M and of dimers are 0 M. The concentration of the thermodynamic compressor is constant at 0.05 M and a random variance is applied to ΔH^o_{rxn} and $\Delta G^{\ddagger}$.

Figure S3. Normalized number of types of monomers plus dimers versus T or t.

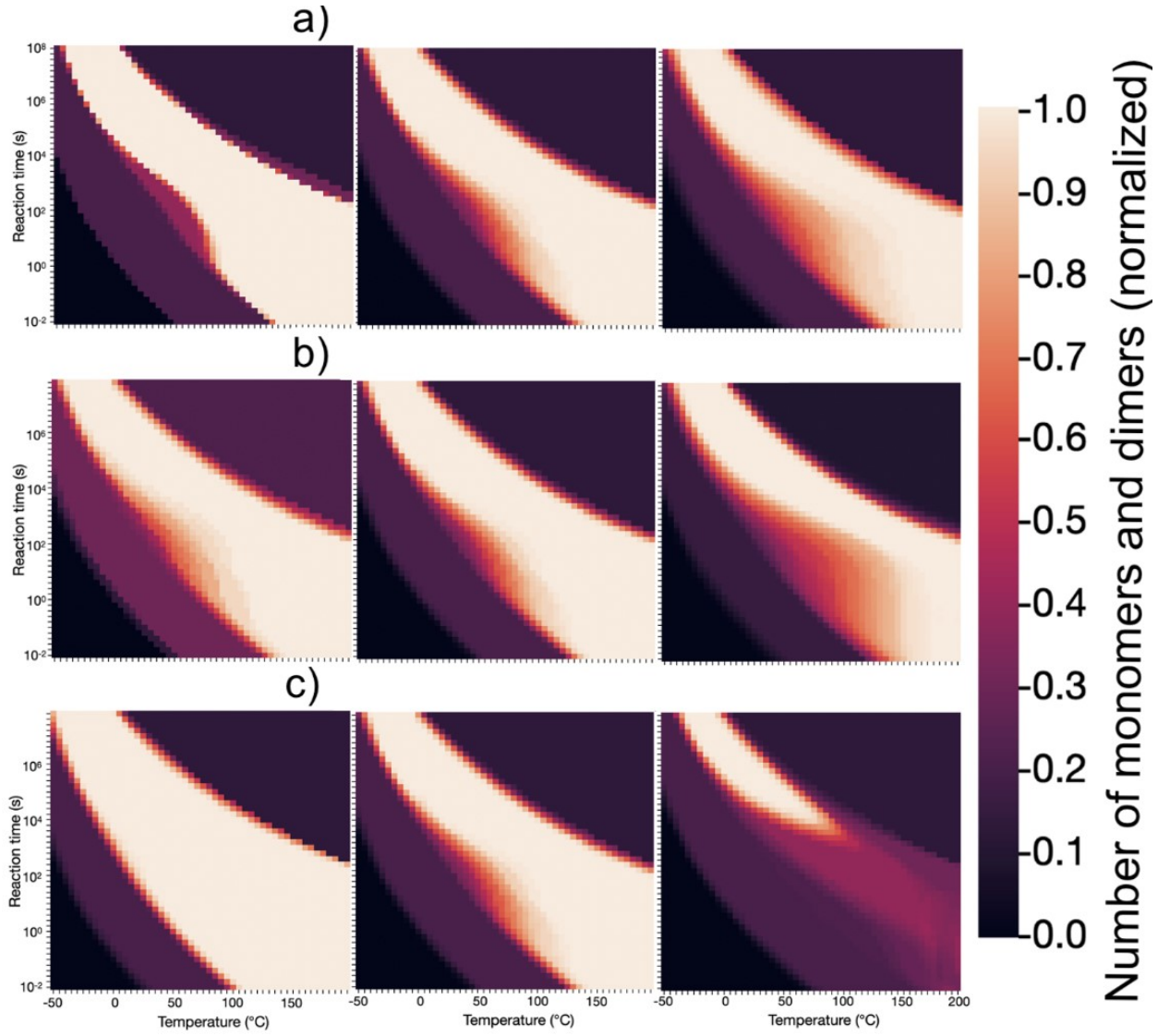

Figure S3. Normalized number of types of monomers plus dimers versus temperature (T) and reaction time (t). a) The variance of ΔH^{o}_{rxn} and $\Delta G^{\ddagger}$ (from left to right) is 0.1, 0.5, and 0.8 kcal/mol. b) The number of starting monomers from left to right is 6, 9, and 12. c) The threshold of detection (from left to right) is 10^{-4} M, 10^{-5} M, and 10^{-6} M. As in Figure S1, the initial concentrations of monomers are 0.05 M and of dimers are 0 M. The concentration of the thermodynamic compressor is constant at 0.05 M. The variance is randomly applied to ΔH^{o}_{rxn} and $\Delta G^{\ddagger}$ and is equal to 0.5 kcal/mol, except for panel a. The number of monomers is 9 except for panel b. The threshold of detection is 10^{-5} M except in panel c.

Figure S4. Changes in the normalized concentrations of monomers and dimers over reaction times.

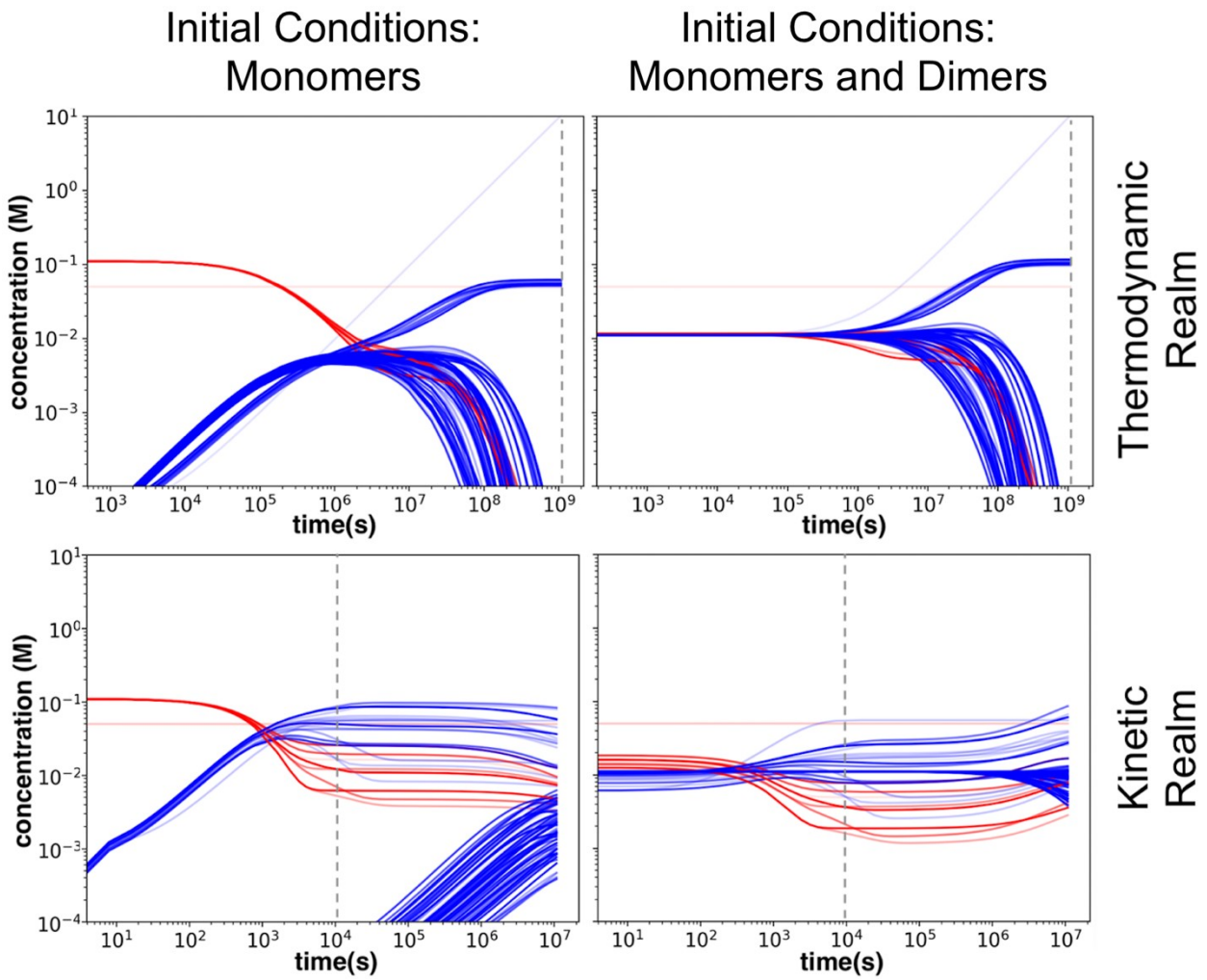

Figure S4. Changes in the normalized concentrations of monomers (red lines) and dimers (blue lines) over reaction times. Thermodynamic dominance is demonstrated in the top row. Kinetic dominance is on the bottom row. On the left column, the initial conditions of the mixture are set to 0 for all dimers and to 0.1 M for each monomer, except for the thermodynamic and kinetic compressor, which is set to 0.05 M. On the right column, the initial conditions are set to 0.05 M for each (monomer and dimer). ΔH^{o}_{rxn} is 8.5 kcal/mol except for the thermodynamic compressor for which it is 0 kcal/mol. ΔS^{o} is 0.033 kcal/mol kcal/mol. $\Delta G^{\ddagger}$ is 26 kcal/mol, except for reactions of the kinetic compressor, for which it is 21 kcal/mol. The thermodynamic regime is identified at time 10^9 s, whereas kinetic regime is identified at time 10^4 s. As in Figure 1 a variance is randomly applied to ΔH^{o}_{rxn} and $\Delta G^{\ddagger}$.

Figure S5. Robustness of combinatorial compression to alternative rate-constant parameterizations.

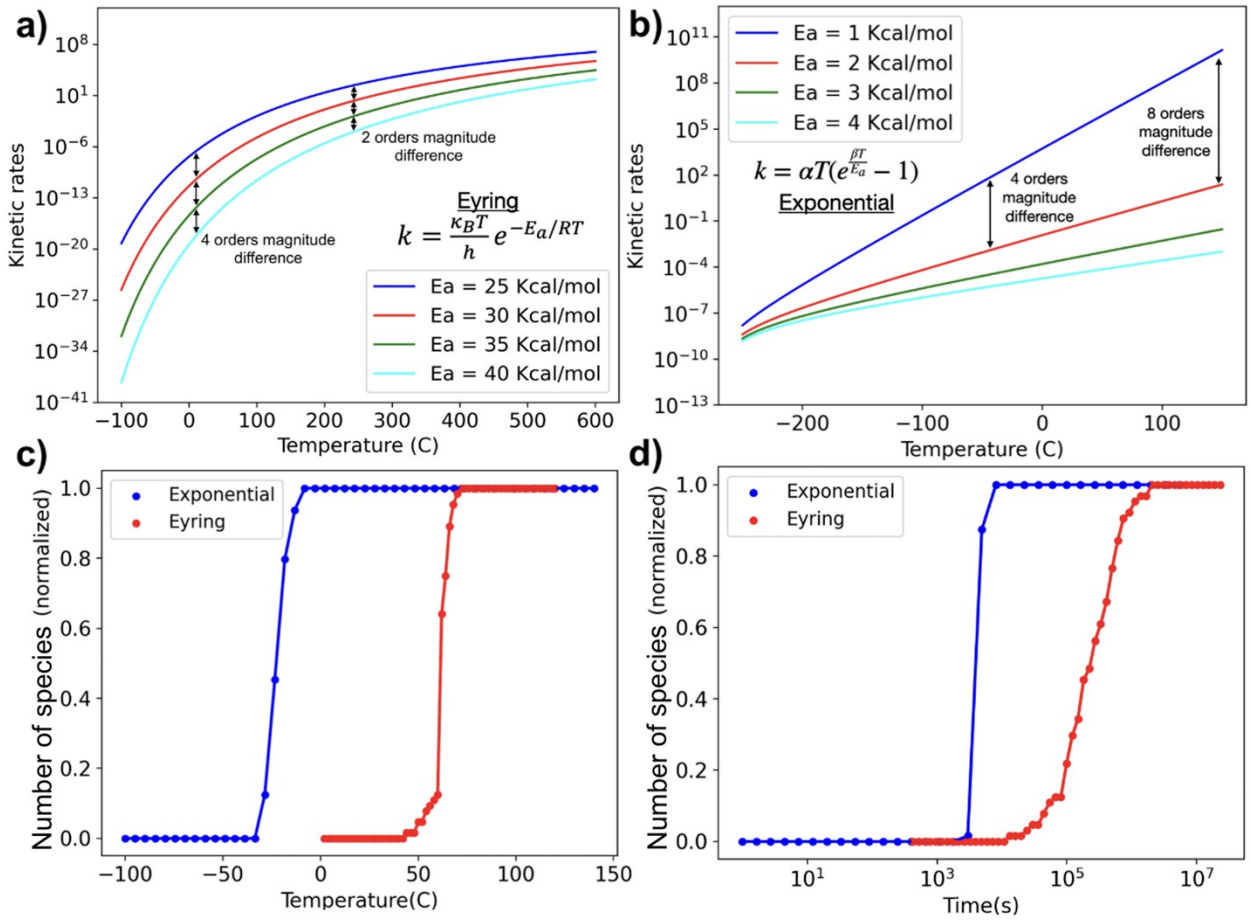

Figure S5. Robustness of combinatorial compression to alternative rate-constant parameterizations. (a) Reaction rates calculated by the Eyring formula versus temperature. Kinetic selectivity declines as T increases. (b) Reaction rates calculated by a non-physical formula that causes kinetic selectivity to increase with T. This artificial effect is inconsistent with the Eyring formula and is intended to verify the independence of our results on the Eyring formula. (c) The number of monomers plus dimers (normalized) versus T when rates are defined by the Eyring formula (red line) and when rates are defined by the alternative method (blue line). In both cases, a transition from compression to explosion is observed. (d) The number of monomers plus dimers (normalized) versus t when rates are defined by the Eyring formula (red line) and when they are defined by the alternative method (blue line). In both cases, transitions from compression to explosion are observed.

Figure S6. Schematic representation of free energy landscapes.

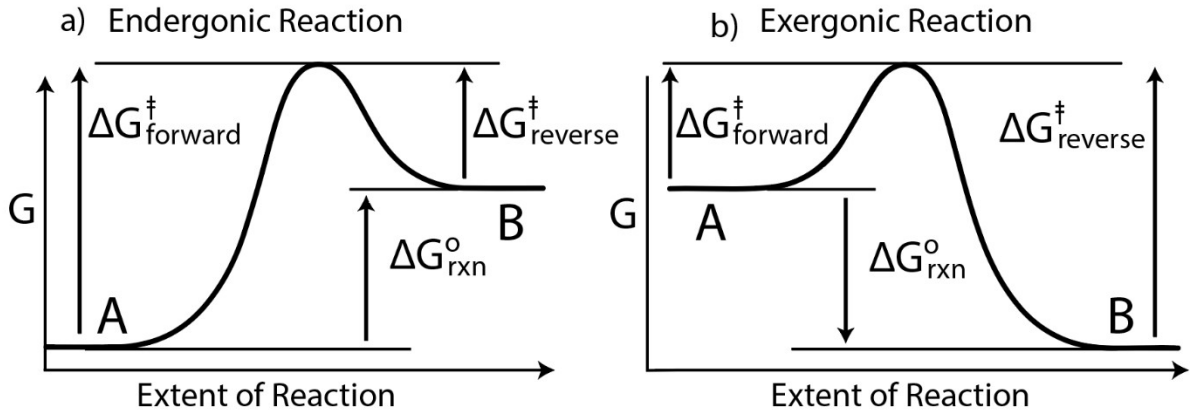

Figure S6. Schematic representation of free energy landscapes in a) an endergonic reaction and b) an exergonic reaction. $\Delta G^0{}_{rxn}$ is the standard change in the free energy of the reaction and $\Delta G^{\ddagger}$'s represents the forward and reverse free energies of activation.

References

1. P. W. Atkins, J. De Paula and J. Keeler, *Atkins' physical chemistry*, Oxford university press, 2023.
2. K. Matange, V. Rajaei, P. Capera-Aragones, J. T. Costner, A. Robertson, J. S. Kim, A. S. Petrov, J. C. Bowman, L. D. Williams and M. Frenkel-Pinter, Evolution of complex chemical mixtures reveals combinatorial compression and population synchronicity, *Nat. Chem.*, 2025, 1-8.