23

Free Energies of Reaction for Aqueous Glycine Condensation Chemistry at Extreme Temperatures

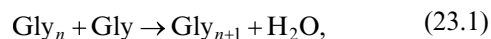**Matthew Kroonblawd[1] and Nir Goldman[1,2]**

ABSTRACT

We have performed high throughput quantum molecular dynamics simulations to determine the free energy surface for aqueous glycine condensation reactions at conditions similar to oceanic hydrothermal vents (1 g cm^{-3} and temperatures ranging from 400 K to 1000 K). Our simulations identify significant changes in the free energy surface topology and subsequent chemical reactivity with increasing temperature. We predict that temperatures at 400 K and below favor dipeptide formation, whereas higher temperatures facilitate the reverse hydrolysis reaction, with solvated glycine molecules showing greater stability. This change in favorability is correlated with a shift in the location and characteristics of specific reaction bottlenecks or barriers. Simultaneously, we observe that relative free energy barriers (total energy plus entropic contributions) for both condensation and hydrolysis reactions generally decrease with increasing temperature. Our results indicate that relatively modest temperatures near 400 K may best facilitate formation of oligoglycine molecules in oceanic systems related to the synthesis of life-building compounds.

23.1. INTRODUCTION

Polypeptide synthesis in prebiotic environments is thought to be a pathway for the formation of more complicated molecules needed for the origins of life (Barbier, Visscher, & Schwartz, 1993; Brack, 2007; Danger, Plasson, & Pascal, 2012). Condensation reactions involving the simplest protein-forming amino acid glycine (Gly) serve as an important model for understanding the formation of more complicated polypeptides and proteins (Borsook, 1953; Faisal et al., 2005; Imai, Honda, Hatori, Brack, et al., 1999; Imai, Honda, Hatori, & Matsuno, 1999; Kitadai, 2014; Lemke, Rosenbauer, & Bird, 2009; Sakata, Kitadai, & Yokoyama, 2010; Schreiner et al., 2009; Shock, 1992; Sugahara, & Mimura, 2014; Van Dornshuld, Vergenz, & Tschumper, 2014). Glycine has been recovered from cometary comas and meteorites and could have been delivered to ancient Earth through an impact event (Altwegg et al., 2016; Blank et al., 2001; Chyba & Sagan, 1992; Goldman et al., 2010; Martins et al., 2013; Sugahara & Mimura, 2014, 2015). Oligoglycine chains can be created during condensation reactions

$$\text{Gly}_n + \text{Gly} \rightarrow \text{Gly}_{n+1} + \text{H}_2\text{O}, \quad (23.1)$$

through formation of C–N peptide bonds and the elimination of water. Peptide bonds in oligoglycine are quite stable once formed, with a 100+ year half-life for the

[1] *Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California, USA*

[2] *Department of Chemical Engineering, University of California–Davis, California, USA*

reverse (hydrolysis) reaction at room temperature and pressure without a catalyst (Wolfenden, 2011). One possibility for the formation of polypeptides or precursor molecules on early Earth is the presence of extreme temperatures and/or pressures (Blank et al., 2001; Goldman et al., 2010; Goldman & Tamblyn, 2013; Martins et al., 2013; Sugahara & Mimura, 2014, 2015). Extreme conditions are a known means to access otherwise unlikely and slow chemical synthetic routes and can prompt the formation of new and exotic products and phases (Goldman et al., 2010; Goldman & Tamblyn, 2013; Imai, Honda, Hatori, Brack, et al., 1999; Imai, Honda, Hatori, & Matsuno, 1999; Koziol & Goldman, 2015; Lemke et al., 2009; Martins et al., 2013; Schreiner et al., 2009; Sugahara & Mimura, 2014, 2015). This could have conceivably involved cycling between high and low temperatures in submarine hydrothermal vents, which can exhibit temperature variations of up to ≈400 K, often reaching $T > 600$ K (Imai, Honda, Hatori, Brack, et al., 1999; Imai, Honda, Hatori, & Matsuno, 1999; Kitadai, 2014; Lemke et al., 2009; Shock, 1992; Schreiner et al., 2009; Tivey et al., 1995; Yanagawa et al., 1990). Individual prebiotic synthesis scenarios thus can span a wide range of relevant states, potentially resulting in large changes to chemical kinetics and equilibria for even simple reactions such as peptide oligomerization. However, determining the specific temperatures and pressures or thermodynamic paths that might favor oligoglycine synthesis through laboratory experiments alone can be slow and expensive.

Quantum-based molecular dynamics (QMD) coupled with enhanced sampling methods such as metadynamics (Laio & Parrinello, 2002) or umbrella sampling (Torrie & Valleau, 1974)) can provide an independent method for predicting prebiotic chemistry in these systems (Kroonblawd et al., 2018; Pietrucci & Saitta, 2015; Schreiner et al., 2009). Ab initio methods such as Kohn-Sham Density Functional Theory (Kohn & Sham, 1965) (DFT) can yield accurate information but are too computationally intensive to simultaneously probe a wide range of thermodynamic thermodynamic conditions. Force-matched semiempirical models (Goldman et al., 2018; Goldman et al., 2015; Kroonblawd et al., 2018) based on the density functional tight binding (DFTB) method (Elstner et al., 1998; Koskinen, & Mäkinen, 2009; Porezag et al., 1995) offer an efficient alternative that is orders of magnitude less computationally intensive while retaining a high degree of accuracy. The computational efficiency of DFTB allows for running many independent simulations concurrently to generate trajectories that can approach chemical equilibrium timescales (Cawkwell et al., 2015; Kroonblawd, & Goldman, 2018; Kroonblawd et al., 2018).

Very recently, we developed a force-matched DFTB model for glycine that was tailored to reproduce DFT predictions for aqueous glycine condensation reactions at $\rho = 1$ g cm^{-3} and $T = 300$ K (Kroonblawd, et al., 2018). In that case, the efficiency of DFTB was found to be critical for performing the multinanosecond long free energy (umbrella sampling) simulations needed to converge predictions for the free energy surface (energy plus entropy) of reaction. Here, we extend these efforts to sample temperatures relevant to potential hydrothermal vent synthesis (i.e. 400 K $\leq T \leq$ 1000 K). Accurate free energy calculations are essential to determine the temperature effects on condensed phase chemical reactivity, which can be significant at the elevated conditions discussed throughout this work. We then use a chemical analysis of the underlying trajectories to obtain clear connections between features on the reaction free energy surface and particular reaction intermediates.

23.2. METHODS

In this section, we detail the number of parameters and variables that were set in our calculations. In addition to brief description of the DFTB method, we include definitions of a reaction-path-specific set of variables that were necessary in order to efficiently monitor chemistry in such a high-dimensionality system, and the schedule of different biasing potentials that were used to determine the underlying free energy surfaces. We also describe simplified analytical tools that were employed to analyze the different types of carbon-nitrogen-oxygen (CNO)-containing backbones produced over the course of our simulations.

23.2.1. Simulation Details

Molecular dynamics (MD) simulations of aqueous glycine condensation were performed using the self-consistent charge DFTB method. DFTB (Elstner et al., 1998; Koskinen & Mäkinen, 2009; Porezag et al., 1995) is a semi-empirical quantum simulation approach that yields a high degree of computational efficiency while potentially retaining the accuracy of the computationally intensive Kohn-Sham DFT. The formalism for DFTB with self-consistent charges has been discussed in detail elsewhere (Aradi et al., 2007; Gaus et al., 2011; Goldman, 2015; Goldman et al., 2018; Goldman et al., 2015; Koskinen & Mäkinen, 2009). Briefly, the method assumes neutral, spherically symmetric charge densities on the atoms and expands the DFT Hamiltonian to second-order in charge fluctuations. The DFTB total energy is expressed as

$$E_{\text{DFTB}} = E_{\text{BS}} + E_{\text{Coul}} + E_{\text{Rep}}. \qquad (23.2)$$

Here, E_{BS} is the electronic band structure energy, which is evaluated within a tight-binding framework (Slater & Koster, 1954), E_{Coul} captures charge transfer between atoms, and E_{Rep} is an empirical term that accounts for ionic repulsion and Kohn-Sham double counting terms. Parameters for E_{BS} and E_{Coul} were taken from the mio-1-1 parameterization (available at http://www.dftb.org), a typical off-the-shelf parameter set for organic systems. Most of the generic E_{Rep} interactions were replaced with force-matched ones (Kroonblawd, et al., 2018) that were specifically tuned to accurately predict the glycine condensation free energy surface at 300 K and 1 g cm^{-3}. The specific parameterization used here is the equal-weight variant of the force-matched DFTB model that is described in detail in Kroonblawd, et al. (2018).

All simulations were performed using a cubic, three-dimensionally periodic simulation cell at constant density 1 g cm^{-3}. The simulation cell had a side length of 12.374 Å and contained two glycine molecules solvated in 55 H_2O molecules, which is equivalent to one diglycine molecule in 56 H_2O molecules. Simulation parameters follow our previous results for glycine condensation at $T = 300$ K. Deuterium masses were substituted for hydrogen masses and the electronic structure was evaluated without spin polarization at the Γ-point only. Umbrella sampling simulations were driven using the PLUMED 1.3 plugin (Bonomi et al., 2009) using collective coordinates that are described in detail in section 23.2.2.

DFTB-MD trajectories were integrated using Extended Lagrangian Born-Oppenheimer dynamics (Niklasson, 2008; Niklasson et al., 2009; Niklasson et al., 2006; Zheng et al., 2011) driven by the LAMMPS simulation software suite (Plimpton, 1995) with forces and stresses evaluated by the DFTB+ code (Aradi et al., 2007). The time step was set to 0.20 fs and isochoric-isothermal (i.e., NVT) sampling was performed with a Nosé-Hoover-style thermostat (Hoover, 1985; Nosé, 1984). The electronic structure was evaluated using four self-consistent charge cycles per time step and Fermi-Dirac thermal smearing (Mermin, 1965) with the electronic temperature set equal to the instantaneous ionic kinetic temperature.

The DFT-MD simulations discussed here were performed using Car-Parrinello dynamics (Car & Parrinello, 1985) driven by the Quantum Espresso ab initio software package (Giannozzi et al., 2009). Similar to previous work (Kroonblawd et al., 2018; Schreiner et al., 2009), the time step and electron mass were set to 0.145 fs (6.0 au) and 700 au, respectively, and both the ionic and electronic degrees of freedom were coupled to a Nosé-Hoover thermostat (Hoover, 1985; Nosé, 1984). We used the Perdew-Burke-Ernzerhof (Perdew et al., 1996) generalized-gradient approximation functional with ultrasoft pseudopotentials (Vanderbilt, 1990) and a 25 Ry planewave cutoff.

23.2.2. Free Energy Calculations

The initial glycine condensation reaction $2\text{Gly} \rightarrow \text{Gly}_2 + H_2O$ was sampled at selected temperatures using the same path collective coordinates that we previously applied to this reaction at $T = 300$ K in Kroonblawd et al. (2018). Path coordinates are particularly useful for describing complicated reactions such as glycine condensation because they significantly reduce the dimensionality of the sampled collective coordinate space. Two path coordinates are defined (Branduardi et al., 2007) in terms of the known reactant and product species as

$$s(t) = \frac{\sum_k k \exp\left(-\lambda D\left[\mathbf{R}(t), \mathbf{R}_k\right]\right)}{\sum_k \exp\left(-\lambda D\left[\mathbf{R}(t), \mathbf{R}_k\right]\right)}, \qquad (23.3)$$

and

$$z(t) = -\frac{1}{\lambda} \ln\left\{\sum_k \exp(-\lambda D[\mathbf{R}(t), \mathbf{R}_k])\right\}, \qquad (23.4)$$

where $D[\mathbf{R}(t), \mathbf{R}_k]$ is some appropriate *distance metric* to measure the displacement of an instantaneous atomic configuration $\mathbf{R}(t)$ from a sequence of reference configurations $\mathbf{R}_k$, and λ is a constant to be specified later. Coordinates $s(t)$ and $z(t)$ are respectively interpreted as measuring the progress along and deviation from the path defined by the sequence of $\mathbf{R}_k$. The distance metric used here,

$$D\left[\mathbf{R}(t), \mathbf{R}_k\right] = \sum_{i\beta}\left[C_{i\beta}(t) - C_{i\beta}^k\right]^2, \qquad (23.5)$$

is a function of the local coordination environment for particular atoms (Pietrucci & Saitta, 2015) that is quantified as

$$C_{i\beta}(t) = \sum_{j\in\beta} \frac{1 - \left[\dfrac{r_{ij}(t)}{d_{\alpha\beta}}\right]^6}{1 - \left[\dfrac{r_{ij}(t)}{d_{\alpha\beta}}\right]^{12}}. \qquad (23.6)$$

Here, $C_{i\beta}(t)$ measures the number of atoms of type β coordinated to atom i of type α at time t. The sum is taken over all atoms j of type β in the system, r_{ij} is the separation distance between atoms i and j, and $d_{\alpha\beta}$ are constants (typically <2 Å) set for each pair type.

Parameter choices for the path variables exactly follow our earlier work on aqueous glycine condensation. Equation (23.5) was evaluated for the four C and two N atoms that form the diglycine backbone, tracking the coordination of those atoms to all other C, N, O, and H atoms in the system. Two reference configurations define the reaction path, with endpoints at $k = 1$ for the two neutral aqueous glycine reactant molecules and $k = 2$ for the single aqueous neutral diglycine product molecule. The $C_{i\beta}^{k}$ corresponding to these endpoints were set using time averages obtained at 300 K from short unbiased MD simulations of neutral glycine and diglycine in explicit H_2O solvent performed using DFTB with the parameter set mio-1-1. The parameters $d_{\alpha\beta}$ were set to 1.8 Å for X–X pairs and 1.5 Å for X–H pairs, where atom type X is either C, N, or O. The parameter λ governs the surface topology and is usually chosen so that $\lambda D[\mathbf{R}_k, \mathbf{R}_{k+1}] \approx 2.3$, which for the present case is satisfied when $\lambda = 0.70$. Neutral glycine and diglycine configurations are approximately mapped to the (s, z) positions $(1.1, -0.1)$ and $(1.9, -0.1)$, and zwitterionic forms are in regions with $z > 0.2$.

Free energy surfaces were computed using umbrella sampling (Torrie, & Valleau, 1974), in which harmonic bias potentials

$$V_{\text{Bias}}(s,z) = \frac{1}{2}K_s(s - s_0)^2 + \frac{1}{2}K_z(z - z_0)^2, \quad (23.7)$$

were applied to the collective coordinates s and z. The free energy

$$F(s,z) = -k_\text{B}T \ln\langle \rho(s,z)\rangle, \quad\quad\quad (23.8)$$

was obtained within an arbitrary constant from an ensemble of independent biased simulations, each with different s_0 and z_0. We used the weighted histogram analysis method (WHAM) (Kumar et al., 1992; Roux, 1995) to compute the average unbiased distribution function $\langle \rho(s, z)\rangle$ from our ensemble of simulations. Histogram bin widths were set to 0.01 for both the s and z dimensions. The self-consistent WHAM equations were evaluated following the discussion in Roux (1995), with the maximum allowed per-cycle change in the free energy constant for each biased distribution set to 10^{-5} kcal mol^{-1}.

Based on our previous experience applying path coordinates to glycine condensation reactions, we were able to devise an efficient schedule of bias force constants and centroids to adequately sample s and z space. Sampling the s dimension is made somewhat complicated by a particularly large gradient along s in the free energy surface in $1.1 < s < 1.3$. We divided s-space into three separate zones covered by 23 simulations for a given z_0, with each zone having a particular grid spacing for s_0 and force constant K_s. The first zone had $s_0 = 1.050$ and 1.100 with $K_s = 1000$ kcal mol^{-1}. The

second zone covered the high-gradient region and had s_0 set in 0.025 increments in $1.125 \le s_0 \le 1.275$ with $K_s = 4000$ kcal mol^{-1}. The third zone had s_0 set in 0.050 increments in $1.300 \le s_0 \le 1.950$ with $K_s = 1000$ kcal mol^{-1}. Gradients along z were more uniform, so we set z_0 in 0.100 increments in $-0.200 \le z_0 \le +0.600$ with $K_z = 1000$ kcal mol^{-1}. Independent initial thermal configurations were prepared at $s_0 = 1.950$, $z_0 = 0.200$ for each temperature. From these starting points, we first expanded along s by seeding new umbrella sampling simulations with configurations taken from simulations with adjacent s_0 to yield a "slice" of the surface. This "slice" was then expanded along z, similarly taking the initial configuration for each new simulation from a simulation with adjacent s_0 and z_0 to yield the full surface.

23.2.3. Chemical Analysis

The diversity of chemical configurations sampled in our umbrella simulations was analyzed through a multipart approach. In the first part of the analysis, the trajectory for each independent simulation was passed through an in-house molecule recognition code to identify underlying stable and transient chemical structures sampled during the simulation. We removed all hydrogen atoms from our trajectories prior to the analysis to limit the search to only unique CNO backbone structures in order to simplify our analysis. Consequently, neutral and zwitterionic forms of glycine and diglycine are treated as equivalent. Molecules sampled in the umbrella simulations were identified using both a bond distance and lifetime criterion. Two atoms were considered connected in each instant of a given trajectory if they were separated by no more than 1.9 Å. The instantaneous connections were used to construct a set of all unique molecules, and those molecules were considered stable if their interatomic connections persisted for at least 40 fs. Molecules persisting for less than 40 fs were classified as transients. We note that our overall conclusions are not affected by these chosen parameter values. Time averages for the concentration of each unique CNO backbone species i were obtained for each umbrella sampling simulation j, with the normalizing volume taken to correspond to the simulation cell. We denote these concentrations as $c_{i,j}$, where each $c_{i,j}$ is a scalar with units mol L^{-1}.

The second part of the analysis combined the information from the independent trajectories to obtain a map of the concentration for each chemical species $c_i(s, z)$ as a function of the path variables s and z. Our approach is shown schematically for two simulations in Figure 23.1, although in practice the analysis was performed on the 207 independent simulations in each umbrella sampling ensemble. First, we mapped the concentrations $c_{i,j}$ to the path variable space (s, z) through the following heuristic. Each umbrella sampling simulation j has associated with it a biased distribution function $\rho_j(s, z)$ corresponding to the region of (s, z) space
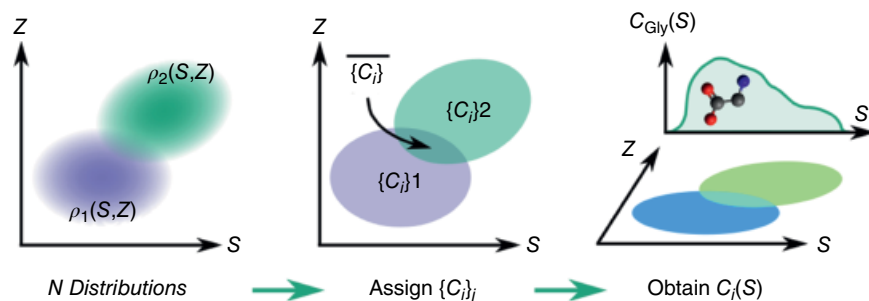
Figure 23.1 Schematic for mapping concentration information generated in two umbrella sampling trajectories to collective coordinate space. See electronic version for color representation of the figures in this book.

that it sampled. We assigned the scalar $c_{i,j}$ uniformly to the exact region of space spanned by $\rho_j(s, z)$ for each simulation j to obtain $c_{i,j}(s, z)$. (That is, $c_{i,j}[s, z] = c_{i,j}$ where $\rho_j[s, z] > 0$ and $c_{i,j}[s, z] = 0$ where $\rho_j[s, z] = 0$.) Overall maps for $c_i(s, z)$ were obtained by combining all of the $c_{i,j}(s, z)$ generated in the ensemble of umbrella sampling simulations. By design, there are regions of (s, z) space that were sampled by multiple simulations, so we took $c_i(s, z)$ to be the arithmetic average of all $c_{i,j}(s, z)$ from simulations sampling such regions. The concentrations $c_i(s, z)$ thus obtained did not vary significantly with respect to the z-coordinate, so we averaged $c_i(s, z)$ over z to obtain the concentration $c_i(s)$ as a function of s only. It should be noted that $\rho_j(s, z)$ and $c_i(s, z)$ are both computed on the same histogram grid employed for computing the the free energy surface $F(s, z)$.

23.3. RESULTS AND DISCUSSION

We begin this section with validation of our DFTB model through comparison to DFT-computed results. In particular, we observe close correspondence between the two methods for the energetics of the reactants (aqueous glcyine) and products (diglycine). We also consider the potential difficulties in converging condensed phase free energy calculations in part due to the large manifold of hydrogen-bonding rearrangements that can occur. We then discuss the temperature dependence of our results and focus on the transition from an exergonic (spontaneous release of free energy) to an endergonic (absorption of free energy) process that occurs between 400 and 500 K. Finally, we determine the different chemical pathways and energetics that are present over the range of temperatures studied here.

23.3.1. Validation and Convergence of Our Calculations

In order to provide a computationally feasible validation of our DFTB model, we have computed the average potential energy difference (e.g. excluding entropy effects) between the reactants and products with comparison to results from DFT (Figure 23.2). Here, values are compared between our DFTB model and DFT results

for a system of two glycine molecules (E_{2Gly}) solvated in 55 H_2O molecules vs. a single diglycine molecule ($E_{Gly\cdot2}$) solvated in 56 H_2O molecules. These results were readily determined from single, unbiased NVT simulations run for 5 ps for each system at each temperature. Uncertainty in the average potential energy for each state was taken to be the standard deviation of the energy fluctuations, and those uncertainties were propagated when computing the average difference $E_{2Gly} - E_{Gly\cdot2}$. Both DFT and DFTB predict that the reactants and products are nominally isoenergetic over the entire temperature interval considered. Although DFTB consistently predicts the energy difference to be more negative than does DFT, our results agree within uncertainty. Similarities between the DFT and DFTB predictions lend confidence in the transferability of our DFTB model, which was originally developed for glycine condensation reactions at 300 K.

Free energy surfaces $F(s, z)$ for glycine condensation were computed as functions of the path coordinates s and z using DFTB and umbrella sampling simulations following the approach described in section 23.2.2. Calculations were performed over temperatures from 400 K $\leq T \leq$ 1000
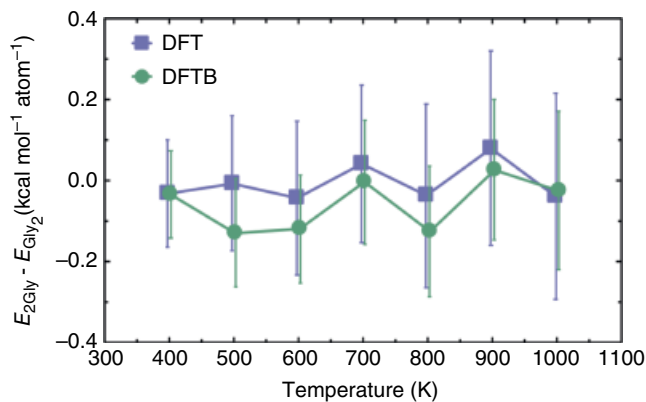

Figure 23.2 Average potential energy difference between reactant and product configurations computed as a function of temperature obtained using DFT and DFTB. Error bars correspond to propagated uncertainties. See electronic version for color representation of the figures in this book.

K in 100 K increments. A total of 207 independent 20 ps simulations were performed to generate the surface at each temperature (≈4.1 ns combined trajectory per surface), which corresponds to nearly 29 ns of combined trajectory overall. $F(s, z)$ was computed in the domain $1.0 \leq s \leq 2.0$ and $-0.2 \leq z \leq 0.6$ and captures both neutral and zwitterionic forms for the reactants and products. Convergence of our free energy calculations was assessed through a blocking analysis (Figure 23.3). Our blocked predictions for the free energy of reaction ΔF_{rxn} exhibit drift towards more positive values on the order of a few kcal mol^{-1} at 400, 800, 900, and 1000 K, though the predicted barriers are perhaps more steady. Despite the very long combined QMD simulation time, our analysis reveals that our predictions are quasi-steady over the last ten time blocks.

23.3.2. Temperature Dependence

Temperature dependence for ΔF_{rxn} and the reaction barrier were assessed by computing the free energy
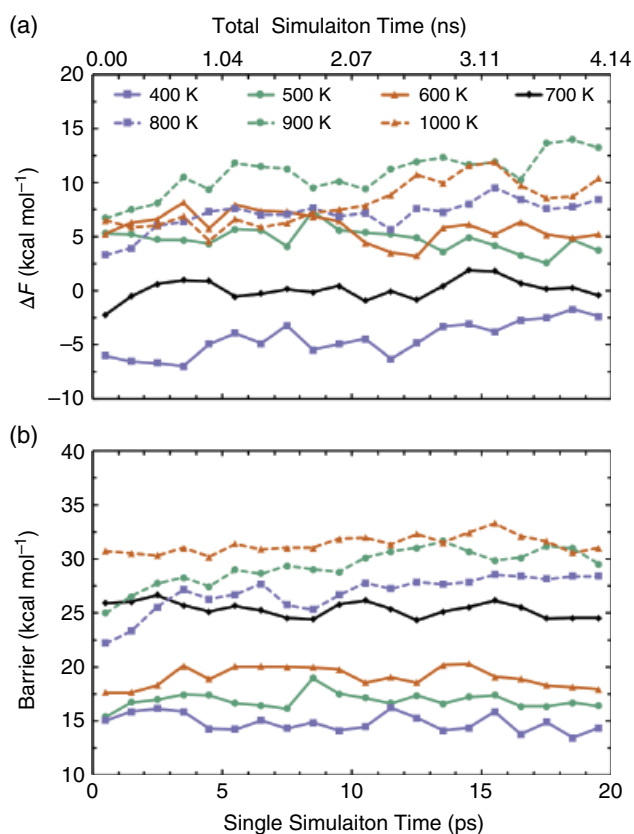

Figure 23.3 Blocking analysis for the predicted free energy difference ΔF_{rxn} and barrier height of the glycine condensation reaction $2Gly \rightarrow Gly_2 + H_2O$. Each data point in a given panel corresponds to a single $F(s, z)$ surface generated with the weighted histogram analysis method using biased histograms computed in nonoverlapping 1 ps windows. See electronic version for color representation of the figures in this book.

surfaces using the last 10 ps of each biased umbrella sampling trajectory. Figure 23.4 shows the temperature dependence for ΔF_{rxn} in panel (a) and for the forward and backward reaction barriers in units of kcal mol^{-1} and $k_B T$ in panels (b) and (c), respectively. Values at $T = 300$ K were previously obtained with this model in Kroonblawd et al. (2018). Uncertainties in our calculations were computed as the standard deviation of the mean obtained from the last 10 ps in the blocking analysis described above and were propagated (where appropriate) when isolating
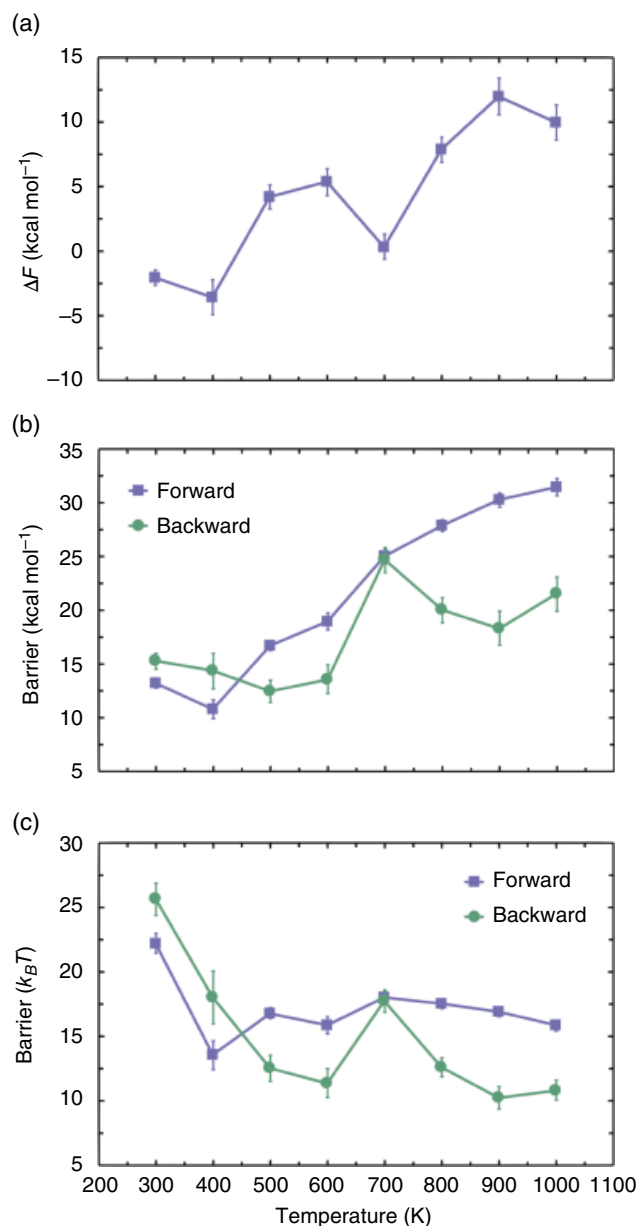

Figure 23.4 Temperature dependence of (a) the free energy difference ΔF_{rxn} and the reaction barrier in units of (b) kcal mol^{-1} and (c) $k_B T$. See electronic version for color representation of the figures in this book.

forward/backward reaction barriers. We note that due to the quasi-steady nature of the computed free energy surfaces, the reported uncertainties may not fully capture fluctuations anticipated in the limit of infinite simulation time.

Several qualitative trends are apparent in the isochoric temperature dependence of the free energy surface. The first trend is that ΔF_{rxn} generally increases with increasing temperature. A transition from a nominally exergonic ($\Delta F_{rxn} < 0$) process at low temperatures to an endergonic ($\Delta F_{rxn} > 0$) one at high temperatures occurs between 400 K and 500 K. Possible explanations for the abrupt decrease in ΔF_{rxn} at 700 K are discussed in section 23.3.4 in the context of the underlying chemical pathways. In terms of absolute values, the reaction barrier for both the forward and backward reaction generally increase with increasing temperature. However, in terms of $k_B T$ (the inherent kinetic energy of a system at a given temperature), the barrier decreases substantially going from 300 K to 400 K, after which it more or less forms a plateau. Commensurate with the transition from an exergonic to endergonic process, the barrier for the back reaction is lower than the barrier to the forward reaction at elevated temperatures.

Glycine condensation reactions have been studied under a diverse set of reaction conditions, with variations in temperature, pressure/density, pH, presence of contaminants or catalysts, and under dynamically changing states (Borsook, 1953; Faisal et al., 2005; Imai, Honda, Hatori, Brack, et al., 1999a; Imai, Honda, Hatori, & Matsuno, 1999b; Kitadai, 2014; Kroonblawd et al., 2018; Lemke et al., 2009; Radzicka & Wolfenden, 1996; Sakata et al., 2010; Schreiner et al., 2009; Shock, 1992; Sugahara & Mimura, 2014; Van Dornshuld et al., 2014). In drawing comparisons, it is important to consider that the present study focuses on the isochoric temperature dependence of glycine condensation at neutral pH in pure water. The average pressure predicted by DFTB at the highest temperature considered here is ≈15 kbar (due to thermal pressurization), although the small system size leads to significant fluctuations of a similar magnitude. With few exceptions, most experimental and theoretical treatments that specifically discuss the free energy difference between glycine and diglycine have found ΔF_{rxn} to be nominally zero within ≈5 kcal mol^{-1} across a wide range of conditions (Borsook, 1953; Danger et al. 2012; Kitadai, 2014; Lemke et al., 2009; Shock, 1992; Wolfenden, 2011). Thus, our results for ΔF_{rxn} are generally consistent in magnitude with the established literature.

The typical maximum bulk temperature and pressure at hydrothermal vents are $T_{max} \approx 660$ K and $P_{max} \approx 0.2$ kbar, which are both well below the maximums considered here. Several studies found ΔF_{rxn} to slightly decrease with increasing temperature at sub-kbar pressures (Kitadai, 2014; Lemke et al., 2009; Shock, 1992).

However, equation of state models predict that ΔF_{rxn} increases with increasing pressure under isothermal conditions from the saturation pressure up to at least 5 kbar (Shock, 1992). A subset of hydrothermal experiments performed at 0.2 kbar and 513 K ≤ T ≤ 573 K showed spontaneous and nearly complete hydrolysis of diglycine into glycine and other products (Faisal et al., 2005), consistent with the results presented here. Similar condensation reaction barrier heights to those computed here have been found in other quantum molecular dynamics calculations (Schreiner et al., 2009), computed over conditions of 300 K, 1.00 g cm^{-3} and 500 K, 0.85 g cm^{-3}. However, the absolute values reported here for the reverse or back reaction (hydrolysis) are significantly lower due to the very large exergonic free energy change ($\Delta F_{rxn} = -20$ kcal mol^{-1}) predicted in that study. In terms of $k_B T$ units, both studies predict that an increase in temperature will lower the effective barrier.

23.3.3. Exergonic-to-Endergonic Transition

We now focus on detailed results of the transition from a spontaneous process to one that requires heat absorption, between 400 and 500 K (Figure 23.5). Once again, free energy surfaces were computed using the last 10 ps of the umbrella trajectories that correspond to a nominal steady state. Maps of regions accessible at an energy equivalent to the minimum barrier height were generated by "filling" the free energy surface to the corresponding energy level. No significant differences in chemical concentrations were observed with respect to our z path coordinate, so we averaged results to obtain the concentration $c_i(s)$ as a function of s only. Our analysis yielded five unique CNO backbone species during the course of reactivity, including glycine **1**, diglycine **2**, a linear intermediate **3** that we previously identified at 300 K, and two different cyclic intermediates **4** and **5**. The concentration of **1** is roughly twice that of **2** because there are two glycine reactant molecules and a single diglycine product molecule. At 400 K, the general sequence for the reaction is $1 \rightarrow 3 \rightarrow 4 \rightarrow 2$. A similar sequence is seen at 500 K, but with small quantities of **5** detected early in the transition to diglycine. The elimination of water occurs during the transition from **3** to **4**.

The two starkest changes in the free energy surface with temperature are the location of the absolute minimum and the overall reaction barrier (or bottleneck) in relation to the underlying chemistry. At 400 K, the free energy minimum corresponds to zwitterionic diglycine and the barrier is near the onset of the transition from glycine to intermediate **3**. In contrast, at 500 K, the free energy minimum of the reaction is neutral glycine and the barrier corresponds to the elimination of water from **3** and the subsequent transition to a mixture of **2** and **4**. Another key difference is that the
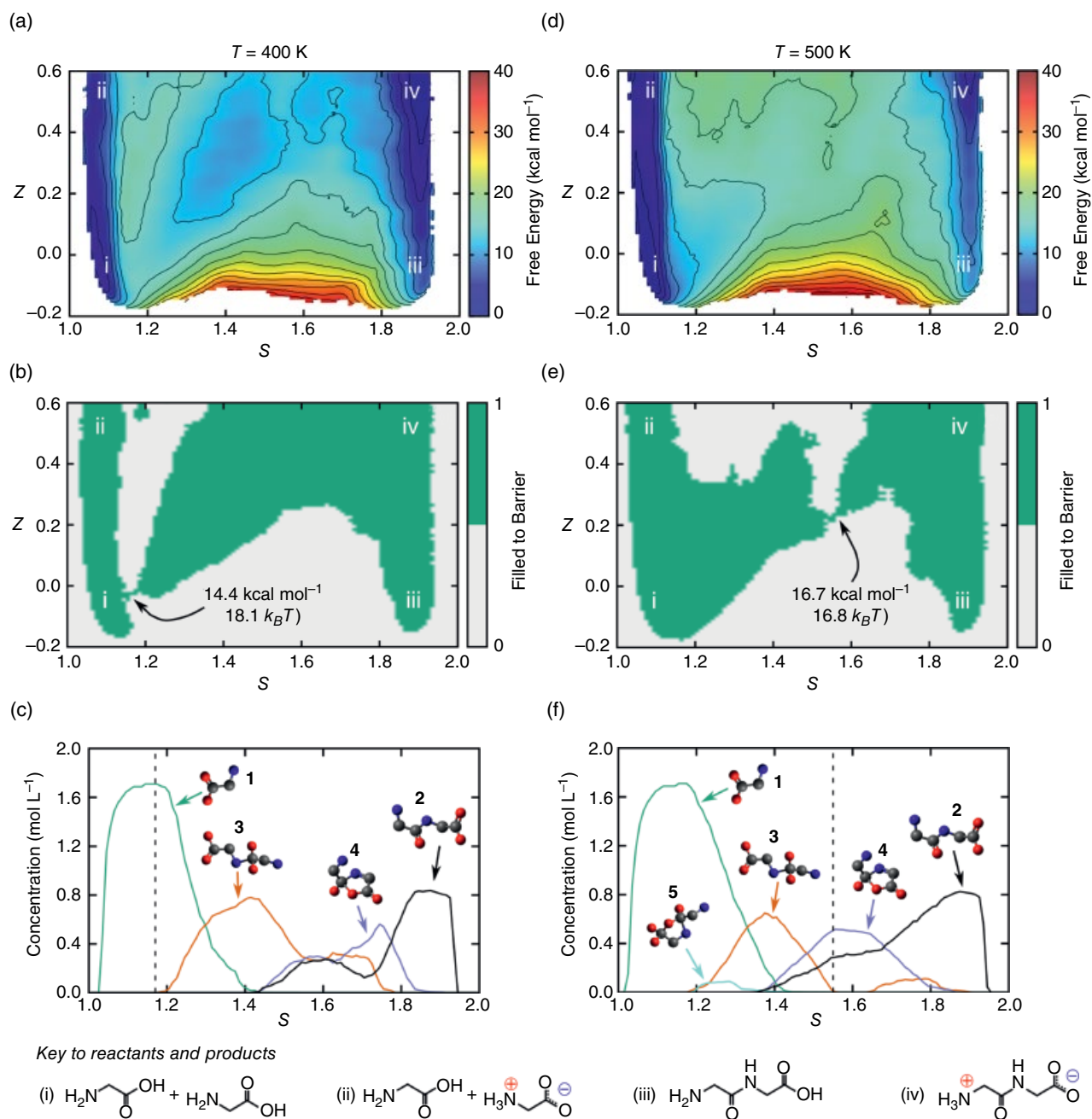
Figure 23.5 Comparison of free energy surface characteristics at 400 K and 500 K. The first column corresponds to 400 K and shows (a) the free energy surface $F(s, z)$, (b) approximate minimum free energy barrier heights, and (c) the concentrations $c_i(s)$ of unique CNO backbone species. The second column shows those same properties at 500 K in panels (d) through (f). Contour lines on $F(s, z)$ are drawn in 2.5 kcal mol^{-1} increments, and locations on $F(s, z)$ that correspond to particular neutral and zwitterionic reactant and product species are indicated by roman numerals. The s-coordinate for the minimum barrier location (or bottleneck) is demarcated in the concentration plots with vertical dashed lines. Atom colors are grey, blue, and red for C, N, and O. See electronic version for color representation of the figures in this book.

400 K surface has an apparent free energy minimum corresponding roughly to intermediate **3** centered at approximately $s = 1.4$, $z = 0.4$. A second local maximum at $s \approx 1.6$ is close in value to the bottleneck at $s = 1.17$ and has similar underlying chemistry to the 500 K bottleneck. One possibility is that solute/solvent interactions are strong enough to stabilize **3** at 400 K but not at 500 K. Previous theoretical studies have shown that even modest increases

in temperature can significantly disrupt the hydrogen bonding network in liquid water (Sahle et al., 2013), thus affecting solubility and the stability of charged species.

23.3.4. Characteristics at Higher Temperatures

We assessed the the changes in the free energy landscape for glycine condensation in 100 K increments up to a maximum temperature of 1000 K. As mentioned, the maximum temperature at hydrothermal vents is typically near 700 K. The higher temperatures considered here approach conditions that are relevant to more extreme prebiotic scenarios, such as during cometary impacts. Free energy surfaces were computed using the last 10 ps of our umbrella sampling trajectories and are shown for each temperature in Figure 23.6. Although not shown, the product and reactant species (i–iv) identified in Figure 23.5 map to the same analogous regions of $F(s, z)$ as at lower temperatures.

A distinct transition in the relative free energies for the diglycine end product itself occurs as the temperature in increases from 500 K to 700 K (see Figure 23.5d for the 500 K surface). At 500 K, the zwitterionic form (iv) is ≈5 kcal mol^{-1} lower in free energy than the neutral form (iii). When the temperature is increased to 600 K, these two forms have practically the same free energy and there is no appreciable barrier to interchange between them. Finally, at 700 K the neutral form becomes more favorable, being >5 kcal mol^{-1} lower in free energy than the zwitterion. Concurrently, there is a subtle increase in the favorability of neutral glycine (i) over its zwitterionic form (ii) with increasing temperature. One possible cause for the apparent minimum in the global $\Delta F_{rxn}(T)$ previously noted at 700 K in Figure 23.4a could be due to these changes and the overall energetic favorability of neutral molecules. The apparent minimum in $\Delta F_{rxn}(T)$ is relatively shallow (≈5 kcal mol^{-1}), which is on the same order of magnitude as the relative differences in free energy between the neutral and zwitterionic forms. Above 700 K, the surface topology does not change substantially with increasing temperature, consistent with the approximately monotonic behavior of $\Delta F_{rxn}(T)$ in 700 K ≤ T ≤ 1000 K.

A chemical concentration analysis was performed on the biased sampling trajectories for each temperature in 600 K ≤ T ≤ 1000 K and is shown in Figure 23.7. The qualitative chemical transformation process ($1 \rightarrow 3 \rightarrow 4 \rightarrow 2$) and location of the bottleneck are similar to the results obtained at 500 K in Figure 23.5f. Perhaps the most significant outlier is the response at 800 K, which exhibits a significantly reduced concentration of linear intermediate **3** relative to the other temperatures. Given that **3** and cyclic intermediate **4** appear similar in terms of the biased collective coordinates, the apparent scatter in their relative concentrations as a function of temperature

is likely a consequence of limited sampling trajectory. Similarly, cyclic intermediate **5** is present at both 500 K and for 700 K ≤ T ≤ 1000 K, but it is absent from our simulations at 600 K. When present, the concentration of **5** is
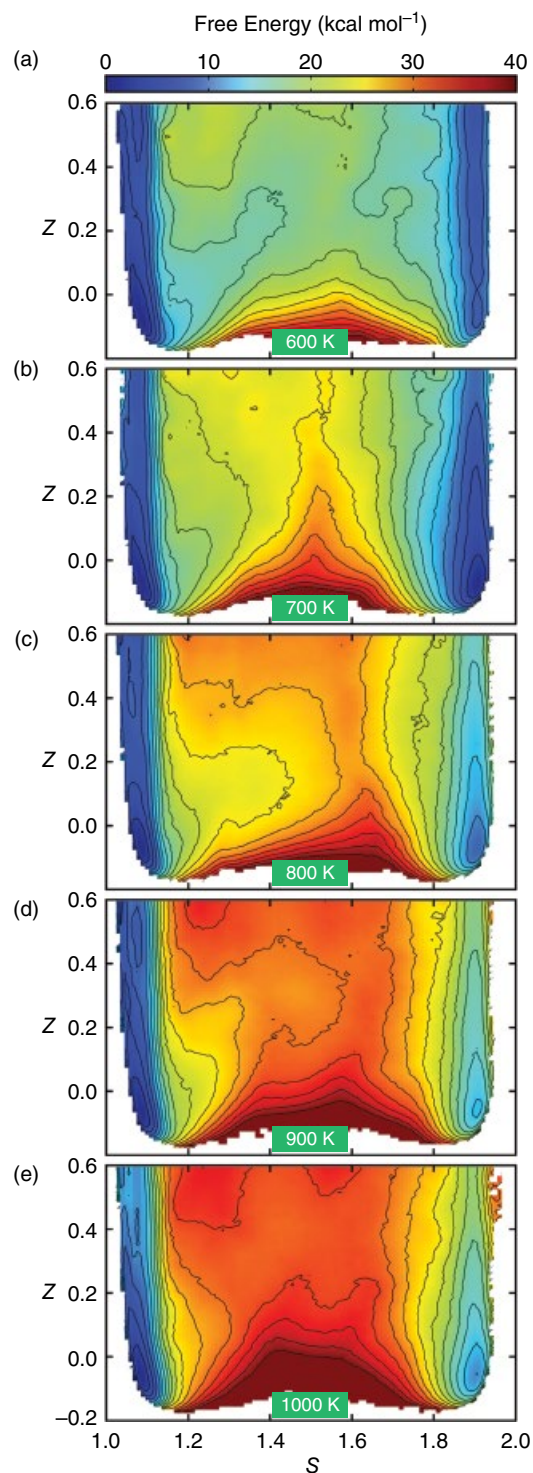

Figure 23.6 Free energy surfaces for glycine condensation at temperatures in 600 K ≤ T ≤ 1000 K. See electronic version for color representation of the figures in this book.
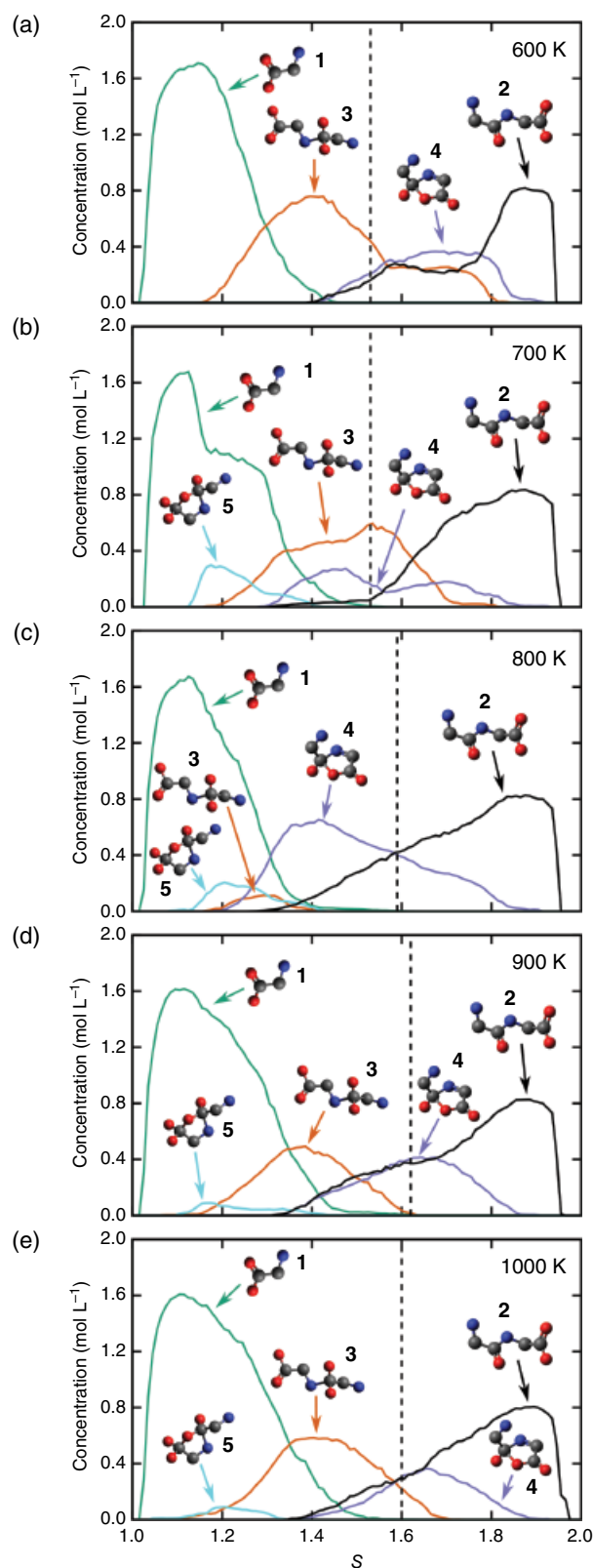
Figure 23.7 Concentration analysis of CNO backbone species at temperatures in 600 K ≤ T ≤ 1000 K. See electronic version for color representation of the figures in this book.

generally quite small compared to the other intermediates. However, the combined trajectory length generated here is substantially longer than would presently be obtainable with a purely ab initio approach such as DFT.

23.4. CONCLUSIONS

We have performed quantum-based molecular dynamics simulations to predict free energy surfaces (i.e. the manifold of temperature effects on chemical reactivity) for aqueous glycine condensation reactions at 1.0 g cm^{-3} and at temperatures ranging from 400 K up to 1000 K. Diglycine (glycylglycine) is predicted to be lower in free energy than glycine at ≈400 K and below, but higher temperatures are predicted to promote hydrolysis to form glycine as the thermodynamically preferred product. Our calculations show that the changes in the relative free energies between glycine and diglycine are likely entropic (temperature driven) in origin. In terms of a factor of k_BT (e.g. normalizing energies relative to the system temperature), the free energy barriers for the forward (condensation) and reverse (hydrolysis) reaction are predicted to decrease substantially with increasing temperature, promoting overall chemical reactivity. Chemical analysis of the underlying trajectories reveals that the nature of our determined reaction barrier changes between 400 K and 500 K, coinciding with an exergonic to endergonic transition for the total reaction. At 400 K, it is likely that hydrogen-bonding networks can help stabilize the formation of different charged intermediates, thus facilitating diglycine synthesis. In contrast, these solvent-mediated effects are less prevalent at higher temperatures, promoting hydrolysis of the diglycine peptide. Our results indicate that elevated temperatures near 400 K might best promote facile formation of peptide bonds, while more extreme temperatures (and pressures) may result in the potential destruction of these types of life-building molecules. Our detailed chemical free energy analysis can be useful for developing coarse-grain models such as equation of state and kinetic simulation methods for larger scale computations. Our overall goal with these efforts is to provide a "bottom-up" approach for understanding organic chemistry under extreme conditions, where there is a strong need for knowledge of the kinetics and thermodynamics of chemical bond formation and breaking over extended periods of time

REFERENCES

Altwegg, K., Balsiger, H., Bar-Nun, A., Berthelier, J.-J., Bieler, A., Bochsler, P., et al. (2016). Prebiotic chemicals—amino acid and phosphorus—in the coma of comet 67p/Churyumov-Gerasimenko. *Sci. Adv.*, *2*(5). Retrieved from http://advances.sciencemag.org/content/2/5/e1600285. doi: 10.1126/sciadv.1600285

Aradi, B., Hourahine, B., & Frauenheim, T. (2007). Dftb+, a sparse matrix-based implementation of the dftb method. *J. Phys. Chem. A*, *111*(26), 5678–5684. Retrieved from http://dx.doi.org/10.1021/jp070186p (PMID: 17567110). doi: 10.1021/jp070186p

Barbier, B., Visscher, J., & Schwartz, A. W. (1993). Polypeptide-assisted oligomerization of analogs in dilute aqueous solution. *J. Mol. Evol.*, *37*(6), 554–558. Retrieved from https://doi.org/10.1007/BF00182740. doi: 10.1007/ BF00182740

Blank, J. G., Miller, G. H., Ahrens, M. J., & Winans, R. E. (2001). Experimental shock chemistry of aqueous amino acid solutions and the cometary delivery of prebiotic compounds. *Orig. Life Evol. Biosph.*, *31*(1), 15–51. Retrieved from https://doi.org/10.1023/A:1006758803255 doi: 10.1023/A:1006758803255

Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., et al. (2009). PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.*, *180*(10), 1961–1972. Retrieved from http://www.sciencedirect.com/science/article/pii/S001046550900157X (PLUMED is available at http://www.plumed.org/). doi: https://doi.org/10.1016/j.cpc.2009.05.011

Borsook, H. (1953). Peptide bond formation. *Adv. Protein Chem.*, *8*, 127–174. Retrieved from http://www.sciencedirect.com/science/article/pii/S0065323308600923. doi: http://dx.doi.org/10.1016/S0065-3233(08)60092-3

Brack, A. (2007). From interstellar amino acids to prebiotic catalytic peptides: A review. *Chem. Biodivers.*, *4*(4), 665–679. Retrieved from http://dx.doi.org/10.1002/cbdv.200790057. doi: 10.1002/cbdv.200790057

Branduardi, D., Gervasio, F. L., & Parrinello, M. (2007). From a to b in free energy space. *J. Chem. Phys.*, *126*(5), 054103. Retrieved from http://dx.doi.org/10.1063/1.2432340, doi: 10.1063/1.2432340

Car, R., & Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, *55*, 2471–2474. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.55.2471 doi: 10.1103/ PhysRevLett.55.2471

Cawkwell, M. J., Niklasson, A.M.N., & Dattelbaum, D. M. (2015). Extended lagrangian Born-Oppenheimer molecular dynamics simulations of the shock-induced chemistry of phenylacetylene. *The Journal of Chemical Physics*, *142*(6), 064512. Retrieved from https://doi.org/10.1063/1.4907909. doi: 10.1063/1.4907909

Chyba, C., & Sagan, C. (1992). Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: An inventory for the origins of life. *Nature*, *355*, 125–132. Retrieved from http://dx.doi.org/10.1038/355125a0. doi: 10.1038/355125a0

Danger, G., Plasson, R., & Pascal, R. (2012). Pathways for the formation and evolution of peptides in prebiotic environments. Chem. Soc. Rev., *41*, 5416–5429. Retrieved from http://dx.doi.org/10.1039/C2CS35064E. doi: 10.1039/C2CS35064E

Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., et al. (1998). Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B: Condens. Matter Mater. Phys.*, *58*, 7260–7268. Retrieved from https://link.aps.org/doi/10.1103/PhysRevB.58.7260. doi: 10.1103/PhysRevB.58.7260

Faisal, M., Sato, N., Quitain, A. T., Daimon, H., & Fujie, K. (2005). Hydrolysis and cyclodehydration of dipeptide under hydrothermal conditions. *Ind. Eng. Chem. Res.*, *44*(15), 5472–5477. Retrieved from http://dx.doi.org/10.1021/ie0500568. doi: 10.1021/ie0500568

Gaus, M., Cui, Q., & Elstner, M. (2011). DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.*, *7*, 931.

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., et al. (2009). Quantum espresso: A modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter*, *21*(39), 395502. Retrieved from http://stacks.iop.org/0953-8984/21/i=39/a=395502 (Quantum Espresso is available at http://www.quantum-espresso.org/).

Goldman, N. (2015). Multi-center semi-empirical quantum models for carbon under extreme thermodynamic conditions. *Chem. Phys. Lett.*, *622*, 128–136.

Goldman, N., Aradi, B., Lindsey, R. K., & Fried, L. E. (2018). Development of a multicenter density functional tight binding model for plutonium surface hydriding. *Journal of Chemical Theory and Computation*, *14*(5), 2652–2660. Retrieved from https://doi.org/10.1021/acs.jctc.8b00165. doi: 10.1021/acs.jctc.8b00165

Goldman, N., Fried, L. E., & Koziol, L. (2015). Using force-matched potentials to improve the accuracy of density functional tight binding for reactive conditions. *J. Chem. Theory Comput.*, *11*(10), 4530–4535. Retrieved from http://dx.doi.org/10.1021/acs.jctc.5b00742. doi: 10.1021/acs.jctc.5b00742

Goldman, N., Reed, E. J., Fried, L. E., Kuo, I.-F. W., & Maiti, A. (2010). Synthesis of glycine-containing complexes in impacts of comets on early earth. *Nat. Chem.*, *2*(11), 949–954. Retrieved from http://dx.doi.org/10.1038/nchem.827. doi: 10.1038/nchem.827

Goldman, N., & Tamblyn, I. (2013). Prebiotic chemistry within a simple impacting icy mixture. *J. Phys. Chem. A*, *117*(24), 5124–5131. Retrieved from http://dx.doi.org/10.1021/jp402976n. doi: 10.1021/jp402976n

Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys.*, *31*, 1695–1697. Retrieved from https://link.aps.org/doi/10.1103/PhysRevA.31.1695. doi: 10.1103/ PhysRevA.31.1695

Imai, E.-I., Honda, H., Hatori, K., Brack, A., & Matsuno, K. (1999). Elongation of oligopeptides in a simulated submarine hydrothermal system. *Science*, *283*(5403), 831–833. Retrieved from http://science.sciencemag.org/content/283/5403/831. doi: 10.1126/science.283.5403.831

Imai, E.-I., Honda, H., Hatori, K., & Matsuno, K. (1999). Autocatalytic synthesis of oligoglycine in a simulated submarine hydrothermal system. *Orig. Life Evol. Biosph.*, *29*, 249–259.

Kitadai, N. (2014). Thermodynamic prediction of glycine polymerization as a function of temperature and pH consistent with experimentally obtained results. *J. Mol. Evol.*, *78*(3), 171–187. Retrieved from https://doi.org/10.1007/s00239-014-9616-1. doi: 10.1007/s00239-014-9616-1

Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, *140*, A1133–A1138. Retrieved from https://link.aps.org/doi/10.1103/PhysRev.140.A1133. doi: 10.1103/PhysRev.140.A1133

Koskinen, P., & Mäkinen, V. (2009). Density-functional tight-binding for beginners. *Comput. Mater. Sci.*, *47*(1), 237–253. Retrieved from http://www.sciencedirect.com/science/article/pii/S0927025609003036. doi: http://doi.org/10.1016/j.commatsci.2009.07.013

Koziol, L., & Goldman, N. (2015). Prebiotic hydrocarbon synthesis in impacting reduced astrophysical icy mixtures. *The Astrophysical Journal*, *803* (2), 91. Retrieved from http://stacks.iop.org/0004-637X/803/i=2/a=91.

Kroonblawd, M. P., & Goldman, N. (2018, May). Mechanochemical formation of heterogeneous diamond structures during rapid uniaxial compression in graphite. *Phys. Rev. B*, *97*, 184106. Retrieved from https://link.aps.org/doi/10.1103/PhysRevB.97.184106. doi: 10.1103/PhysRevB.97.184106

Kroonblawd, M. P., Pietrucci, F., Saitta, A. M., & Goldman, N. (2018). Generating converged accurate free energy surfaces for chemical reactions with a force-matched semiempirical model. *Journal of Chemical Theory and Computation*, *14*(4), 2207–2218. Retrieved from https://doi.org/10.1021/acs.jctc.7b01266. doi: 10.1021/acs.jctc.7b01266

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules: I. the method. *J. Comput. Chem.*, *13*(8), 1011–1021. Retrieved from http://dx.doi.org/10.1002/jcc.540130812. doi: 10.1002/jcc.540130812

Laio, A., & Parrinello, M. (2002). Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.*, *99*(20), 12562–12566. doi: 10.1073/pnas.202427399

Lemke, K. H., Rosenbauer, R. J., & Bird, D. K. (2009). Peptide synthesis in early earth hydrothermal systems. *Astrobiology*, *9*(2), 141.

Martins, Z., Price, M. C., Goldman, N., Sephton, M. A., & Burchell, M. J. (2013). Shock synthesis of amino acids from impacting cometary and icy planet surface analogues. *Nat. Geosci.*, *6*(12), 1045–1049. Retrieved from http://dx.doi.org/10.1038/ngeo1930 doi: 10.1038/ngeo1930

Mermin, N. D. (1965). Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, *137*, A1441–A1443. Retrieved from https://link.aps.org/doi/10.1103/PhysRev.137.A1441. doi: 10.1103/PhysRev.137.A1441

Niklasson, A.M.N. (2008). Extended Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.*, *100*, 123004. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.100.123004. doi: 10.1103/PhysRevLett.100.123004

Niklasson, A.M.N., Steneteg, P., Odell, A., Bock, N., Challacombe, M., Tymczak, C. J., et al. (2009). Extended lagrangian born–oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.*, *130*(21), 214109. Retrieved from http://dx.doi.org/10.1063/1.3148075. doi: 10.1063/1.3148075

Niklasson, A.M.N., Tymczak, C. J., & Challacombe, M. (2006, Sep). Time-reversible Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.*, *97*, 123001. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.97.123001. doi: 10.1103/PhysRevLett.97.123001

Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, *81*(1), 511–519. Retrieved from http://dx.doi.org/10.1063/1.447334. doi: 10.1063/1.447334

Perdew, J. P., Burke, K., & Ernzerhof, M. (1996, Oct). Generalized gradient approximation made simple. *Phys. Rev. Lett.*, *77*, 3865–3868. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.77.3865. doi: 10.1103/PhysRevLett.77.3865

Pietrucci, F., & Saitta, A. M. (2015). Formamide reaction network in gas phase and solution via a unified theoretical approach: Toward a reconciliation of different prebiotic scenarios. *Proc. Natl. Acad. Sci. U.S.A.*, *112*(49), 15030–15035. Retrieved from http://www.pnas.org/content/112/49/15030.abstract doi: 10.1073/pnas.1512486112

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.*, *117*(1), 1–19. Retrieved from http://www.sciencedirect.com/science/article/pii/S002199918571039X (LAMMPS is available at http://lammps.sandia.gov). doi: http://dx.doi.org/10.1006/jcph.1995.1039

Porezag, D., Frauenheim, T., Köhler, T., Seifert, G., & Kaschner, R. (1995). Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.*, *51*, 12947–12957. Retrieved from https://link.aps.org/doi/10.1103/PhysRevB.51.12947. doi: 10.1103/PhysRevB.51.12947

Radzicka, A., & Wolfenden, R. (1996). Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *Journal of the American Chemical Society*, *118*(26), 6105–6109. Retrieved from https://doi.org/10.1021/ja954077c. doi: 10.1021/ja954077c

Roux, B. (1995). The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.*, *91*(1), 275–282. Retrieved from http://www.sciencedirect.com/science/article/pii/001046559500053I. doi: http://dx.doi.org/10.1016/0010-4655(95)00053-I

Sahle, C. J., Sternemann, C., Schmidt, C., Lehtola, S., Jahn, S., Simonelli, L., et al. (2013). Microscopic structure of water at elevated pressures and temperatures. *Proceedings of the National Academy of Sciences*, *110*(16), 6301–6306.

Sakata, K., Kitadai, N., & Yokoyama, T. (2010). Effects of pH and temperature on dimerization rate of glycine: Evaluation

of favorable environmental conditions for chemical evolution of life. *Geochim. Cosmochim. Acta*, *74*(23), 6841–6851. Retrieved from http://www.sciencedirect.com/science/article/pii/S0016703710004813. doi: http://dx.doi.org/10.1016/j.gca.2010.08.032

Schreiner, E., Nair, N. N., & Marx, D. (2009). Peptide synthesis in aqueous environments: The role of extreme conditions on peptide bond formation and peptide hydrolysis. *J. Am. Chem. Soc.*, *131*(38), 13668–13675. Retrieved from http://dx.doi.org/10.1021/ja9032742. doi: 10.1021/ja9032742

Shock, E. L. (1992). Stability of peptides in high-temperature aqueous solutions. *Geochim. Cosmochim. Acta*, *56*(9), 3481–3491. Retrieved from http://www.sciencedirect.com/science/article/pii/001670379290392V. doi: http://dx.doi.org/10.1016/0016-7037(92)90392-V

Slater, J. C., & Koster, G. F. (1954). Simplified LCAO method for the periodic potential problem. *Phys. Rev.*, *94*, 1498–1524. Retrieved from https://link.aps.org/doi/10.1103/PhysRev.94.1498. doi: 10.1103/PhysRev.94.1498

Sugahara, H., & Mimura, K. (2014). Glycine oligomerization up to triglycine by shock experiments simulating comet impacts. *Geochem. J.*, *48*(1), 51–62. doi: 10.2343/geochemj.2.0285

Sugahara, H., & Mimura, K. (2015). Peptide synthesis triggered by comet impacts: A possible method for peptide delivery to the early Earth and icy satellites. *Icarus*, *257*, 103–112. Retrieved from http://www.sciencedirect.com/science/article/pii/S0019103515001724. doi: http://dx.doi.org/10.1016/j.icarus.2015.04.029

Tivey, M. K., Humphris, S. E., Thompson, G., Hannington, M. D., & Rona, P. A. (1995). Deducing patterns of fluid flow and mixing within the tag active hydrothermal mound using mineralogical and geochemical data. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 12527–12555. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JB00610. doi: 10.1029/95JB00610

Torrie, G. M., & Valleau, J. P. (1974). Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem. Phys. Lett.*, *28*(4), 578–581. Retrieved from http://www.sciencedirect.com/science/article/pii/0009261474801090. doi: http://dx.doi.org/10.1016/0009-2614(74)80109-0

Vanderbilt, D. (1990). Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B: Condens. Matter Mater. Phys.*, *41*, 7892–7895. Retrieved from https://link.aps.org/doi/10.1103/PhysRevB.41.7892. doi: 10.1103/PhysRevB.41.7892

Van Dornshuld, E., Vergenz, R. A., & Tschumper, G. S. (2014). Peptide bond formation via glycine condensation in the gas phase. *J. Phys. Chem. B*, *118*(29), 8583–8590. Retrieved from http://dx.doi.org/10.1021/jp504924c. doi: 10.1021/jp504924c

Wolfenden, R. (2011). Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu. Rev. Biochem.*, *80*(1), 645–667. Retrieved from https://doi.org/10.1146/annurev-biochem-060409-093051. doi: 10.1146/annurev-biochem-060409-093051

Yanagawa, H., Kojima, K., Ito, M., & Handa, N. (1990, Sep 01). Synthesis of polypeptides by microwave heating: I. Formation of polypeptides during repeated hydration-dehydration cycles and their characterization. *J. Mol. Evol.*, *31*(3), 180–186. Retrieved from https://doi.org/10.1007/BF02109494. doi: 10.1007/BF02109494

Zheng, G., Niklasson, A. M. N., & Karplus, M. (2011). Lagrangian formulation with dissipation of Born-Oppenheimer molecular dynamics using the density-functional tight-binding method. *J. Chem. Phys.*, *135* (4), 044122. Retrieved from http://dx.doi.org/10.1063/1.3605303 doi: 10.1063/1.3605303