

A Reaction Network for Triglycine Synthesis under Hydrothermal Conditions

Shigenobu Mitsuzawa* and Tetsuyuki Yukawa[1]

Frontier Research System for Extremophiles, Japan Marine Science and Technology Center, 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061

[1]Coordination Center for Research and Education, Graduate University for Advanced Studies, Shonan Village, Hayama, Kanagawa 240-0193

Received September 30, 2003; E-mail: shigenobum@jamstec.go.jp

It has been hypothesized that chemical evolution leading to the origin of life might have occurred in hydrothermal environments on primitive Earth. To examine this hypothesis, we investigated how the polymerization of amino acids proceeds under high-temperature and high-pressure conditions. We investigated a reaction network consisting of glycine and oligoglycines up to trimer, and the condensation/hydrolysis reactions among these molecules. We determined the rate constants of these reactions in experiments employing a flow reactor at 200 °C and 25 MPa. We found that two condensation reactions of glycine, which yield diglycine and diketopiperazine as products, have larger equilibrium constants under these conditions than at 25 °C. This result supports the hypothesis that hydrothermal conditions are thermodynamically favorable for chemical evolution. We also found that triglycine formation is mediated by diketopiperazine at 200 °C and 25 MPa. This implies that diketopiperazine acts as an important intermediate in the polymerization process of amino acids, which might have occurred in hydrothermal environments on primitive Earth.

It has been hypothesized that life might have originated in high-temperature (probably higher than 100 °C) and high-pressure (probably greater than 20 MPa) seas, similar to the present submarine hydrothermal environments.[1–6] Since these systems are rich in thermal energy, metal ions and reducing gases, they are considered to be ideal sites for abiogenetic synthesis of organic compounds. Imai et al.[7] reported the synthesis of triglycine (GGG) from glycine (G) in 30 min at 225 °C and 25 MPa and that of hexaglycine when copper ion was added. However, it should be noted that these extreme conditions also accelerate decomposition reactions. Qian et al.[8] conducted an experiment using 20 mM (1 M = 1 mol dm^{-3}) diglycine (GG) as the starting material at 220 °C and 27 MPa and found that the hydrolysis of the peptide bonds and decomposition (mainly deamination and decarboxylation) of the amino acids were so rapid that larger oligoglycines were not detected. Shock[9] predicted theoretically that thermodynamic equilibrium would shift to the polymerization of amino acids as the temperature increased. To date, to the best of our knowledge, no experimental study to confirm this prediction has been conducted.

In the present submarine hydrothermal environments, water circulates through the systems from a high-temperature zone (350–400 °C) to an intermediate zone and then to a low-temperature zone (ca. 0–2 °C). It has been proposed that organic monomers, such as amino acids, are produced in the high-temperature zone from gases, such as NH_3, CH_4, CO, H_2, and N_2.[6,10] Amino acids synthesized in this manner may polymerize as they are transported by the seawater circulation and cooled. We should therefore investigate which temperature area in the wide range of 0–400 °C is the most conductive to amino acid polymerization. Decomposition of various amino acids within a few minutes has been reported at temperatures above 220 °C.[8,11–13] Thus, lower temperatures appear to be appropriate for amino acid polymerization. In this study, we performed experiments on the polymerization of glycine at 200 and 160 °C and 25 MPa.

Diketopiperazines (DKPs) are easily formed in most experiments via prebiotic peptide formation employing condensing agents[14–21] or under hydrothermal conditions.[11,12,22,23] Their formation was considered to be an obstacle to peptide chain elongation beyond dipeptides.[24–26] However, Nagayama et al.[27] found that at 90 °C the reactions of DKPs with amino acids or peptides resulted in the formation and chain elongation of oligopeptides. The validity of this reaction should be investigated at much higher temperatures.

In this study, we performed experiments using six different starting materials, i.e., G, GG, DKP, GGG, a mixture of G and GG, and a mixture of G and DKP, in a flow reactor simulating hydrothermal systems at 200 °C and 25 MPa. We also performed experiments at 160 °C and 25 MPa using three different starting materials, i.e., G, GG, and DKP. From the resulting data, we examined the rate constants for all condensation and hydrolysis reactions among G, GG, DKP, and GGG at 200 °C and among G, GG, and DKP at 160 °C. Based on the rate constants obtained, we discuss the polymerization of G and other amino acids in prebiotic hydrothermal environments. We focus particularly on 1) the temperature dependence of the equilibrium constant for the condensation of G, and 2) the proposed reaction path to produce GGG from G via DKP.

Methods

Experimental. Experiments were performed using a flow

Published on the web May 6, 2004; DOI 10.1246/bcsj.77.965

reactor allowing high-temperature/high-pressure conditions simulating hydrothermal systems (Fig. 1).[28] Starting material solutions at room temperature and ambient pressure were injected using an HPLC pump (PU-1580, JASCO, Hachioji, Japan) into the reaction system through a Hastelloy tube (i.d. 0.5 mm). The high-temperature/high-pressure reaction chamber was made of Inconel (i.d. 10 mm × height 300 mm, with a volume of 23.6 mL). The temperature of the reactor was set at 200 or 160 °C. After passing through the reaction chamber, the solution was quenched by passage through the coiled Hastelloy tube bathed in ethylene glycol at 5 °C. At the end of this cooling tube, the temperature of the solution was about 20 °C. Pressure between the pump and the back-pressure regulator (SCF-Bpg, JASCO) was maintained at 25 MPa. After flowing ample amounts (75 mL, more than three times the volume of the high-temperature/high-pressure reaction chamber) of the solutions to reach the steady state, the solutions were sampled at the end of this cooling tube. We carried out several experiments for each starting material condition, varying the flow rate in the range of 0.3–10 $cm^3\ min^{-1}$. We calculated the reaction time under high-temperature/high-pressure conditions by dividing the volume of the reaction chamber by the flow rate. We employed six starting solutions at 200 °C and three starting solutions at 160 °C (Table 1). None of the solutions had any pH or redox control, or added salts. The pH values measured prior to heating are shown in Table 1. G and oligoglycines were purchased from Wako Pure Chemical Industries (Osaka, Japan).

Samples were analyzed with a JASCO HPLC apparatus employing a JASCO Finepak SIL C18S column. The mobile phase consisted of 95% phosphate buffer (10 mM KH_2PO_4 with pH maintained at 2.5 by adjusting the amount of H_3PO_4 added), 5% CH_3CN, and 10 mM $C_6H_{13}SO_3Na$. The flow rate of the mobile phase was 1.0 $cm^3\ min^{-1}$. The compounds were detected by monitoring the absorbance at 210 nm. Identification and quantification of G, GG, DKP, and GGG produced in the experiments were carried out by referring to the chromatograms of standard solutions of these compounds and the calibration lines of their peak areas to concentrations. The lower detection limit was about 1 μM for both G and oligoglycines. Figure 2 shows a chromatogram from an experiment at 200 °C employing a mixture of 100 mM G and 50 mM DKP as the starting materials with a flow rate of 1 $cm^3\ min^{-1}$. The separation of the compounds can be considered adequate for quantification. The peak of DKP is saturated in Fig. 2. In its actual quantification, we employed a 10-fold diluted sample solution. We also

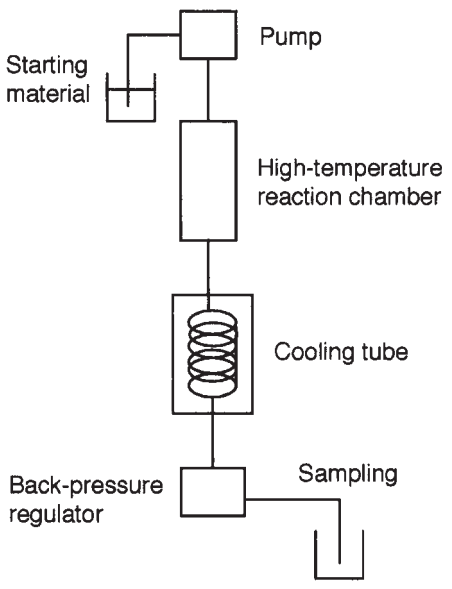

Fig. 1. Scheme of the flow reactor system.

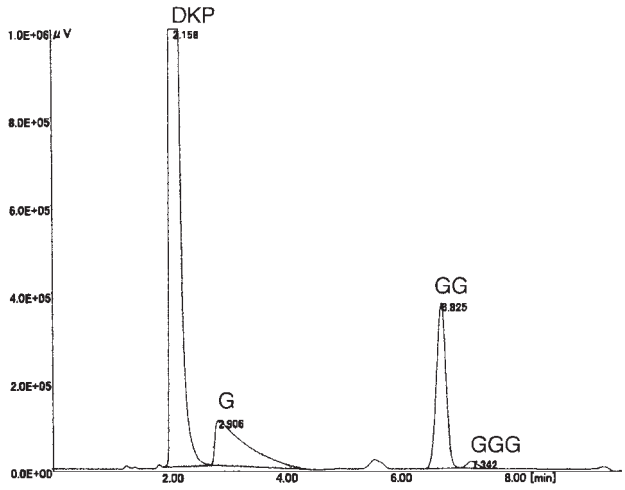

Fig. 2. Chromatogram resulting from the experiment at 200 °C employing a mixture of 100 mM G and 50 mM DKP as the starting material and a flow rate of 1 $cm^3\ min^{-1}$. Each peak is labeled with the compound name and retention time.

Table 1. Experimental Starting Material Conditions, Examined Rate Constants from the Experimental Data, and Molecular Species Most Sensitive to the Examined Rate Constants

		Starting material (Starting pH)	Examined rate constants	Most sensitive molecules
200 °C	(a)	100 mM G (6.1)	a_1, a_2	GG, DKP
	(b)	3.7 mM GG (5.7)	b_1, a_5	G, DKP
	(c)	3.1 mM DKP (5.8)	b_2, b_5	G, GG
	(d)	50 mM GGG (5.6)	b_4, b_3	GGG, GG
	(e)	100 mM G + 100 mM GG (5.7)	a_3	GGG
	(f)	100 mM G + 50 mM DKP (5.9)	a_4	GGG
160 °C	(a)	500 mM G (6.1)	a_1	GG, DKP
	(b)	10 mM GG (5.7)	b_1, a_5	G, DKP
	(c)	10 mM DKP (5.7)	b_5	GG

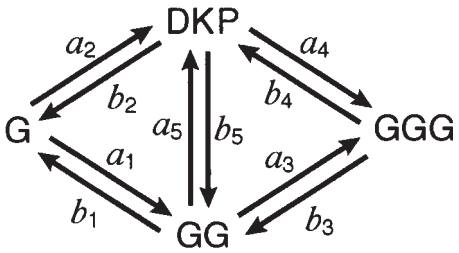

Fig. 3. Condensation and hydrolysis reaction network including G, GG, DKP, and GGG. The pair of a_1 and b_1 represents $G + G \rightleftarrows GG + H_2O$, a_2 and b_2: $G + G \rightleftarrows DKP + 2H_2O$, a_3 and b_3: $G + GG \rightleftarrows GGG + H_2O$, a_4 and b_4: $G + DKP \rightleftarrows GGG$, and a_5 and b_5: $GG \rightleftarrows DKP + H_2O$.

diluted the other sample solutions appropriately for the acquisition of each peak area.

Numerical Model. For the analysis of the experimental data, we consider five pairs of condensation and hydrolysis reactions including G, GG, DKP, and GGG:

$$G + G \rightleftarrows GG + H_2O \ (a_1, b_1), \tag{1}$$

$$G + G \rightleftarrows DKP + 2H_2O \ (a_2, b_2), \tag{2}$$

$$G + GG \rightleftarrows GGG + H_2O \ (a_3, b_3), \tag{3}$$

$$G + DKP \rightleftarrows GGG \ (a_4, b_4), \tag{4}$$

and $$GG \rightleftarrows DKP + H_2O \ (a_5, b_5). \tag{5}$$

Each condensation reaction is denoted by a_i ($i = 1$–5) and each hydrolysis reaction by b_i ($i = 1$–5), which are shown in parentheses in the above reaction equations. We also denote the reaction rate constants by the symbols. Figure 3 shows these reactions in a graphic manner to help understand the connections among them. Here we note that two reaction pairs concerning DKP, $2G \rightleftarrows DKP$ and $G + DKP \rightleftarrows GGG$, must be included to explain the experimental data fully. The analytical results and the reaction mechanisms are discussed in subsequent sections.

The reaction rates are assumed to be first order in each of the reactants. The pair of reactions a_i and b_i ($i = 1, 2, 3$, or 4) is in the form of $G + A_i \rightleftarrows B_i$, where A_i and B_i represent the molecular species of reactants. For example, in the case of $i = 3$, A_i is GG and B_i is GGG. Assuming the reference state of the solutions and the unit activity of H_2O, the reaction rate of a_i ($i = 1, 2, 3$, or 4) is

$$v_{a_i} = a_i x_G x_{A_i}, \tag{6}$$

and that of b_i ($i = 1, 2, 3$, or 4) is

$$v_{b_i} = b_i x_{B_i}, \tag{7}$$

where x_G, x_{A_i}, and x_{B_i} denote the concentration of molecular species G, A_i, and B_i, respectively. The reaction rate of a_5 ($GG \rightarrow DKP$) is

$$v_{a_5} = a_5 x_{GG}, \tag{8}$$

and that of b_5 ($DKP \rightarrow GG$) is

$$v_{b_5} = b_5 x_{DKP}, \tag{9}$$

where x_{GG} and x_{DKP} denote the concentrations of GG and DKP, respectively. Using these reaction rates, the rate equation of each molecular species is expressed as

$$\frac{dx_G}{dt} = -2(v_{a_1} - v_{b_1}) - 2(v_{a_2} - v_{b_2}) - (v_{a_3} - v_{b_3}) - (v_{a_4} - v_{b_4}), \tag{10}$$

$$\frac{dx_{GG}}{dt} = (v_{a_1} - v_{b_1}) - (v_{a_3} - v_{b_3}) - (v_{a_5} - v_{b_5}), \tag{11}$$

$$\frac{dx_{DKP}}{dt} = (v_{a_2} - v_{b_2}) - (v_{a_4} - v_{b_4}) - (v_{a_5} - v_{b_5}), \tag{12}$$

and

$$\frac{dx_{GGG}}{dt} = (v_{a_3} - v_{b_3}) + (v_{a_4} - v_{b_4}). \tag{13}$$

By solving these simultaneous differential equations, we can obtain the time course of the concentration of each molecular species. Numerical simulations were done using Mathematica ver. 3.0 on a Win98 PC. Using the rate constants and the concentrations of molecular species j when the reaction network is at equilibrium, $\bar{x}_j$, the equilibrium constant for the pair of reactions a_i and b_i ($i = 1, 2, 3$, or 4) is expressed as

$$K_i = \frac{a_i}{b_i} = \frac{\bar{x}_{B_i}}{\bar{x}_G \bar{x}_{A_i}}, \tag{14}$$

and that for the pair of reactions and a_5 and b_5 is

$$K_5 = \frac{a_5}{b_5} = \frac{\bar{x}_{DKP}}{\bar{x}_{GG}}. \tag{15}$$

From these equations, the relations between the equilibrium constants for two reaction loops, composed of G–GG–DKP and GG–DKP–GGG, are

$$K_2 = K_1 K_5, \tag{16}$$

and

$$K_3 = K_5 K_4. \tag{17}$$

Results

Experimental Results. Figure 4 summarizes the experimental results at 200 °C and the results of the numerical simulations. Plots represent the experimental data, and solid and dashed lines are based on the numerical simulations using the rate constants discussed later. Some condensation reactions are shown in Fig. 4a, e, and f. Islam et al.[12] carried out the polymerization of 0.1 M G in 2 min at 200 °C and 25 MPa, which is identical to the conditions shown in Fig. 4a in our study, and obtained 1.78×10^{-5} M DKP, 2.4×10^{-6} M GG, 1.4×10^{-7} M GGG, and trace amounts of tetraglycine. Our result, shown in Fig. 4a, is in good agreement with the results of Islam et al. The lower detection limit of our HPLC apparatus was about 1 μM, so we were unable to detect the synthesis of GGG. When the G and GG mixture (Fig. 4e) and the G and DKP mixture (Fig. 4f) were used as the starting materials, GGG was synthesized by condensation of the two starting materials. When GG (Fig. 4b), DKP (Fig. 4c), and GGG (Fig. 4d) were used as the starting materials, primarily hydrolysis reactions were observed. Comparing Fig. 4b with Fig. 4c, it is notable that DKP was much more stable than GG. Figure 4d shows that GGG rapidly hydrolyzed into G and DKP.[29] We noted that in Fig. 4a, the case when G was the starting material, GG and

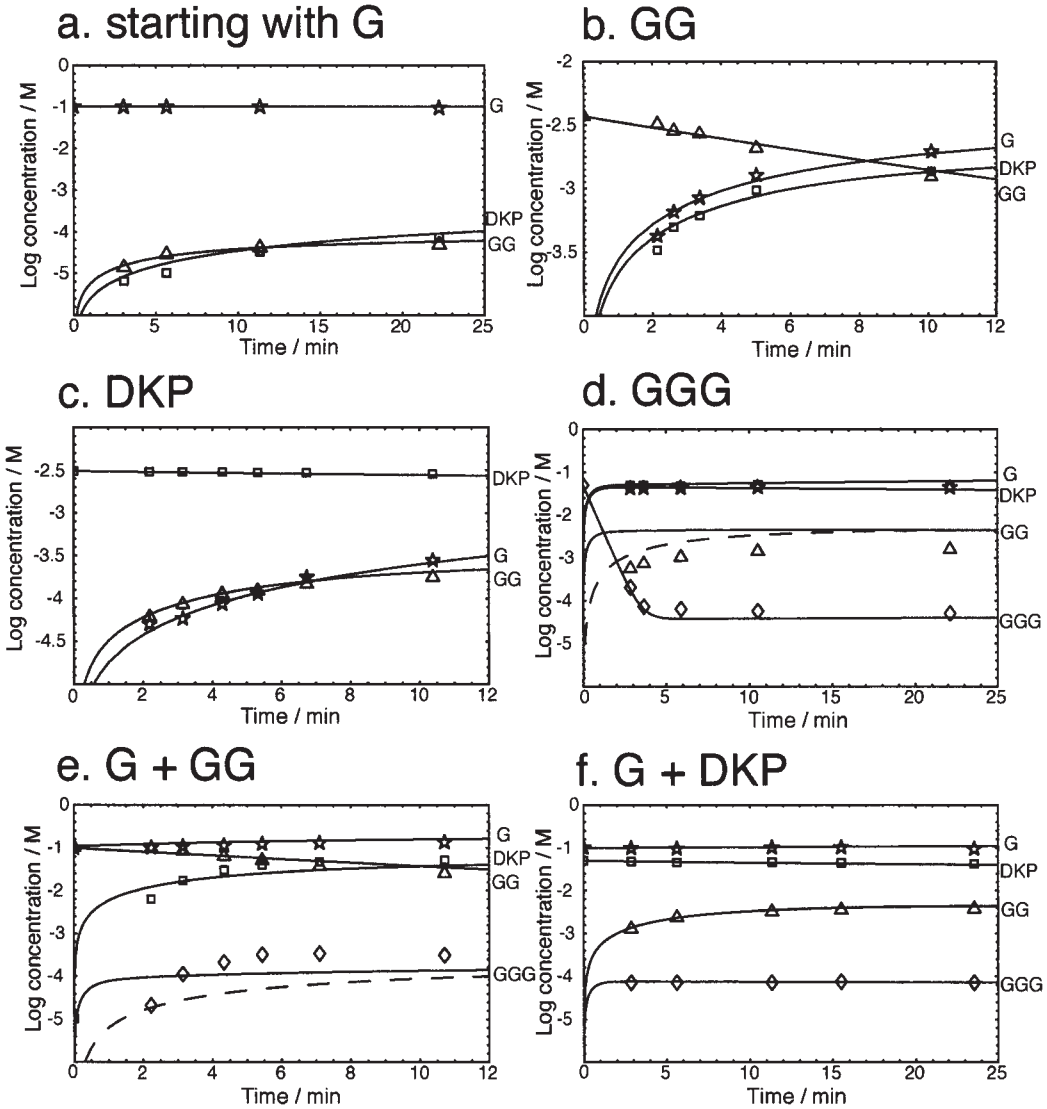

Fig. 4. Time courses of the reaction yields obtained from experiments at 200 °C and numerical simulations. Symbols are the experimental results: stars denote G, triangles GG, squares DKP, and diamonds GGG. Solid lines are the results of numerical simulations using $a_2 = 2.4 \times 10^{-4}$ $M^{-1}\,min^{-1}$, $(a_3/M^{-1}\,min^{-1}, b_3/min^{-1}) = (1.6 \times 10^{-2}, 1.7 \times 10^{-1})$, and the optimized values listed in Table 2 for the other rate constants. Dashed lines are the results of numerical simulations using $(a_3/M^{-1}\,min^{-1}, b_3/min^{-1}) = (0, 0)$ and keeping the other rate constants unchanged. Starting with (a) G, (b) GG, (c) DKP, (d) GGG, (e) a mixture of G and GG, and (f) a mixture of G and DKP.

DKP were synthesized at almost the same rate. When DKP hydrolyzes (Fig. 4c), the yield of G is almost the same as that of GG. These data cannot be explained fully by the two-step reaction path of G + G $\rightleftarrows$ GG and GG $\rightleftarrows$ DKP. We found that an additional reaction path, G $\rightleftarrows$ DKP, is necessary. This will be discussed further using numerical analyses. The total concentration of G calculated by $x_G + 2x_{GG} + 2x_{DKP} + 3x_{GGG}$ was constant during the time course of each experiment. This indicates that in the short reaction time examined, other side reactions, i.e., decomposition of G and oligomerization beyond the trimer level, were negligible. Figure 5 shows the experimental results and the numerical simulations at 160 °C. Unlike the results at 200 °C, a much slower production of DKP from G compared with that of GG was achieved through a two-step reaction, i.e., 2G $\rightarrow$ GG and GG $\rightarrow$ DKP. In the same manner, the much slower production of G from DKP compared with that of GG demonstrates that G is produced via GG.

Examination of Rate Constants. To examine the rate constants based on the experimental data, we executed the following procedures. First, for each experimental run, we calculated the initial reaction rates of the reactions as shown in Table 1 by linearly fitting the concentration of the concerned products against the reaction time and obtaining the slopes. Then, by inserting the calculated v values and the initial concentrations of the starting materials into Eqs. 6–9, we obtained rough values of the rate constants. Next, we fine-tuned the rate constants one by one. For each rate constant, we performed numerical simulations employing Eqs. 10–13 using the experimental starting material conditions (listed in Table 1) and the previously obtained rough values of the rate constants excluding the rate constant being examined. For example, to examine a_1, we employed the initial conditions of $x_G = 100$ mM and 0 mM for the

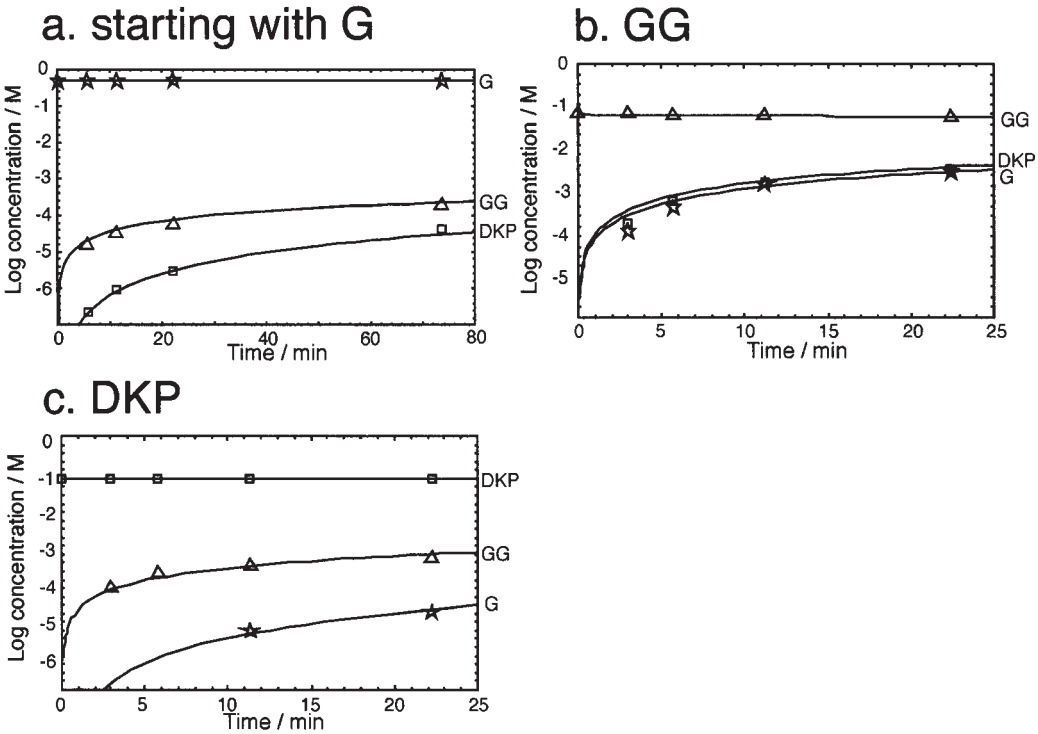

Fig. 5. Time courses of the reaction yields obtained from experiments at 160 °C and numerical simulations. Symbols are the experimental results: stars denote G, triangles GG, and squares DKP. Lines are the results of numerical simulations best fitting the experimental plots using the rate constants listed in Table 2. Starting with (a) G, (b) GG, and (c) DKP.

Table 2. Determined Rate Constants

	Rate constant	Optimized	Minimized RMSE	Lower limit	Upper limit
200 °C	a_1	6.0×10^{-4}	0.066	3.0×10^{-4}	12.0×10^{-4}
	b_1	3.9×10^{-2}	0.091	1.9×10^{-2}	8.5×10^{-2}
	a_2	1.2×10^{-4}	0.063	0.05×10^{-4}	3.8×10^{-4}
	b_2	2.7×10^{-3}	0.060	0.87×10^{-3}	6.6×10^{-3}
	a_3	1.6×10^{-2}	0.349	—	—
	b_3	0.0	0.283	—	5.4×10^{-3}
	a_4	3.1×10^{-2}	0.012	1.5×10^{-2}	6.4×10^{-2}
	b_4	1.9	0.192	1.6	2.3
	a_5	6.2×10^{-2}	0.079	2.9×10^{-2}	14.2×10^{-2}
	b_5	1.1×10^{-2}	0.024	0.54×10^{-2}	2.2×10^{-2}
160 °C	a_1	1.5×10^{-5}	0.032	0.6×10^{-5}	2.2×10^{-5}
	b_1	2.2×10^{-3}	0.166	1.2×10^{-3}	3.7×10^{-3}
	a_5	3.5×10^{-3}	0.116	1.7×10^{-3}	6.2×10^{-3}
	b_5	7.8×10^{-4}	0.074	4.0×10^{-4}	15.3×10^{-4}

$[a_i] = \mathrm{M}^{-1}\,\mathrm{min}^{-1}$ and $[b_i] = \mathrm{min}^{-1}$ for $i = 1$–4. $[a_5] = [b_5] = \mathrm{min}^{-1}$.

other concentrations of compounds. Repeating the simulations by varying the value of the concerned rate constant, we determined the value that minimizes the root mean square of the errors (RMSE) between the experimental and simulated logarithmic concentrations for the molecular species of which the time course is the most sensitive to the varying rate constant (shown in Table 1), e.g., GG for a_1. The rate constants optimized in this way and minimized RMSEs are shown in Table 2. Although we used the roughly obtained rate constants in the simulations, we found that the optimized values of the rate constants were not significantly affected by changing the values of the other rate constants. To evaluate the degree of confidence for each optimized rate constant, we determined its lower and upper limit values that give a RMSE of $\log 2$ (≈ 0.3), where the experimental concentration of the molecular species for which the time course is the most sensitive to the rate constant is larger than the simulated value by a factor of 2 or smaller by a factor of 0.5. We considered the criterion of an RMSE below $\log 2$ appropriate for this study because we examined the logarithmic concentrations of compounds and rather roughly estimated

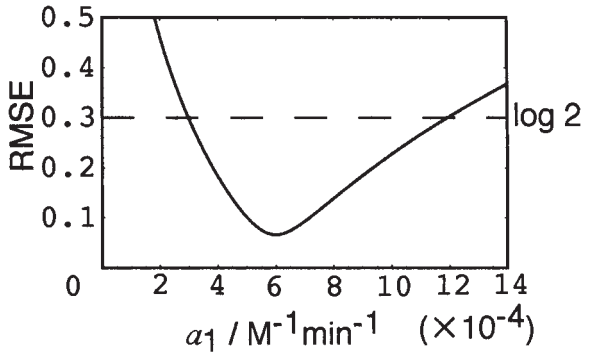

Fig. 6. RMSE of x_G vs a_1 at 200 °C.

the rate constants.

Results at 200 °C. Figure 6 shows the RMSE of x_G vs a_1 at 200 °C. The lower limit value of a_1 is smaller than its optimized value by a factor of 0.5, and the upper limit is larger by a factor of 2. The other rate constants have the same trend except for a_3 and b_3. Inserting the obtained a_1, b_1, b_2, a_5, and b_5 values into Eqs. 14–16 gives $a_2 = 2.4 \times 10^{-4}$ M^{-1} min^{-1}. The fact that this value is close to the independently optimized value, 1.2×10^{-4} M^{-1} min^{-1}, and in the range of an RMSE below log 2 supports the correct determination of the six rate constants concerning the reaction loop of G–GG–DKP. The results, shown in Table 2, that a_3 yields an optimized RMSE of 0.349 means that the experimental data on GGG in the case of G + GG as the starting materials cannot be fully explained by the simulation. The optimized b_3 value of 0 means that the synthesis of GG from GGG can be better explained by the two-step reaction path of GGG → G + DKP and then DKP → GG than by GGG → G + GG. The values of a_3 and b_3 can be calculated as dependent on each other by inserting the optimized value, together with a_4, b_4, a_5, and b_5, into Eqs. 14, 15, and 17. In this way, using $a_3 = 1.6 \times 10^{-2}$ M^{-1} min^{-1} gives $b_3 = 1.7 \times 10^{-1}$ min^{-1}, and $b_3 = 0$ min^{-1} yields $a_3 = 0$ M^{-1} min^{-1}. Figure 4 shows the results of the simulations using $a_2 = 2.4 \times 10^{-4}$ M^{-1} min^{-1}, (a_3/M^{-1} min^{-1}, b_3/min^{-1}) = (1.6×10^{-2}, 1.7×10^{-1}) or (0, 0), and the optimized values for the other rate constants. The simulations fit the experimental data very well except for GG in Fig. 4d and GGG in Fig. 4e, which are caused by the ambiguous determination of b_3 and a_3. We could not determine these values in this study. The results show that the simulation better fits the experimental data on GG in Fig. 4e when $b_3 = 0$ min^{-1} than when $b_3 = 1.7 \times 10^{-1}$ M^{-1} min^{-1}, and that the experimental and simulated GGG are on the same order in Fig. 4e when $a_3 = 0$ M^{-1} min^{-1}, although they are closer when $a_3 = 1.6 \times 10^{-2}$ M^{-1} min^{-1}, suggesting that the reaction path GGG ⇄ G + GG can be neglected.

The results of the G + DKP mixture (Fig. 4f) and GGG (Fig. 4d) as starting materials show that G + DKP ⇄ GGG proceeds effectively under hydrothermal conditions. More interestingly, 2G ⇄ DKP should proceed without GG as an intermediate. Figure 4a shows that GG and DKP are synthesized from G with almost the same initial reaction rates. The indispensability of 2G → DKP can be more explicitly understood by the simulation, taking $a_2 = 0$ and keeping the other rate constants optimized. In this case we would not be able to achieve the rapid synthesis of GG (Fig. 7a). Here, we should note that we must not arbitrarily make the values of a_1 or a_5 larger to increase DKP production because they are determined so as to fit the experimental data, i.e., GG production from G (Fig. 4a), or DKP production from GG (Fig. 4c). In the same manner, Figs. 7b–d shows the inconsistencies between the experimental data and

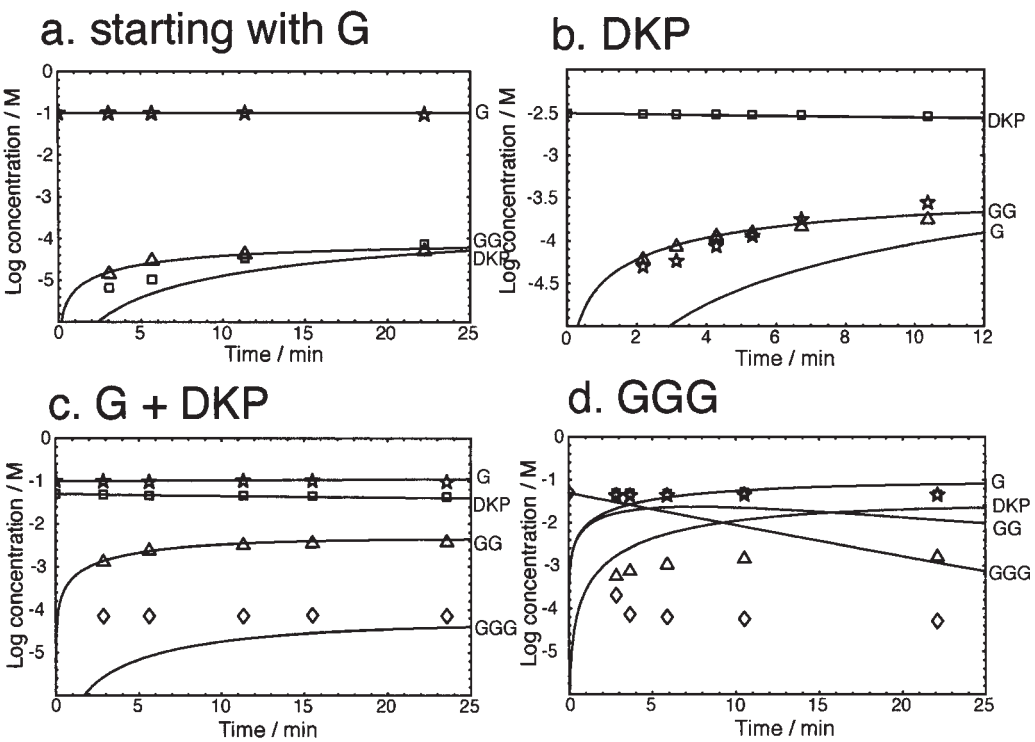

Fig. 7. Comparison of time courses of the reaction yields obtained from experiments at 200 °C and the simulation taking $a_2 = 0$ (a), $b_2 = 0$ (b), $a_4 = 0$ (c), and $b_4 = 0$ (d). The other rate constants were kept optimized.

the simulations when taking $b_2 = 0$, $a_4 = 0$, or $b_4 = 0$, respectively. These results confirm that G + G $\rightleftarrows$ DKP and G + DKP $\rightleftarrows$ GGG should proceed at 200 °C.

Results at 160 °C. We did not need the reaction pair G + G $\rightleftarrows$ DKP to fit the simulations to the experimental data well, and the rate constants a_1, b_1, a_5, and b_5 were determined using the above-mentioned procedures taking $a_2 = b_2 = 0$. Figure 5 shows the results of the simulations using the optimized rate constants (Table 2). The simulations fit the experimental data very well in all three runs.

Discussion

Equilibrium Constants. Shock[9] estimated the equilibrium constants of the condensation of amino acids at high temperatures and pressures employing thermodynamic calculations. Table 3 shows a comparison of the values K_1 and K_2 from our experiments and from the figures of Shock. At 200 and 160 °C, Shock's and our results are on the same order and can be regarded as in good agreement. This confirms that the high-temperature and high-pressure conditions expected in hydrothermal systems are thermodynamically favorable for condensation. However, because oligopeptides and amino acids decompose under these extreme conditions, oligopeptides cannot exist at high concentrations for a long time in hydrothermal systems without a continuous supply of amino acids. Another strategy to avoid decomposition could be quenching, e.g., injection of high-temperature water carrying oligopeptides into the surrounding cold water in hydrothermal systems.[7,30] Hydrothermal systems appear to be ideal sites for satisfying these conditions because of the inherent flow of seawater.

Synthesis of GGG via DKP. Our experimental data indicated that DKP production from G should be much faster than that via GG, and thus we added the G + G $\rightarrow$ DKP path to the reaction network model. However, the probability that two G molecules undertake two dehydration condensations to form DKP at exactly the same time must be almost nil. The actual mechanism of this reaction, expressed as G + G $\rightarrow$ DKP as an abbreviated form in this study, might be intermediated by *cis*-GG (Fig. 8a). At room temperature, *cis*-GG, which is much less stable than the *trans* form (Fig. 8b), is quantitatively negligible owing to its steric hindrance.[31] This is because we considered only one form of GG in this study. We implicitly referred to *trans*-GG with the notation of GG and presumed that the much more unstable *cis*-GG should exist stochastically only as a transition state in the reaction of *trans*-GG $\rightleftarrows$ DKP. However, this appears not to be the case at higher temperatures. Grathwohl and Wuthrich[32] measured the rate of *cis–trans* interconversion for X–Pro peptide bonds using NMR, where X represents several types of amino acids. Their data showed that the rate of *cis* to *trans* conversion is much larger than that of *trans* to *cis* at room temperature, and that they become closer to each other as the temperature increases up to 80 °C. This evidence suggests that *cis*-peptides are quantitatively very scarce at room temperature, but that they are much more stable at higher temperatures. This can be understood as the energetic instability caused by *cis* steric hindrance being overcome by thermal energy at high temperatures. Thus, taking the distinction of the GG isomers into account, we propose a hypothetical reaction network (Fig. 9). We replace GG in Fig. 3 with *trans*-GG and add *cis*-GG in the middle of the reaction paths, 2G $\rightleftarrows$ DKP and *trans*-GG $\rightleftarrows$ DKP. The condensation of two G molecules can potentially form *cis*-GG as well as *trans*-GG. The two isomeric GG can hydrolyze into G in reverse, and also undergo interconversions to each other. *cis*-GG synthesized directly from G or *trans*-GG could immediately undergo an intramolecular condensation between the free amino and carboxyl groups, which are located very close to each other. The reverse reaction, DKP $\rightarrow$ *cis*-GG should also proceed. The reaction path 2G $\rightleftarrows$ *cis*-GG does not actually proceed at lower temperatures, including 160 °C, because the energetic barrier is too high to overcome. At 200 °C, however, this path proceeds with substantial reaction rates and bypasses the DKP synthesis from

Table 3. Comparison of the Equilibrium Constants Obtained from This Study and from Shock[9]

	Source	200 °C	160 °C	25 °C
2G $\rightleftarrows$ GG (K_1)	This study	1.5×10^{-2}	6.8×10^{-3}	—
	Shock (1992)	4.0×10^{-2}	1.6×10^{-2}	2.0×10^{-3}
2G $\rightleftarrows$ DKP (K_2)	This study	8.9×10^{-2}	—	—
	Shock (1992)	2.5×10^{-2}	—	4.0×10^{-7}

$[K_1] = [K_2] = M^{-1}$.

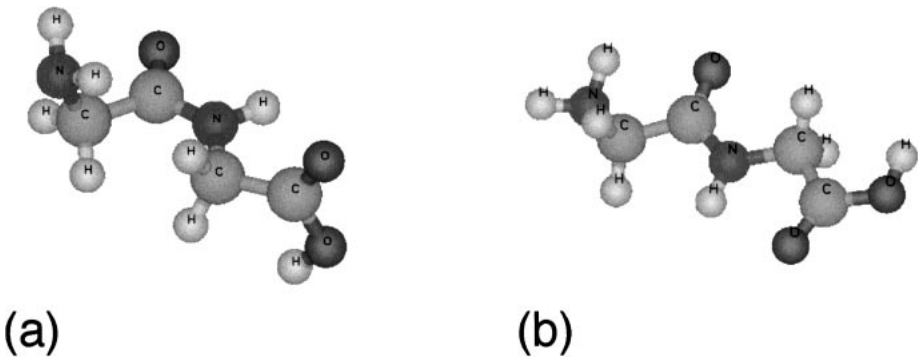

Fig. 8. *cis*- (a) and *trans*- (b) forms of glycylglycine.

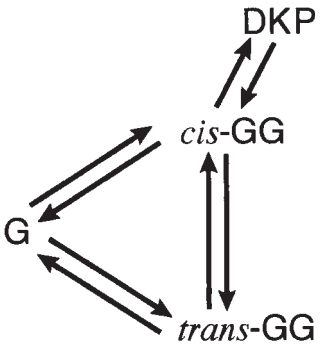

Fig. 9. Hypothetical reaction network, including G, *cis*-GG, *trans*-GG, and DKP.

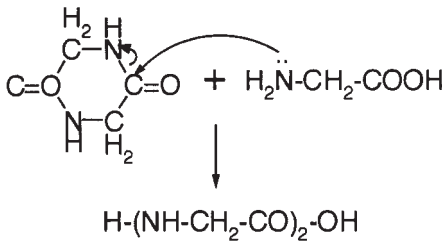

Fig. 10. Probable mechanism of the reaction DKP + G → GGG proposed by Nagayama et al.[27]

G and the reverse hydrolysis, which explains the rapid synthesis of DKP in Fig. 4a and of G in Fig. 4c. Considering that we could construct the simulation based on the simpler reaction model (Fig. 3) to the experimental data, *cis*-GG can be regarded as a transition state for 2G ⇄ DKP and *trans*-GG ⇄ DKP. In this study, the products were analyzed by HPLC at room temperature. This analytical method cannot detect the concentration of *cis*-GG that might exist at 200 °C, because it quickly isomerizes into the *trans* form, hydrolyzes into G, or undergoes intramolecular peptide bond formation into DKP as the sample solution is cooled prior to the HPLC analyses. Direct detection of *cis*-GG at hydrothermal temperatures should be conducted using other analytical techniques, e.g., NMR, in the future. This will judge the validity of our hypothetical reaction mechanism.

Nagayama et al.[27] discussed the mechanism of DKP-mediated peptide formation, i.e., reaction a_4, at 90 °C. They investigated the reaction of alanine with DKP, and the resulting tripeptide was NH_2–Gly–Gly–Ala–COOH, not NH_2–Ala–Gly–Gly–COOH. Thus, they concluded that the reaction proceeds through the nucleic attack of an amino group of monomeric amino acids on DKP accompanied by its ring opening (Fig. 10). We expect that the same mechanism proceeds at 200 °C, which will be confirmed by a future experiment employing Ala and DKP.

In this investigation, we confirmed that G polymerization mediated by DKP occurs at 200 °C as well as at 90 °C.[27] Since DKP has been considered to be a chemically inactive and useless compound, its formation has been believed to be an unavoidable obstacle for the elongation of peptide chains beyond dipeptides. However, we suggest that DKP is not a very stable molecule, but serves as a probable intermediate for prebiotic peptide formation under hydrothermal conditions. Moreover, we suggest that the formation of DKP proceeds not only following *trans*-GG and *cis*-GG formation, but also by directly forming *cis*-GG at 200 °C. Thus, under hydrothermal conditions, G and other types of amino acids may polymerize efficiently through the formation of DKPs consisting of various amino acids and the condensations of such DKPs and various amino acids/oligopeptides. It is noteworthy that Nagayama et al.[27] reported that the larger the number n of $(Gly)_n$ in the condensation reaction with DKP, the higher the yield of the resulting peptide, $(Gly)_{n+2}$. Thus, DKPs might enhance abiotic polymerization more rapidly as the elongation proceeds in hydrothermal environments. This possibility should be investigated in more detail in the future.

We thank Tomohiko Yamaguchi, Toshinori Kusumi, and Hajime Mita for their technical suggestions. We are grateful to Kazuhiro Iida and Shigeru Deguchi for their useful comments and to Masatake Akita for drawing the diglycines. We deeply appreciate Hirotaka Sugawara, director general of the National Laboratory for High Energy Physics, for giving us the opportunity to perform experiments there. This research was financially supported by the Group Research (Soken/G-3) Fund of the Graduate University for Advanced Studies.

References

1 J. B. Corliss, J. A. Baross, and S. E. Hoffman, *Ocean. Acta*, **4** (suppl.), 59 (1981).
2 J. A. Baross and S. E. Hoffman, *Origins of Life*, **15**, 327 (1985).
3 E. G. Nisbet, *Nature*, **322**, 206 (1986).
4 J. B. Corliss, *Origins of Life*, **19**, 534 (1989).
5 J. B. Corliss, *Nature*, **347**, 624 (1990).
6 H. Yanagawa and K. Kobayashi, *Origins Life Evol. Biosphere*, **22**, 147 (1992).
7 E. Imai, H. Honda, K. Hatori, A. Brack, and K. Matsuno, *Science*, **283**, 831 (1999).
8 Y. Qian, M. H. Engel, S. A. Macko, S. Carpenter, and J. W. Deming, *Geochim. Cosmochim. Acta*, **57**, 3281 (1993).
9 E. V. Shock, *Geochim. Cosmochim. Acta*, **56**, 3481 (1992).
10 N. G. Holm, *Origins Life Evol. Biosphere*, **22**, 5 (1992).
11 D. K. Alargov, S. Deguchi, K. Tsujii, and K. Horikoshi, *Origins Life Evol. Biosphere*, **32**, 1 (2002).
12 Md. N. Islam, T. Kaneko, and K. Kobayashi, *Bull. Chem. Soc. Jpn.*, **76**, 1171 (2003).
13 R. H. White, *Nature*, **310**, 430 (1984).
14 G. D. Steinman, R. M. Lemmon, and M. Calvin, *Proc. Natl. Acad. Sci. U.S.A.*, **52**, 27 (1964).
15 G. D. Steinman, D. H. Kenyon, and M. Calvin, *Biochim. Biophys. Acta*, **124**, 339 (1966).
16 J. Rabinowitz, J. Flores, R. Krebsbach, and G. Rogers, *Nature*, **224**, 795 (1969).
17 N. Chung, R. Lohrmann, L. E. Orgel, and J. Rabinowitz, *J. Mol. Evol.*, **8**, 307 (1976).
18 J. Rabinowitz and A. Hampai, *Helv. Chim. Acta*, **61**, 1842 (1978).
19 J. R. Hawker, Jr. and J. Oro, *J. Mol. Evol.*, **17**, 285 (1981).
20 M. Sakurai and H. Yanagawa, *Origins of Life*, **14**, 171 (1984).
21 J. Yamanaka, K. Inomata, and Y. Yamagata, *Origins of Life*, **18**, 165 (1988).
22 E. Imai, H. Honda, K. Hatori, and K. Matsuno, *Origins Life*

Evol. Biosphere, **29**, 249 (1999).
23 Y. Ogata, E. Imai, H. Honda, K. Hatori, and K. Matsuno, *Origins Life Evol. Biosphere*, **30**, 527 (2000).
24 A. Brack, K. W. Ehler, and L. E. Orgel, *J. Mol. Evol.*, **8**, 307 (1976).
25 A. L. Weber and L. E. Orgel, *J. Mol. Evol.*, **11**, 189 (1978).
26 A. L. Weber and L. E. Orgel, *J. Mol. Evol.*, **13**, 185 (1979).
27 M. Nagayama, O. Takaoka, K. Inomata, and Y. Yamagata, *Origins Life Evol. Biosphere*, **20**, 249 (1990).
28 S. Mitsuzawa and T. Yukawa, *Origins Life Evol. Biosphere*, **33**, 163 (2003).
29 S. M. Steinberg and J. L. Bada, *J. Org. Chem.*, **48**, 2295 (1983).
30 E. V. Shock, *Origins Life Evol. Biosphere*, **20**, 33 (1990).
31 C. Branden and J. Tooze, "Introduction to Protein Structure, Second Edition," Garland Publishing, New York (1999).
32 C. Grathwohl and K. Wuthrich, *Biopolymers*, **20**, 2623 (1981).