6-Membered ring intermediates in polymerization of *N*-carboxyanhydride-L-α-arginine in H₂O

XIN Liang[1], REN Jie[1], XIANG JunFeng[2†], YAN Qin[1], XIE Yi[1], WANG KongJiang[1†],
LAI PingAn[3] & BAI YaDuo[3]

[1] Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;
[2] Beijing National Laboratory for Molecular Sciences (BNLMS), Centre for Molecular Sciences, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, China;
[3] Beijing Entry-Exit Inspection and Quarantine, Beijing 100026, China

In polymerization of *N*-carboxyanhydride-L-α-arginine (L-Arg-NCA) in H₂O, nucleophilic reaction of guanidine group with the carbonyl group of L-Arg-NCA leads to quick intramolecular rearrangement, yielding a 6-membered ring intermediate 1-amidino-3-amino-2-piperidone, which is either elongated by another L-Arg-NCA yielding arginyl-1-amidino-3-amino-2-piperidone or hydrolyzed to L-α-arginine. The oligoarginines are formed mainly through hydrolysis of arginyl-1-amidino-3-amino-2-piperidones. This is a unique pathway in polymerization of L-Arg-NCA with regard to the usual pathway of elongations by reaction of *N*-carboxyanhydride-L-α-amino acid with L-α-amino acid or oligopeptides.

1 Introduction

Polymerization of L-α-amino acids via *N*-carboxyanhydrides (Leuchs anhydride) serves as a model for both biotic[1−4] and prebiotic[5−11] peptide formation. In spite of its simplicity, the polymerization of *N*-carboxyanhydride-L-α-amino acid presents a few aspects which have not been clarified[12]. L-α-arginine is the most basic natural amino acid (pK_a of the guanidinium is ≈ 13). Polymerization of L-α-arginine is one of the concerns because of the "exceptions" found in reactions (Scheme 1). The natural α-amino acids found in protein yield polymers of the amino acids upon polymerization of L-α-amino acid by phosgene in dioxane except L-α-arginine[13−15]. In Ag⁺-assisted cyclization of unprotected peptides, only peptides of *C*-terminal arginine (GDFVR-thioester) give rise to hydrolysis as the major product in addition to the formation of arginine side-chain-to-end cyclic peptides. The unusual high yield of hydrolysis was also attributed to be an "exception"[16]. The other exception we found

is that in H₂O, polymerization of L-Arg-NCA yields oligoarginines as the dominant products, while the products from polymerization of L-α-lysine and L-α-histidine include complicated polymers linked by both amide (—NH—CO—) and ureido (—N—CO—N—) linkage, indicating the involvement of ω-NH₂ of L-α-lysine and —NH— of L-α-histidine in the polymerization[17,18].

During monitoring the reaction process of polymerization of L-Arg-NCA in H₂O using High Performance Liquid Chromatography (HPLC), we incidentally found a series of new intermediates (Figure 1). This suggests that there is a new pathway of reaction in addition to the known *N*-carboxyanhydride intermediates. Here we present new 6-membered ring intermediates found in polymerization of L-Arg-NCA in H₂O.
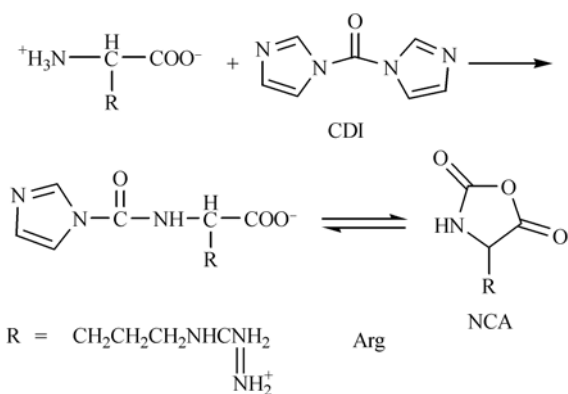
Scheme 1 The activation of L-α-arginine by *N,N'*-carbonyldiimidazole.

2 Experimental

2.1 Materials

L-α-arginine (≥99%), *N*-α-acetyl-L-α-arginine, *N,N'*-carbonyldiimidazole (CDI, ≥97%) and other major compounds were from Sigma.

2.2 Polymerization

The reaction started by adding the solution (pH 8.0, 2℃) of L-α-arginine prepared by titrating the free base of L-α-arginine using 1 M HCl to more than two-fold solid CDI. After quick vortex and storage at 2℃ for 5 min, the tubes were kept in a 20℃ water bath for 24 h[7]. The activation of *N*-α-acetyl-L-α-arginine by CDI was achieved by transferring the solution of 50 mM *N*-α-acetyl- L-α-arginine (pH 8.0 at 2℃) into more than two-fold solid CDI. The products of the polymerization were analyzed by the established method[17].

Preparation of 2,5-diketopiperazine. 2,5-diketopiperazine of L-α-arginine was prepared from the polymerization products of 20 mM L-α-arginine by HPLC on Zorbax 300-SCX (4.6×50 mm, 5 μ) using $NH_2CH_2CH_2$- $NH_2 \cdot 2HCl$ gradient (0%－100% B in 60 min; buffer A: 0.02 M NaH_2PO_4 at pH 5.2; buffer B is 1 M NH_2CH_2- $CH_2NH_2 \cdot 2HCl$ in buffer A at pH 3.2). The fractions were then desalted using RP HPLC (Altima C18, 4.6×60 mm, 5 μ; buffer: water). After being frozen at about –40℃, the sample was lyophilized.

2.3 Preparation of intermediate C1, C2 and C3

Because of their good resolution on HPLC profiles (Figure 1), **C1**, **C2** and **C3** were prepared using cation-exchange HPLC[17]. The preparation of the major intermediates is very hard because they are labile to in-

organic ions, and also to organic additives such as methanol. Fortunately, we found that $NaClO_4$ gradient can be employed to prepare **C1** without significant hydrolysis of the intermediates.

The **C1** was prepared from the reaction solution for 3 min (in a water bath at 20℃ for 3 min) by HPLC on Zorbax 300-SCX (4.6×50 mm, 5 μ) using $NaClO_4$ gradient (2%－80% B in 80 min; buffer A is 0.02 M NaH_2PO_4 at pH 5.2; buffer B is 2 M $NaClO_4$ in buffer A at pH 5.2). The samples were frozen in liquid nitrogen immediately. After repeating tedious collections, the collected fractions were defrosted and then desalted using RP HPLC (Altima C18, 4.6×60 mm, 5 μ; buffer is water). The fractions of the compound were frozen immediately. After lyophilization at about –50℃, the solid power was kept frozen until it was used for other experiments.

The preparation of **C2**, **C3** was basically similar to that for the preparation of **C1**. However, the relative low yield of **C2**, **C3** relative to **C1** and their liability prevented us from getting enough samples for detailed NMR characterization.

2.4 Techniques section

HPLC was performed on a Hitachi L7100 pump with UV-vis L7420 detector using gradient elution at the flow rate of 1 mL/min. The HPLC elutions were detected at 214 nm.

Electrospray ionization mass spectrometry (ESI-MS) spectra were obtained on a Shimadzu LC-MS 2010 mass spectrometer. The samples were dissolved in 50% methanol aqueous solution. Matrix-assisted laser desorption ionization linear mass spectrometry (MALDI-TOF-MS) was carried out on an AXIMA-CFR plus (Kratos Analytical, Shimadzu Group Company) equipped with a pulsed nitrogen laser emitting at 337 nm. The detector was operated in positive mode and the pulse voltage was set to 20000 V. All spectra were recorded in a reflection mode. α-Cyano-4-hydroxycinnamic acid was dissolved in 50% acetonitrile and 0.1% trifluoroacetic acid to get saturated solution. Then the saturated solution was diluted once and mixed with the peptide sample at 1∶1. The mixture (0.5 μL) was dropped onto a stainless steel plate and the solvent was evaporated.

All nuclear magnetic resonance (NMR) measurements were carried out at 600.13 MHz using a Bruker AVANCE 600 spectrometer equipped with a 5 mm BBI probe capa-

ble of delivering z-field gradients up to 50 G·Cm^{-1}. The temperature was controlled by a Bruker BVT 3200 temperature control unit. The 1D proton spectra were recorded by the standard Bruker pulse program Zgig 30 with 8 k scans. The 1D carbon spectra were recorded by the standard Bruker Pulse Program Zgig30 with 8 k scans. The DEPT135 and DEPT90 spectra were recorded by the standard Bruker pulse program dept135 and dept 90 with 3 k scans, respectively. The relaxation delay for each above measurement is 2 s.

The Heteronuclear Single Quantum Correlation (HSQC) and Heteronuclear Multiple Bond Coherence (HMBC) spectra were obtained using hsqcetgp and hmbcgplpndqf pulse programs, respectively. For each FID, 32 transients were collected with 2 s relaxation delays. C-H coupling constant was 145 Hz, while the long range C-H coupling constant for HMBC was 6.25 Hz. The acquisition data size was 2048 × 256, 1024 × 1024 points were used in Fourier transformation.

3 Results and discussion

HPLC and MS detection of the intermediates in polymerization. Just incubating the solution of L-Arg-NCA in a H$_2$O bath at 20℃ for 5 min gave rise to a major intermediate **C1** (Figure 1(a)), and other intermediates (**C2**, **C3** and **C4**) after 10 min of incubation (Figure 1(b) and (c)). Each intermediate, the major product within about 30 min of the polymerization at 20℃, was followed by an oligoarginine (Figure 1(b)). After 2.5 h reaction, the products are dominated by oligoarginines (Figure 1(e)), and finally, virtually all products are oligoarginines (Figure 1(f)) after proceeding of the reaction for 8 h at 20℃. This is unique in the known polymerization reactions of NCA of L-α-amino acids. The activation of α-NH$_2$ group by CDI, other than that of guanidine group on the side-chain of L-α-arginine, was confirmed by the reaction of CDI with N-α-acetyl-L-α-arginine, which yielded no detectable CDI-activated products determined by both HPLC and MS analysis.

Corresponding to the intermediates and oligoarginines detected by HPLC (Figure 1(b)), both ESI-MS (M$^+$ mode) and MALDI-TOF MS (plus) analysis of the sample after incubating for 5 min at 20℃ of activated 50 mM L-α-arginine disclosed the presence of a compound with molecular weight (MW) being 156, and 18 less than that of L-α-arginine. Both ESI and MALDI-TOF MS analysis of the sample after incubating for 30 min at
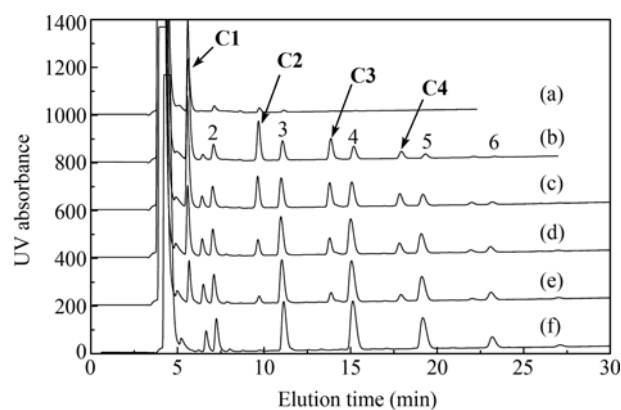

Figure 1 HPLC profiles of the products after polymerization of 20 mM L-α-arginine at 20℃ for 5 min (a), 30 min (b), 60 min (c), 105 min (d), 150 min (e) and 8 h (f). The numbers denote the residues of the formed oligoarginines, and **C1**, **C2**, **C3** and **C4** refer to the four intermediates.

20℃ detected a series of compounds with MW being 18 less than that of oligoarginines, in correspondence to Figure 1(b) and (c). After completion of the reaction, only oligoarginines were detected, except a stable compound with MW being 18 less than dipeptide of L-α-arginine (Arg-Arg). Therefore, both HPLC and MS data consistently confirmed the formation of a series of intermediates (such as **C1**−**C4** in Figure 1) in addition to oligoarginines. The unusual high yields of intermediates compared to the oligoarginines in the first 30 min of reaction indicate their major roles in the polymerization compared to the usual intermediate Leuch anhydride. And the increasing yields of oligoarginines corresponding to the decreasing yields of the intermediates during the polymerization suggest that oligoarginines were formed from the corresponding intermediates through hydrolysis.

3.1 Characterization of 2,5-diketopiperazine

It is noted that in addition to oligoarginines, which increases steadily with the proceeding of the reaction, the yield of a minor stable product between **C1** and Arg-Arg lso increases with time (Figure 1). The purified compound was regarded as 2,5-diketopiperazine of L-α-arginine. This is interesting with regard to the fact that in polymerization of L-amino acid-NCA, 2,5-diketopiperazine can hardly be formed because carboxyl-activated dipeptides are not generated (Scheme 1)[7,19−21]. In fact, we did not detect the formation of 2,5-diketopiperazine of amino acids in the polymerization of L-Glu-NCA and Gly-NCA. This also indicates that there is a new pathway of reaction in the polymerization of L-Arg-NCA, different from that of the usual polymerization where

oligopeptides are elongated by reacting with activated monomers (Leuchs anhydrides).

3.2 Characterization of C1

We found that $NaClO_4$ gradient can be employed to prepare **C1** using cation exchange HPLC without significant hydrolysis of the intermediates if the fractions can be kept frozen as soon as possible. We were rewarded by enough quantity of the intermediates **C1** after the tedious repeated HPLC collection.

Independent MS analysis of the prepared intermediates disclosed that these three compounds (**C1**—**C3**) were corresponding to the molecules with MW 18 less than that of L-α-arginine, Arg-Arg and Arg-Arg-Arg respectively. And after standing the intermediates at room temperature for 10 days, separate MS analysis detected three compounds with MW the same as L-α-arginine, Arg-Arg and Arg-Arg-Arg respectively. Figure 2a and b shows the ^{1}H NMR spectrum of freshly prepared **C1** and that after hydrolysis. The labile property of the intermediate **C1** was demonstrated by the time-dependent ^{1}H NMR spectra shown in Supporting Information. After 10 days, intermediate **C1** is completely hy-

drolyzed to L-α-arginine (Figure 2).

Since the molecular weight of the intermediate **C1** is 18 less than that of L-Arg, both six- and eight-membered heterocyclic structures might be formed from the intramolecular condensation reaction of the reactant. In the case of six-membered heterocycle, the proton in δ-CH_2 should have long-range correlation with both carbons of guanazole and carbonyl, while the eight-membered heterocycle should only give long-range correlation with the carbon of guanazole. In order to obtain a definite answer, the proton, carbon, ^{13}C-DEPT135, ^{13}C-DEPT90, ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC), ^{1}H-^{13}C heteronuclear multiple bond coherence (HMBC) spectra of intermediate **C1** have been obtained. The chemical shifts and assignments of both **C1** and the hydrolytic product are summarized in Tables 1 and 2. It is observed that the proton in δ-CH_2 group in HMBC spectra gives rise to long-range correlation with both C in guanazole and C in C≡O. That means that a 6-membered ring intermediate **C1** (1-amidino-3-amino-2-piperidone) was formed through intramolecular rearrangement of L-Arg-NCA (Scheme 2).
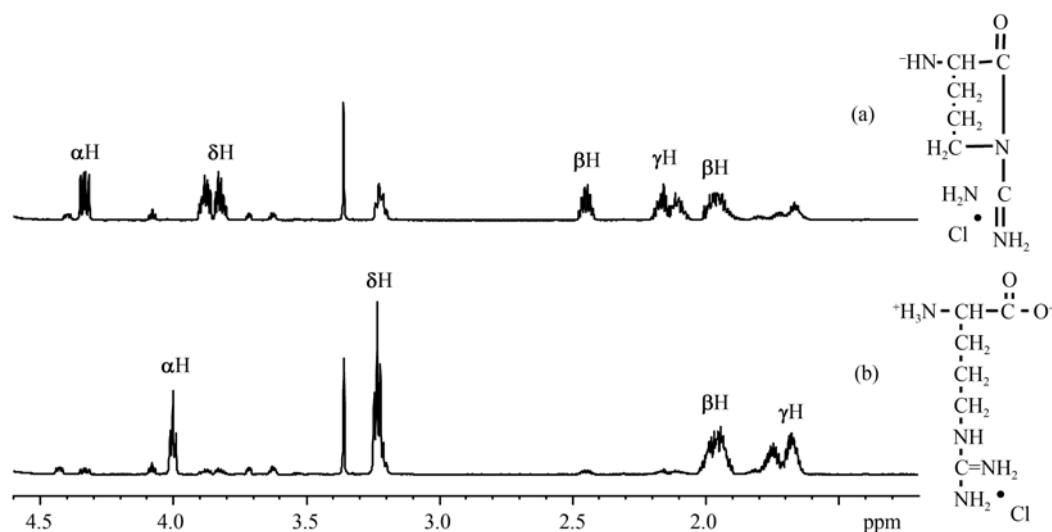

Figure 2 The ^{1}H NMR spectra of fresh sample **C1** (a) and that stored for 10 days (b) in D_2O at 298 K.

Table 1 The summary of the chemical shifts and assignments of the intermediate **C1**

Group	δ_H (ppm)	δ_C (ppm)	^{1}H-^{13}C long-distance correlation
CH (α)	4.34(4)	51.4	19.1, 23.1, 171.4
CH_2 (β)	2.45 (m)/1.98 (m)	23.1	19.1, 46.7, 51.4, 171.4
CH_2 (γ)	2.13 (m)	19.1	23.1, 46.7, 51.4
CH_2 (δ)	3.86 (m)	46.7	19.1, 23.1, 156.7, 171.4
C (guanidine)	—	156.7	—
CO	—	171.4	—

Table 2 The summary of the chemical shifts and assignments of the hydrolysis product

Group	δ_H (ppm)	δ_C (ppm)	^{1}H-^{13}C long-distance correlation
CH (α)	4.05(3)	52.1	22.8, 26.1, 171.4
CH$_2$ (β)	1.96(m)	26.1	22.8, 39.5, 52.1, 171.4
CH$_2$ (γ)	1.71(m)	22.8	26.1, 39.5, 52.1
CH$_2$ (δ)	3.23(3)	39.5	22.8, 26.1, 155.9
C (guanidine)	—	155.9	—
CO	—	171.4	—
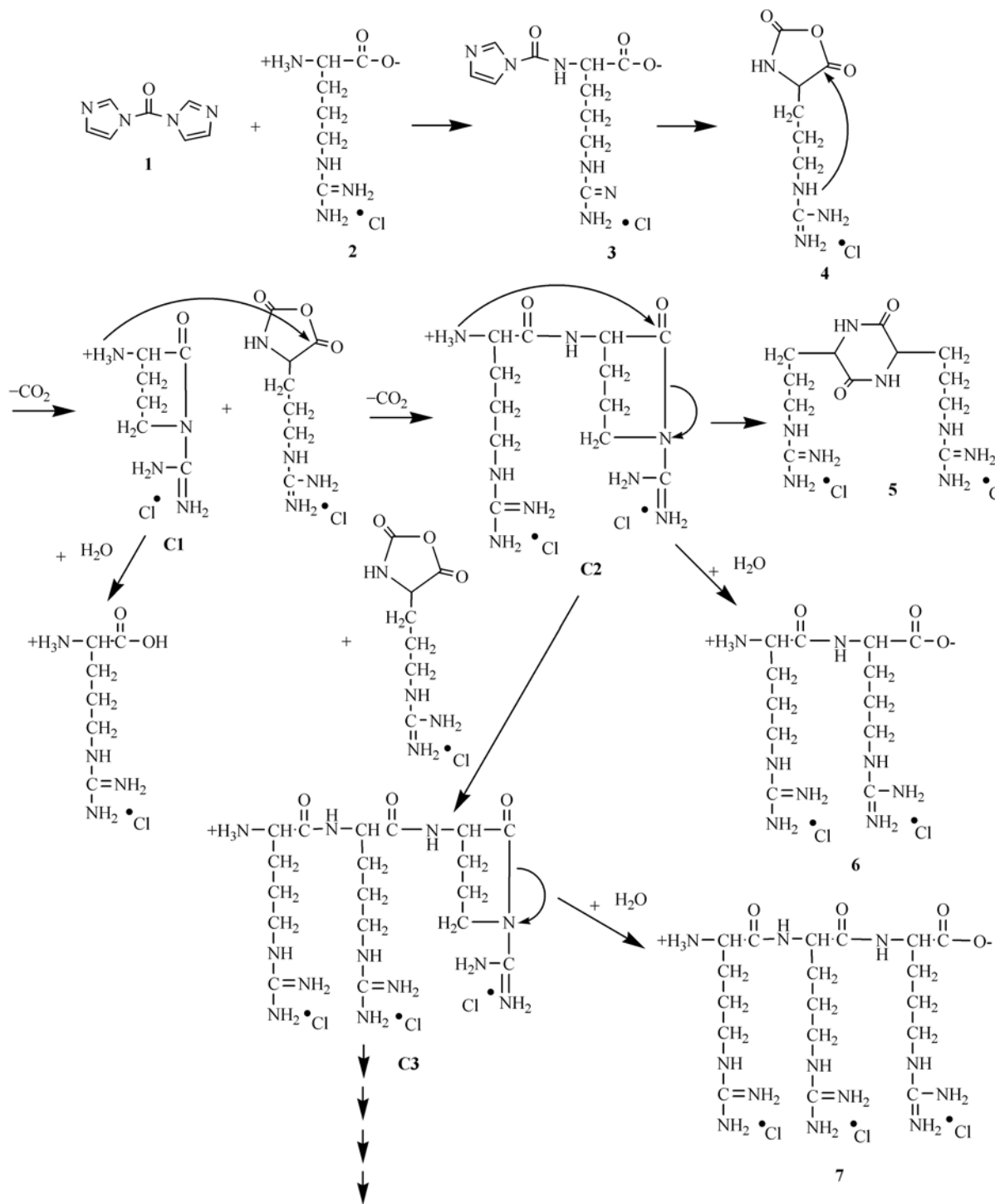
Scheme 2 Polymerization of L-α-arginine induced by CDI in H$_2$O.

3.3 Characterization of C2 and C3

Because of the technical difficulties in preparation of intermediate **C2** and **C3** in the procedures, we did not get enough **C2** and **C3** for detailed NMR characterization. However, both independent MS and MALDI-TOF determination of the reaction solution clearly uncovered that like the MW of **C1**, MW of **C2** and **C3** is 18 less than that of Arg-Arg and Arg-Arg-Arg respectively. After being kept at room temperature for 24 h, Arg-Arg and Arg-Arg-Arg were formed. Cation exchange HPLC profiles (Figure 1) clearly indicate that the positive charges of **C1**, **C2** and **C3** are in the order of one positive charge increase. Therefore, we speculated that the **C2** and **C3** possess the similar structures to the structure of **C1**.

3.4 Mechanism of polymerization of L-Arg-NCA

In summary, the polymerization of L-Arg-NCA is significantly different from that of the other L-amino acid-NCA (Scheme 2). The reaction of CDI **1** with L-α-arginine **2** results in the activation of α-NH$_2$ yielding **3** and the subsequent cyclized anhydride L-Arg-NCA **4**. The usual pathway of polymerization is that the hydrolyzed L-amino acid or the oligopeptides can be elongated by reacting with L-amino acid-NCA. However because of the existence of guandinium, the intramolecular nucleophilic reaction of NH-group of guanidine group with carbonyl group of **4** leads to quick intramolecular rearrangement, yielding a new 6-membered ring intermediate 1-amidino-3-amino-2-piperidone **C1**, which is elongated by either another L-Arg-NCA yielding arginyl-1-amidino-3-amino-2-piperidone **C2** or hydrolyzed to L-α-arginine. Product **C2** could be elongated by another L-Arg-NCA yielding **C3**, or hydrolyzed to Arg-Arg **6**. As a minor pathway, α-NH$_2$ of **C1** could attack the α-carbonyl group itself resulting in cyclization to **5**, a stable 6-membered diketopiperazine of L-α-arginine. Further elongation of product **C3** by another L-Arg-NCA gives rise to longer arginyl-1-amidino-3-amino-2-piperidones, or **C3** hydrolyzes to Arg-Arg-Arg **7**.

There are two pathways in the reaction: either via an intramolecular rearrangement of L-Arg-NCA to **C1** or via intermolecular reaction of L-Arg-NCA with the amino acid or the formed peptides yielding oligoarginines. Because intramolecular reactions entropically favor intermolecular reactions, it is suggested that the intramolecular rearrangement of L-Arg-NCA to **C1** significantly benefits based on the yields of oligoarginyl-L-amidino-amino-2-piperidones significantly higher than the oligoarginines in the first hour of polymerization (Figure 1(a)−1(c)). The usual pathway of elongation by reaction of NCA with free L-amino acids or the formed peptides is a minor one. Therefore, the formation of oligoarginines appears to be mainly from the hydrolysis of oligoarginyl-1-amidino-3-amino-2-piperidones.

The formation of **C1**, **C2** and **C3** is equivalent to blockage of guanidinium group, therefore, preventing the occurrence of the side reactions, occurring in polymerization of L-Lys-NCA and L-His-NCA[17]. The mechanism proposed might also explain the high yield of hydrolysis of GDfVR-thioester[16]. It is proposed that nucleophilic reaction of the activated carbonyl thioester with neighboring guanidinium of arginine should have happened to give **C1**, which will mainly hydrolyze to GDFVR. However for peptides with non-neighboring arginine (DFVRG-thioester), the steric distance will hinder the formation of **C1**. The major product will be end-to-end cyclic peptides.

4 Conclusions

In the polymerization of L-amino acid-NCA in H$_2$O, the elongation is achieved by reaction of L-amino acid-NCA with L-amino acid or the peptides[19−21]. However, in the polymerization of L-Arg-NCA, the intramolecular rearrangement of L-Arg-NCA to a 6-membered 1-amidino-3-amino-2-piperidone **C1** benefits more than the intermolecular reactions. Here the major pathway of elongation is the reaction of L-Arg-NCA with the intermediates and the oligoarginines are formed through hydrolysis of oligoarginyl-L-amidino-amino-2-piperidones. The formation of the 6-membered ring intermediates is equivalent to protection of the guanidinium group of L-α-arginine yielding oligoarginines as the dominant products, avoiding the complicated products in the polymerization of L-Lys-NCA and L-His-NCA.

1 Stahmann M. Polyamino Acids, Polypeptides and Proteins. Madison, Wisconsin: Univ Wisconsin Press, 1962

2 Katchelski E, Sela M, Silman H I, Berger A. Polyamino Acids as

Protein Models. The Proteins. Vol Ⅱ. 2nd ed. New York: Academic Press, 1964. 405−581

3 Kricheldorf H R. α-Aminoacid-N-Carboxy-Anhydrides and Related

Heterocycles: Synthesis, Properties, Peptide Synthesis. Polymerization. Berlin: Springer-Verlag, 1987. 1－209

4　Kricheldorf H R. Polypeptides and 100 years of chemistry of α-amino acid N-carboxyanhydrides. Angew Chem Int Ed, 2006, 45: 5752－5784

5　Ehler K W. Reactions of carbamylimidazoles with nucleophiles－An example of an intramolecular acyl transfer reaction. J Org Chem, 1976, 41: 3041－3042

6　Brack A. Selective emergence and survival of early polypeptides in water. Orig Life Evol Biosph, 1987, 17: 367－379

7　Hill R., Orgel L E. Oligomerization of negative-charged amino acids by carbonyldiimidazole. Origins Life Evol Biosph, 1996, 26: 539－545

8　Blocher M, Liu D, Walde P, Luisi P L. Liposome-assisted selective polycondensation of α-amino acids and peptides. Macromolecules, 1999, 32: 7332－7334

9　Pascal R, Boiteau L, Commeyras A. From the prebiotic synthesis of α-amino acids towards a primitive translation apparatus for the synthesis of peptides. Top Curr Chem, 2005, 259: 69－122

10　Plasson R, Biron J P, Cottet H, Commeyras A, Taillades J. Kinetic study of the polymerization of α-aminoacid N-carboxyanhydrides in aqueous solution using capillary electrophoresis. J Chromatograph A, 2002, 952: 239－248

11　Hitz T, Luisi P L. Spontaneous onset of homochirality in oligopeptide chains generated in the polymerization of N-carboxyanhydride amino acids in water. Origins Life Evol Biosph, 2004, 34: 93－110

12　Kricheldorf H R. Polypeptides. In: Penczek S, ed. Models of Bio-

polymers by Ring-Opening Polymerization. Boca Raton, Florida: CRC Press, 1990. 1－132

13　Zervas L, Winitz M, Greenstein J P. Studies on arginine peptides. Ⅰ. Intermediates in the synthesis of N-terminal and C-terminal arginine peptides. J Org Chem. 1957, 22: 1515－1521

14　Hayakawa T, Kondo Y, Yamamoto H, Murakami Y. The synthesis of poly-L-arginine hydrobromide and copolymers of L-arginine and other amino acids. Bull Chem Soc Jpn, 1969, 42: 479－482

15　Hayakawa T, Windsor C R, Fox S W. Copolymerization of the Leuchs anhydrides of the eighteen amino acids common to protein. Arch Biochem Biophys, 1967, 118: 265－272

16　Zhang L S, Tam J P. Lactone and lactam library synthesis by silver ion-assisted orthogonal cyclization of unpro-tected peptides. J Am Chem Soc, 1999, 121: 3311－3320

17　Xin L, Ma Y L, Liu Y N, Yan Q, Wang K J. Sodium chloride enhanced oligomerization of L-arginine. Biopolymers, 2006, 81: 1－7

18　Ren J, Xin L, Liu Y N, Wang K J. Copolymerization of mixed L-alpha-arginine with L-alpha-glutamic acid. Macromolecules, 2008, 41: 1996－2002

19　Bartlett P D, Jones R H. A kinetic study of the Leuchs anhydrides in aqueous solution. J Am Chem Soc, 1957, 79: 2153－2159

20　Ehler K W, Orgel L E. N,N'-carbonyldiimidazole-induced peptide formation in aqueous solution. Biochim Biophys Acta, 1976, 434: 233－243

21　Liu R H, Orgel L E. Polymerization of beta-amino acids in aqueous solution. Origins Life Evol Biosph, 1998, 28: 47－60