Copolymerization of Mixed L-α-Arginine with L-α-Glutamic Acid

Jie Ren, Liang Xin, Yi-Nan Liu, and Kong-Jiang Wang*

Protein & Peptide Pharmaceutical Lab, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Received August 27, 2007; Revised Manuscript Received December 16, 2007

ABSTRACT: The finding of the enhanced polymerization of *N,N*-carbonyldiimidazole (CDI) activated L-α-glutamic acid (Glu) by alkali cations like Na^+ and K^+ and CDI-activated L-α-arginine (Arg) by halide anions like Cl^- and other common ions indicates the importance of the weak electrostatic coordination of these common ions on the polymerization of charged amino acids in H_2O. Here the copolymerization of L-arginine-*N*-carboxyanhydride (L-Arg-NCA) mixed with L-glutamic acid-*N*-carboxyanhydride (L-Glu-NCA) produced much longer peptides than that of the same concentration of L-Arg-NCA or L-Glu-NCA, and the optimal proportion of L-Glu-NCA over L-Arg-NCA for the longest peptide formation was 1:1. NaCl affected the sequences of the peptides formed in the mixed L-Arg-NCA with L-Glu-NCA, yielding peptides with significantly more Arg residues than Glu residues. Enzymatic hydrolysis of the formed peptides uncovered significantly more two residues of the same charges. These data strengthen the importance of weak electrostatic interactions in the polymerization of charged amino acids in H_2O.

Introduction

Since the finding of amino acid production in the classic discharge experiment by Miller in 1953,[1] the prebiotic synthesis of the longer heteropeptides[2] in H_2O was recognized as one of the major problems in the origins of life.[3−6] Although various pathways of peptide formation have been reported, the formed peptides were generally less than 5-mers and rarely exceeded 10-mers.[7,8]

It was reported that just heating a dry unactivated mixture of the 18 free amino acids at 180 °C for 2−5 h yielded as high as 10000 Da polymers,[9,10] although further studies found the actual molecular weights are ca. 2000 Da.[11] However the proteinoids also formed unnatural peptide linkages and therefore were contested.[12−15] The other pathways of thermal conversion of unactivated amino acids into polypeptides like repetitive evaporation of the solutions of unactivated amino acids[11,16] or direct hydrothermal synthesis also gave rise to short homopeptides.[17,18]

Catalysis enhances prebiotic peptides in H_2O to different extents. For example, Cu^{2+},[3] and CdS or ZnS catalyze peptide formation, yielding homopeptides up to heptapeptide.[19] To date, the longest homopeptides (∼55 residues) with natural amide linkage formed were achieved by repeated 50 times "feeding" the polymerization with the activated monomers under the catalysis of illites at the cost of low rate utility of the activated monomers.[7]

The role of electrostatic interactions in prebiotic peptide formation has been noted. The yield of ligation of Arg pentapeptide (Arg_5) and Arg dipeptide (Arg_2) to cysteine residue of a scaffold peptide (Cys-S-S-Cys-Gly-Glu_{10}) differs by a factor of 18, presumably owing to higher electrostatic affinity of Arg_5 to the Glu_{10} scaffold than Arg_2.[20] The enhanced polymerization of both the negative-charged amino acids by positive-charged micelles[21] and the positive-charged dipeptides by negative-charged liposome[22,23] were assumed to be due to the electrostatic and/or hydrophobic concentration of the substrates. Na^+ (K^+) and Cl^- are known for their noncatalytic functions in biological systems[24,25] because their weak coordination on most biological ligands yields negligible stability constants and almost diffusion-controlled ligand exchange rates of H_2O (∼10^9−10^{10} s^{-1}).[26] The universal presence of these simple ions in the primitive oceans reminded us to consider their possible role in the polymerization of amino acids during our exploration of the prebiotic peptide formation. We found that the polymerizations of L-Arg-NCA and L-Glu-NCA were enhanced by Cl^- and Na^+ (K^+) respectively, yielding much longer peptides and less short peptides.[27,28] The notion that the enhancement of Cl^- is at least comparable to FeS_2, the best solid catalyst for the polymerization of L-Arg-NCA,[29] suggests that the weak electrostatic coordination of ions on the oppositely charged peptides significantly affected the processes of the reaction. However they have yet led to the expected heteropeptides.

Although the most plausible system of prebiotic peptide evolution should be the mixed systems of amino acids, both the theoretical difficulties concerning the copolymerization of a mixture of different monomers,[30] and technical difficulties in characterization of the formed peptides have severely limited the progress of this approach. Here we extend our research to investigate the copolymerization of Arg mixed with other amino acids such as glycine (Gly), especially of the oppositely charged Glu because of the relatively strong electrostatic interactions between carboxylate and guanidinium,[31] and also to explore the effect of NaCl on the copolymerization of the mixed L-Arg-NCA with L-Glu-NCA in view of their enhanced separate polymerizations of charged amino acids. We are not aware of any existing experimental data.

Experimental Section

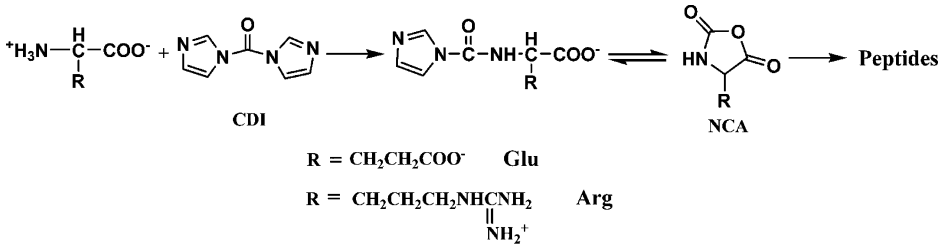
Materials. L-Amino acids (≥99%, TLC), Gly (>99% titration), Glu-Glu (EE, >99%), *N,N′*-carbonyldiimidazole (CDI, ≥97%) and other major compounds were from Sigma. Trypsin (from bovine pancreas, 9000 units/mg) was from Fluka. Endoproteinase Glu-C (proteinase V8) was from both Sigma (500−1000 units/mg) and New England Biolabs (38.3 (μmol/min)/mg).

Polymerizations. L-Arg-NCA, L-Glu-NCA, L-leucine-*N*-carboxyanhydride (L-Leu-NCA) and glycine-*N*-carboxyanhydride (Gly-NCA) were gained by activating Arg, Glu, L-α-leucine (Leu), and Gly using CDI (Scheme 1).[32] For example, this may be done by

* Corresponding author. E-mail: wangkj@moon.ibp.ac.cn.

10.1021/ma7019286 CCC: $40.75
Published on Web 02/27/2008

Scheme 1. Polymerization of Amino Acids Induced by CDI



adding 1 mL of solution of 20 mM Arg at 2 °C (pH 8.0) into twice the excess of solid CDI (6.8 mg). After vortexing, the solution was kept at 2 °C for 5 min and then stored at room temperature for 24 h to allow the polymerization to complete. In the polymerization of L-Arg-NCA with sodium acetate, Glu, L-Glutamic acid γ-methyl ester as additives, pH of the solution containing the additives was adjusted to pH 8.0 before reacting with L-Arg-NCA. The samples with NaCl as additive were first dialyzed overnight using Spectra/Pro MWCO 500 before mass spectrometry identification. The polymerization products were determined by matrix-assisted laser desorption ionization linear mass spectrometry (MALDI−TOF−MS).

Elongation Reactions. In order to determine the competitive selectivity for the activated monomers, L-Arg-NCA or L-Glu-NCA was set to react with mixed 1:1 Arg with Glu or their homopeptides. For example, Arg solution (0.5 mL, 20 mM, pH 8.0) was first activated by 3.4 mg CDI at 2 °C for 5 min. Then the solution was added into 0.5 mL mixed 100 mM Glu with 100 mM Arg. After incubating the samples at room temperature for 24 h, the products were analyzed by reversed-phase high performance liquid chromatography (RP−HPLC).

Enzymatic Hydrolysis. The peptides formed in the copolymerization of mixed L-Arg-NCA with L-Glu-NCA were first desalted by Sephadex G-25 chromatography. The lyophilized peptides formed in copolymerization of 1 mL 10 mM mixed L-Arg-NCA with L-Glu-NCA were dissolved in 1 mL of 100 mM NH_4HCO_3, 20 mM $CaCl_2$, pH 7.8, and then incubated with trypsin (0.6 mg) for 3 h at 37 °C. Digestion was stopped by heating the samples at 100 °C for 5 min. In proteinase V8 reactions, Glu tetrapeptides (0.27 mM) dissolved in 50 μL 100 mM NH_4HCO_3 buffer at pH 7.8 were digested by V8 at a ratio of 60:1 (M/M) for 18 h at 25 °C. The hydrolytic products were analyzed using RP−HPLC.

Synthesis and Characterization. Homopeptides Arg-Arg (RR), Arg-Arg-Arg (RRR), and Glu-Glu-Glu (EEE) were prepared and identified by the established methods.[28,33]

Glu-Arg (ER), Glu-Glu-Arg (EER), and Glu-Glu-Glu-Arg (EEER) were prepared from the reaction of 10 mM L-Glu-NCA with 50 mM Arg. They were separated by RP−HPLC. After lyophilization, the isolated peptides were confirmed by electrospray ionization mass spectrometry (ESI−MS): ER, calculated 304.3, detected 304.64; EER, calulated 431.2, detected 431.2; EEER calculated 560.24, detected 560.3.

Arg-Glu (RE). A 10 mM sample of L-Arg-NCA was added into 50 mM Glu and 50 mM Arg and then was kept at room temperature for about 24 h. The sample was analyzed by RP−HPLC. Out of the two major products, RR was confirmed using the cation-exchange HPLC.[28] RE was identified using ESI−MS (calculated 304.3, detected 304.4).

Sequence Assignments. The assignment of the unknown sequence of peptides from the copolymerization of the mixed amino acids is based on their molecular weights from MALDI−TOF−MS. The assignment of RE, ER, EER, and EEER is based on co-injection of the identified compounds with the samples.

We found that the retention time (RT) of RE on the anion-exchange HPLC is shorter than EE, and that of ER on the cation-exchange HPLC is shorter than that of RR.

REE. The products from the reaction of 5 mM L-Arg-NCA + 25 mM RR + 25 mM EE were first analyzed using RP−HPLC. Two major products were collected. RRR was identified using cation-exchange HPLC, the other product with RT shorter than EEE on anion-exchange HPLC is assigned to be REE.

ERR. The products from the reaction of 5 mM L-Glu-NCA + 25 mM RR + 25 mM EE were first analyzed using RP−HPLC. Two major products were collected. EEE was identified using anion-exchange HPLC, the other product with RT shorter than RRR on cation-exchange HPLC is assigned to be ERR.

The assignment of REEE and ERRR is based on the similar processes to that of REE and ERR.

Quantification of the Peptides. The quantification of the short peptides is based on UV absorption of the amide bond at 214 nm using HPLC. Because the molar extinction coefficients of short peptides are not available, molar extinction coefficients of the heteropeptides of the same length as ER and RE were assumed to be the same. The concentration of EE was determined by weighing authentic samples from Sigma, and the concentration of RR was determined by first hydrolyzing RR in 6 M HCl at 110 °C for 40 h and the amino acid determination using classical ninhydrin method. The molar extinction coefficients of EEE and RRR were assumed to be twice that of EE and RR based on the number of the amide bond in the molecules. Subsequently the concentration of the short peptides can be quantified by comparing their UV absorption at 214 nm using either cation/anion-exchange or RP−HPLC.

Techniques Section. HPLC was performed on Hitachi L7100 pump with UV−vis L7420 detector using gradient elution. The RP−HPLC column for elongation products analysis is an Alltima C18 column (Alltech, 5 μm, 4.6 × 250 mm), and the elution gradient is from 0% B to 100% B in 60 min at the flow rate of 1 mL/min (solvent A: 0.1% trifluoroacetic acid, solvent B: 0.1% trifluoroacetic acid in 50% acetonitrile). The RP−HPLC analysis of enzymatic hydrolysates was performed on Diamonsil C18 column (Dikma Technologies, 5 μm, 4.6 × 250 mm), and the gradient is from 2% B to 60% B in 60 min. The analysis conditions of anion-exchange HPLC and cation-exchange HPLC were the same as reported.[28,33]

MALDI−TOF−MS was carried out on 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 reflectron mode. α-Cyano-4-hydroxycinnamic acid was dissolved in 50% acetonitrile 0.1% trifluoroacetic acid to get saturated solution. Then the saturated solution was diluted 1 time 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. ESI−MS spectra were obtained on Shimadzu LC-MS 2010 mass spectrometer. The samples were dissolved in 50% methanol aqueous solution.

Results and Discussion

Copolymerization of L-Arg-NCA Mixed with L-Glu-NCA. Our previous results demonstrated that the polymerization of charged amino acids[27,28] including phospho-*O*-L-serine were enhanced by oppositely charged common ions like Na^+ (K^+) and Cl^-. The weak coordination of the oppositely charged ions was assumed to be responsible for the enhancement. We are especially interested in the possible role of the interactions between carboxylate and guanidinium on the polymerization

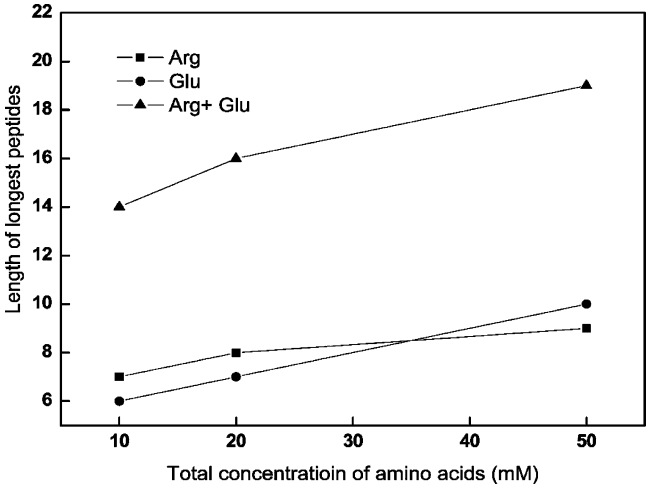


Figure 1. Longest peptide formation of homopolymerization of L-Arg-NCA (squares) or L-Glu-NCA (circles) and that of the copolymerization of mixed 1:1 L-Arg-NCA with L-Glu-NCA (triangles). The length of the longest peptide is based on MALDI−TOF MS chromatographs.

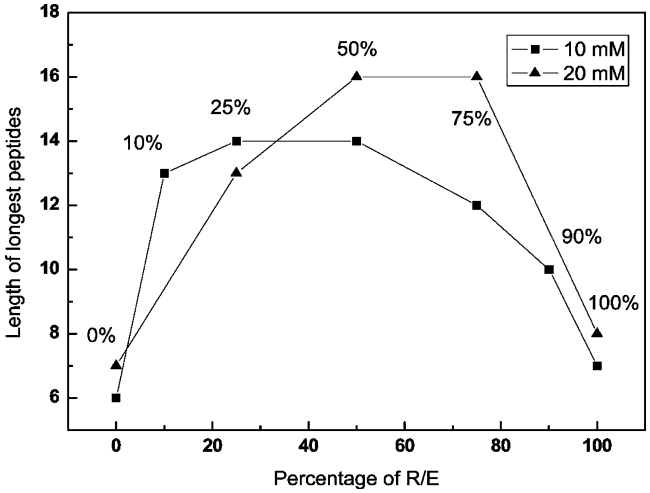


Figure 2. Longest peptide formation of polymerization of different proportion of total concentration of 10 mM (squares) and 20 mM (triangles) mixed L-Arg-NCA with L-Glu-NCA.

because the dissociation energy for a dimer molecular system with double guanidinium−carboxylate linkages (ionic H-bonds) is up to 51 to 64 kJ mol^{-1} in H_2O,[31] significantly more than that between point charges. Also, both carboxylate and guanidinium, the groups of the natural amino acids, constitute the major groups of electrostatic interactions in proteins.

Indeed, the copolymerization of mixed L-Arg-NCA with L-Glu-NCA was significantly enhanced (Figure 1). The homopolymerization of 10 mM L-Arg-NCA or L-Glu-NCA yielded successive homopeptides[2] with 7-mer or 6-mer as the longest peptides.[27−29,32] The longest peptide formed in reactions of mixed 5 mM L-Arg-NCA with 5 mM L-Glu-NCA is up to 14-mer (Figure 1). In contrast, homopolymerizations of 50 mM L-Arg-NCA and 50 mM L-Glu-NCA yielded 9-mer and 10-mer as the longest peptides respectively. The copolymerization of mixed 50 mM of 1:1 L-Arg-NCA with L-Glu-NCA gave rise to 19-mer as the longest peptides, significantly more than that formed in the homopolymerization of the same concentration of L-Arg-NCA or L-Glu-NCA (Figure 1).

Copolymerization of L-Arg-NCA in the Presence of Different Proportions of L-Glu-NCA. The longest peptide formed in reactions of total 10 mM mixed L-Arg-NCA at 100%, 90%, 75%, 50%, 25%, 10%, and 0% proportion with L-Glu-NCA was 7-mer, 10-mer, 12-mer, 14-mer, 14-mer, 13-mer, and 6-mer (Figure 2) respectively. The copolymerization of 20 mM mixed solution of L-Arg-NCA at 75%, 50%, and 25% proportion with L-Glu-NCA gave rise to 16-mer, 16-mer, and 13-mer as the longest peptide respectively. The optimal proportion for the longest peptide formation is the reaction of 1:1 L-Arg-NCA with L-Glu-NCA.

Amino Acid Composition in the Formed Peptides. There are generally two or three major kinds of peptides of certain length in the polymerization products (Figure 3). The heteropeptides are the dominant products even in the reaction of 9 mM L-Arg-NCA with 1 mM L-Glu-NCA (Figure 3A). Peptides longer than 7-residues are heteropeptides. And the longer peptides were dominantly composed of similar proportion of Glu and Arg in the reaction of 5 mM L-Arg-NCA with 5 mM L-Glu-NCA (Figure 3C). For example, out of the peptides of 11-residues formed in 5 mM L-Arg-NCA with 5 mM L-Glu-NCA, three peptides (yields in the order of R6E5 > R5E6 > R7E4) are the major products. Similarly, two peptides (R6E6 > R7E5) are the products of peptides of 12 residues in the same reaction (Figure 3C). The number of the amino acid residues in the formed peptides is directly proportional to the comonomer concentration.

Effect of NaCl on the Polymerization. Because electrostatic interactions are weakened by increasing ionic strength, the presence of a high concentration of NaCl should decrease the enhancement in the mixed L-Glu-NCA with L-Arg-NCA. However the polymerization of L-Glu-NCA and L-Arg-NCA was enhanced by Na^{+}[27] and Cl^{-},[28] respectively. Therefore, it will be interesting to probe the effect of NaCl on the copolymerization of L-Arg-NCA with L-Glu-NCA. The polymerization degree of mixed 5 mM L-Arg-NCA with 5 mM L-Glu-NCA was not significantly affected by the presence of 0.1, 0.5, 0.8, 1, 2, 3, and 4 M NaCl. This may be understandable because of the mentioned significantly stronger electrostatic interactions between carboxylate and guanidinium than that between point charges of Na^{+} and carboxylate or Cl^{-} and guanidinium of the amino acids.[31]

However with increasing NaCl concentration, amino acid residues of the major peptides gradually deviated from 1:1 of L-Arg over L-Glu, and finally in the presence of 4 M NaCl (Figure 4), major peptides are those significantly deviating from 1:1 of L-Arg over L-Glu. For example, the major 6-mer without NaCl is R3E3 (Figure 4A), and that in the presence of 1 M NaCl is R3E3 and R4E2 (Figure 4B); however, the major 6-mer in the presence of 4 M NaCl is R4E2 and R5E1 (Figure 4C). This shows the complications of electrostatic interactions in the reactions. We have no solid explanation for this.

Copolymerization of L-Arg-NCA with Other Amino Acids. The polymerization of 2.5 mM Gly-NCA in H_2O is highly efficient, yielding 16-mer as the longest peptides,[28] compared with that of charged amino acids. And the polymerization of 3 mM and higher concentration of Gly-NCA yielded precipitates of oligoglycines. This might suggest that the electrostatic repulsion between the side chains with the same charge is the reason for significantly less efficient polymerization of L-Arg-NCA and L-Glu-NCA. The addition of 5 mM Gly-NCA into 5 mM L-Arg-NCA yielded the 17-mer as the longest peptides. Considering the polymerization of the Gly-NCA itself, it seems that there is no enhancement by the mixed L-Arg-NCA with Gly-NCA. In the formed peptides longer than 10-mer, such peptides as R4G7, R4G8, R4G9, R5G7, R5G8, R5G8, and R5G9 are the major products. Similarly, the polymerization of 5 mM L-Arg-NCA with 5 mM L-Leu-NCA gave rise to 12-mer

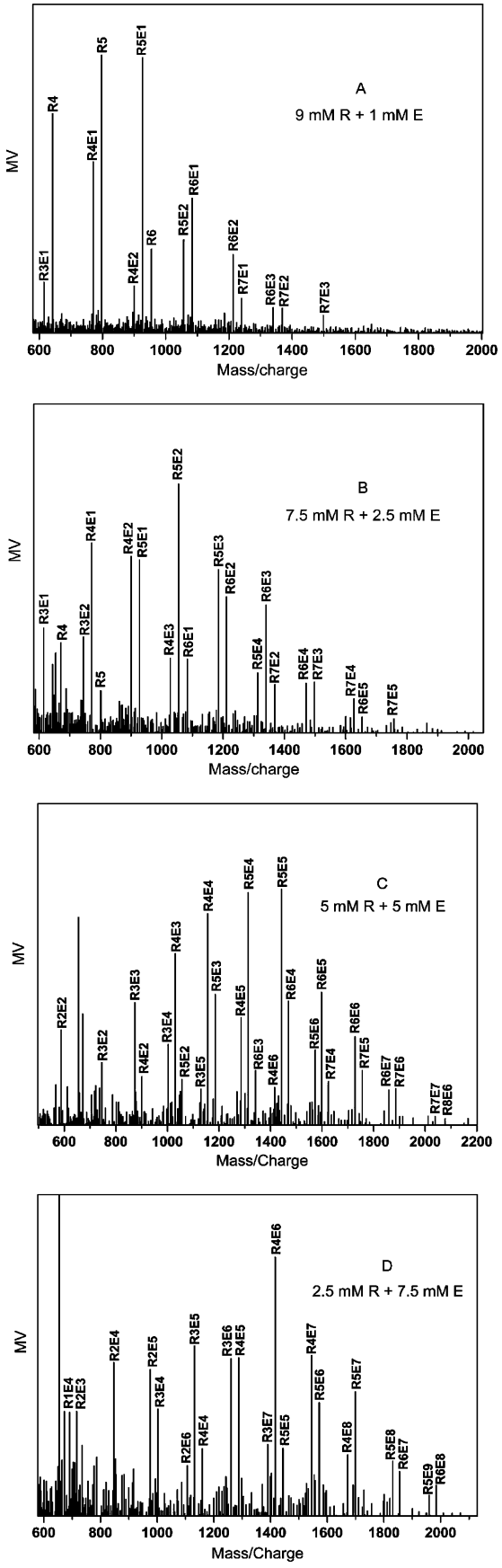


Figure 3. MALDI−TOF−MS chromatographs of different proportion of total 10 mM mixed L-Arg-NCA with L-Glu-NCA. Abbreviations for the amino acid residues: R = Arg; E = Glu.

as the longest peptide. In the short peptides Leu residues are dominant; major products are peptides like R1L4, R1L5, R2L5, R2L6, and R2L7. However in the longer peptides, the proportion of arginine residues increases, R4L6, R4L7, and R5L7 are the major products. The addition of Gly-NCA into the mixed system of L-Arg-NCA with L-Glu-NCA yielded very complicated peptides (Figure 5). Gly residues in the formed major peptides accounted for significantly less percentage than that of Arg or Glu. Therefore, copolymerization of mixed amino acids does not necessarily give rise to much longer peptides, and the enhancement peptide formation in the mixed Arg with Glu should owing to their electrostatic interactions between their side chains.

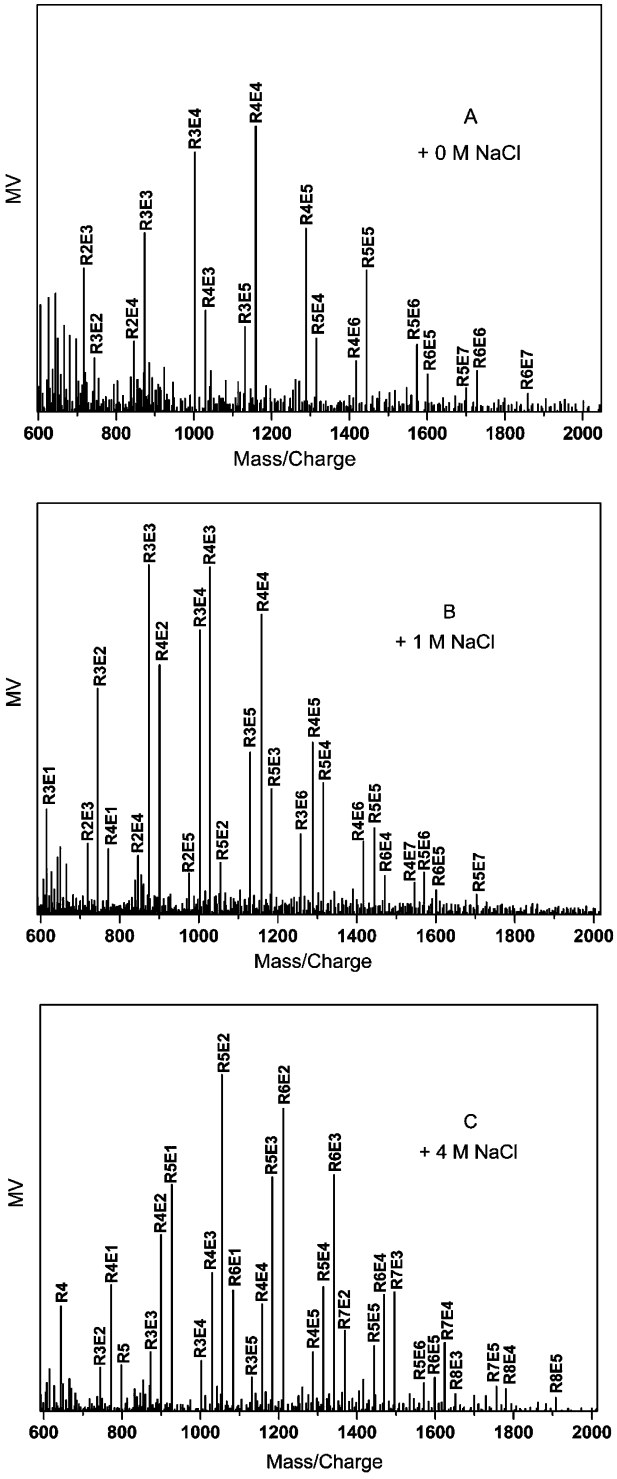


Figure 4. MALDI−TOF−MS chromatographs of copolymerization of 5 mM L-Arg-NCA + 5 mM L-Glu-NCA in the presence of 0, 1, and 4 M NaCl.

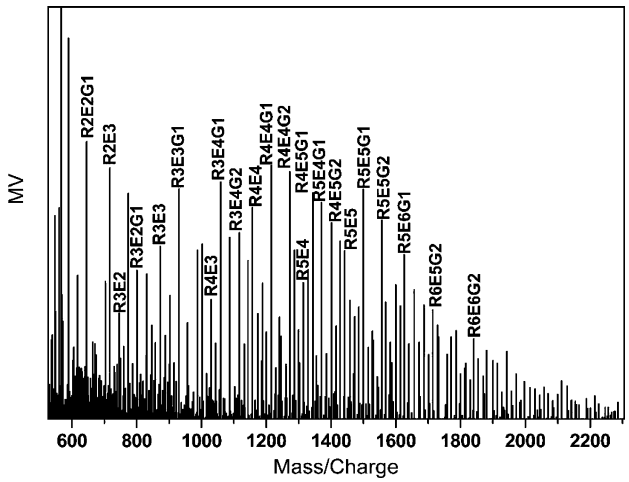


Figure 5. MALDI−TOF−MS chromatographs of 10 mM L-Arg-NCA + 10 mM L-Glu-NCA + 4 mM Gly-NCA. Abbreviations for the amino acid residues: R = Arg; E = Glu; G = Gly.

Table 1. Yields of the Elongated Peptides by the Activated Monomers (X*)[a]

reactions	rates of peptides
10 mM E* + 50 mM R + 50 mM E	ER/EE = 2.98
5 mM E* + 25 mM RR + 25 mM EE	ERR/EEE = 4.81
5 mM E* + 25 mM RRR + 25 mM EEE	ERRR/EEEE = 2.01
10 mM R* + 50 mM R + 50 mM E	RE/RR = 0.44
5 mM R* + 25 mM RR + 25 mM EE	REE/RRR = 1.07
5 mM R* + 25 mM RRR + 25 mM EEE	REEE/RRRR = 1.38

[a] Abbreviations for the amino acid residues: R = Arg; E = Glu. The rates of products were based on the peak areas determined by HPLC.

The polymerization of 10 mM L-Arg-NCA in the presence of 10 mM sodium acetate gives rise to R7 as the longest peptides. The polymerization of L-Arg-NCA in the presence of Glu should give rise to polymers with Glu or Arg on the C-terminal because the polymerization proceeds as a stepwise elongation through L-Arg-NCA attacking of the α-NH_2 of the amino acids or that of the peptides. Indeed, we found that the polymer formed in 10 mM L-Arg-NCA with 10 mM Glu consisted of two kinds of peptides: major peptides are those with one Glu residue R2E1, R3E1, R4E1, R5E1, R6E1, and minor peptides are R3, R4, and R5. Similarly, the presence of 10 mM L-Glutamic acid γ-methyl ester in 10 mM L-Arg-NCA yielded R2E1, R3E1, R4E1, and R5E1 as major products and R3, R4, and R5 as minor peptides. Obviously, the presence of either Glu or carboxylate-blocked Glu on the C-terminal of the peptide is not enough to enhance the polymerization.

Sequential Selectivity of the Homopeptides. In order to explore the sequential selectivity, Arg and Glu or their purified homopeptides were set to be elongated by 10 mM L-Arg-NCA or 10 mM L-Glu-NCA to investigate the competitive elongations, for example:

$$\text{L-Glu-NCA (10 mM)} + \text{Glu (50 mM)} + \text{Arg (50 mM)} \rightarrow \text{Glu-Glu} + \text{Glu-Arg}$$

On the basis of the yields of the two major products determined using HPLC, significantly sequential selectivity was observed in the elongations (Table 1). Arg, Arg-Arg, and Arg-Arg-Arg are always preferentially elongated by L-Glu-NCA respectively, showing preference of oppositely charged substrates. However, for the elongations by L-Arg-NCA, Arg was elongated more efficiently than Glu; Glu-Glu and Arg-Arg were elongated similarly; Glu-Glu-Glu was elongated slightly more efficient than Arg-Arg-Arg. This might be related to significantly higher rates of polymerization of L-Arg-NCA because of the formation of new six-membered intermediates through intramolecular rearrangement of L-Arg-NCA.[44]

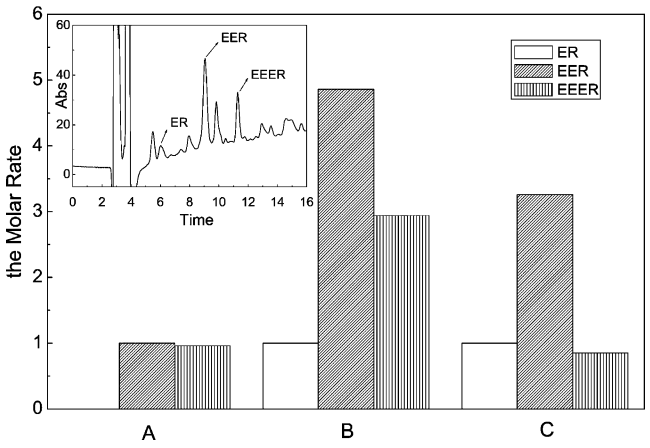


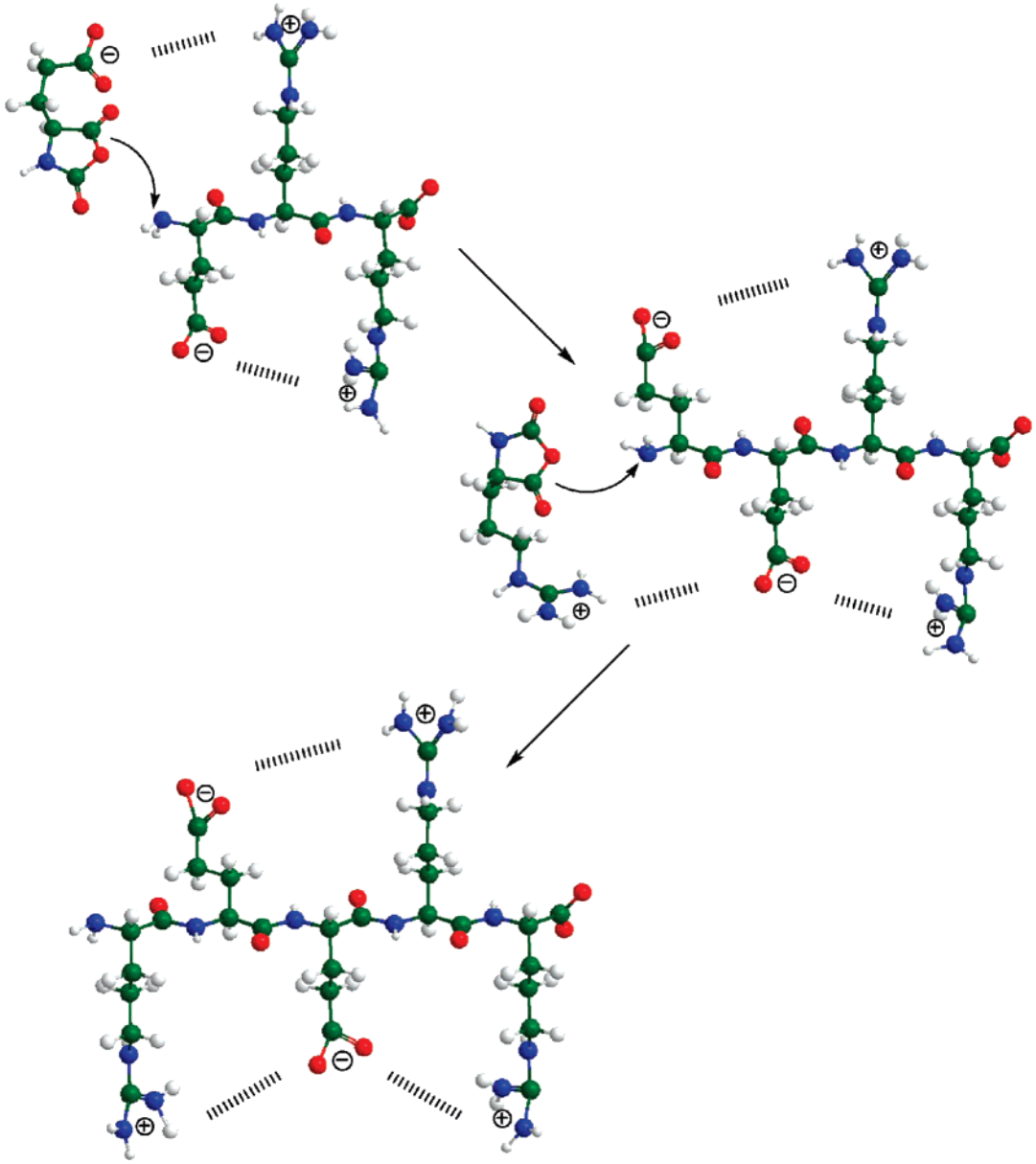
Figure 6. Molar rates of the products after trypsin digestion of the peptides formed in 1 mL 10 mM mixed L-Arg-NCA with L-Glu-NCA determined by HPLC. (A) L-Arg-NCA/L-Glu-NCA is 1:9, where EER/EEER = 1: 0.96, ER is not detected. (B) L-Arg-NCA/L-Glu-NCA is 1:3, where ER/EER/EEER is 1:4.86:2.94. (C) L-Arg-NCA/L-Glu-NCA is 1:1, where ER/EER/EEER is 1:3.26:0.85. The inset HPLC profile shows the products after trypsin digestion of the products from the polymerization of 1 mL 10 mM mixed L-Arg-NCA with L-Glu-NCA (1:1). We assume that the molar extinction coefficiencies of ER, EER and EEER at 214 nm are in direct proportion to the number of the amide bond of the peptides.

Enzymatic Analysis of the Sequences of the Formed Peptides. Although previous data suggest the sequential selectivity, it is technically difficult to characterize the sequence distribution of all the formed peptides to establish the sequential selectivity because the polymerization, e.g., mixed 5 mM L-Arg-NCA with 5 mM L-Glu-NCA, yields hundreds of peptides of different sequences. Therefore, sequentially specific enzymes (trypsin and V8-protease) were applied to detect the sequence selectivity.

Trypsin cleaves peptide bonds at lysine and arginine C-terminal, so the major hydrolytic peptides after trypsin hydrolysis are ER, EER, and EEER. EER was always the major product compared with ER and EEER, even in polymerization of 1:9 ratio of L-Arg-NCA to L-Glu-NCA (Figure 6). This means that Glu residues mainly exist in the form of two Glu in the heteropeptides. The attempt to hydrolyze the peptides using V8-proteinase, which cleaves peptide bonds C-terminal to Glu residues, failed. However because the major products detected by MS are roughly 1:1 of Arg residue to Glu residues. It is appropriate to assume the existence of two Arg residues as major sequence in the formed peptides.

Mechanisms of the Enhancement. The polymerization of NCAs of neutral amino acids is much more efficient than that of NCAs of charged amino acids, suggesting the negative role of electrostatic repulsion between the N-terminal residues and the coming charged activated monomer. In the enhanced polymerization of charged amino acids by oppositely charged ions weakly bound ions around the N-terminal residues could partially neutralize the charges of the residues to weaken the electrostatic repulsion, therefore enhancing the polymerizations.[27,28] The much more efficient copolymerization of mixed L-Arg-NCA with L-Glu-NCA, combined with the finding of peptides with two same-charged residues as the major sequences in the formed peptides consistently suggests the importance of electrostatic interactions in the polymerization (Scheme 2).

Scheme 2. Mechanism for the Enhanced Copolymerization of L-Arg-NCA with L-Glu-NCA



As demonstrated in Scheme 2, we speculate the mechanisms of the enhancement based on electrostatic interactions. The elongation of the peptides depends on the competitive approaching α-NH_2 of amino acids or of the peptides between L-Arg-NCA and L-Glu-NCA. Because the dimension and shape of L-Arg-NCA and L-Glu-NCA is similar, random elongation must exist, as demonstrated by the presence of ER and EEER as minor sequences in the peptides. However the electrostatic attractions between the charged group on the side chain of peptides and the charged side chain of the coming activated monomer obviously could attract the activated monomer to N-terminal of the peptides to be elongated. Further, electrostatic attraction between the oppositely charged side chain of the activated monomer and the second residue on N-terminal of the peptide is more sterically favorable for anhydrides of the activated monomer to attack the N atom on N-terminal of the peptides. Therefore, both approximation and orientation, two basic factors contributing to the catalysis of the reaction,[34] could have contributed to the formation of significantly more two residues of the same charge, and the enhanced copolymerization.

Concluding Remarks

The copolymerization of L-Arg-NCA mixed with L-Glu-NCA was significantly enhanced, giving rise to much longer peptide formation compared to that of homopolymerization of the same concentration of L-Arg-NCA or L-Glu-NCA. The optimal proportion is 1:1 L-Arg-NCA over L-Glu-NCA. Higher concentration of NaCl affects the sequence of the formed peptides, yielding peptides with significantly more Arg residues. Both enzymatic hydrolysis and competitive elongations indicate the sequential selectivity. In the formed peptides, two residues of the same charge are the major sequences. These data suggest the importance of electrostatic interactions in the polymerization in H_2O.

The number of all possible sequences of a peptide of 50 amino acid residues is huge ($=20^{50}$). It seems that electrostatic interactions could have been one of the basic factors guiding the prebiotic polymerization to yield limited number of sequential longer peptides, required for the origins of life in H_2O.[3−7]

The electrostatic interaction is greatly diminished in H_2O, leading to free energies of association near zero for an ion-pair

or ion-dipole formation although the interaction of two monovalent ions in vacuum or in crystal is up to −100 kcal mol^{-1}.[35] This is why surface-exposed ion pairs of proteins often contribute very little to protein stability in H_2O because the energy gain from forming an ion pair is just enough or not enough to compensate for the loss of desolvation penalty of the two ions.[36,37] The dissociation energies for a dimer molecular system with double guanidinium−carboxylate linkages (ionic H-bonds) are up to 51 to 64 kJ mol^{-1} in H_2O,[31] significantly more than that between point charges. This explains the null effect of NaCl on the copolymerization of the mixed L-Arg-NCA with L-Glu-NCA.

The importance of electrostatic interactions in living systems has long been a topic of controversial, e.g., that in protein chemistry. Most interpretations on the function of electrostatic interactions in H_2O are ambiguous because of the complications like conformational changes and the dehydration enthalpies due to ion-pair formation in molecules like protein in H_2O. In contrast, the polymerization of NCAs of charged amino acids in H_2O is a simple system consisted of the charged peptides of different length, common ions like Na^+ and Cl^- and the activated charged monomers. Furthermore, this reaction not only provides a simple system for the complicated electrostatic interactions in H_2O but also gives their effects on the kinetic reactions of the peptides. It is the simplicity that facilitates our unambiguous observation of the enhancement by simple ions like Na^+ and Cl^-, and the enhanced copolymerization reported here. We believe that this system might act as a good model for further exploration of weak electrostatic interactions in H_2O.

The mechanism of weak electrostatic interactions is far from being fully understood. Many ion-specific phenomena were considered largely a mystery until now.[38,39] The Hofmeister effects found in 1888, which initially refer to the relative abilities of anions and cations to precipitate proteins and were then found to be applicable in many basic aspects of both chemistry and biology from solubility of salts, to electrolyte activities, surface tensions to pH measurements, microemulsion, proteins and membranes, transport across membranes, enzyme activities, bacterial growth, and so on[43] are still unexplained by present theories of physical chemistry.[35,38−43]

Acknowledgment. We acknowledge the financial support from Natural Basic Research Program of China (Grant 973-2007CB935901).

References and Notes

(1) Miller, S. L. *Science* **1953**, *117*, 528−529.
(2) The heteropeptide refers to peptides composed of different amino acids, while the homopeptide refers to those composed of the same amino acid residues.
(3) Rode, B. M. *Peptides* **1999**, *20*, 773−786.
(4) Pascal, R.; Boiteau, L.; Commeyras, A. *Top. Curr. Chem.* **2005**, *259*, 69−122.
(5) Brack, A. *Chem. Biodivers.* **2007**, *4*, 665−679.
(6) Luisi, P. L. *Chem. Biodivers.* **2007**, *4*, 603−621.
(7) Ferris, J. P.; Hill, A. R.; Liu, R. H.; Orgel, L. E. *Nature (London)* **1996**, *381*, 59−61.
(8) Orgel, L. E. *Orig. Life Evol. Biosph.* **1998**, *28*, 227−234.
(9) Fox, S. W.; Harada, K.; Vegotsky, A. *Experientia* **1959**, *15*, 81−84.
(10) Rohlfing, D. L. *Science* **1976**, *193*, 68−70.
(11) Yanagawa, H.; Kojima, K.; Ito, M.; Handa, N. *J. Mol. Evol.* **1990**, *31*, 180−186.
(12) Saunders, M. A.; Rohlfing, D. L. *Biosystems* **1974**, *6*, 82−92.
(13) Andini, S.; Benedetti, E.; Ferrara, L.; Paolillo, L.; Temussi, P. A. *Origins Life* **1975**, *6*, 147−153.
(14) Temussi, P. A.; Paolillo, L.; Ferrara, L.; Benedetti, E.; Andini, S. *J. Mol. Evol.* **1976**, *7*, 105−110.
(15) Ferris, J. P. *Nature (London)* **1989**, *337*, 609−610.
(16) Ito, M.; Handa, N.; Yanagawa, H. *J. Mol. Evol.* **1990**, *31*, 187−194.
(17) Yanagawa, H.; Kobayashi, K. *Orig. Life Evol. Biosph.* **1992**, *22*, 147−159.
(18) Yanagawa, H.; Ogawa, Y.; Kojima, K.; Ito, M. *Orig. Life Evol. Biosph.* **1988**, *18*, 179−207.
(19) Bertrand, M.; Bure, C.; Fleury, F.; Brack, A. Prebiotic polymerization of amino acid thioesters on mineral surfaces. In *Geochemistry and the Origin of Life*; Nakashima, S., Maruyama, S., Brack, A., Windley, B. F., Eds.; Universal Academy Press: Tokyo, 2001; pp 51−60.
(20) Chu, B. C.; Orgel, L. E. *Orig. Life Evol. Biosph.* **1999**, *29*, 441−449.
(21) Bohler, C.; Hill, A. R., Jr.; Orgel, L. E. *Orig. Life Evol. Biosph.* **1996**, *26*, 1−5.
(22) Blocher, M.; Liu, D.; Walde, P.; Luisi, P. L. *Macromolecules* **1999**, *32*, 7332−7334.
(23) Blocher, M.; Liu, D.; Luisi, P. L. *Macromolecules* **2000**, *33*, 5787−5796.
(24) Kaim, W.; Schwederski, B. *Bioinorgainc Chemistry: Inorganic Chemistry of Life*, 2nd ed.; John Wiley & Sons: New York, 1994; p 267.
(25) Frausto de Siliva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements: the Inorganic Chemistry of Life*, 2nd ed.; Oxford Press: London, 2001; p 231.
(26) Diebler, H.; Eigen, M.; Ilgenfritz, G.; Maa, G.; Winkler, R. *Pure Appl. Chem.* **1969**, *20*, 93−115.
(27) Wang, K. J.; Yao, N.; Li, C. *Orig. Life Evol. Biosph.* **2005**, *35*, 313−322.
(28) Xin, L.; Ma, Y. L.; Liu, Y. N.; Yan, Q.; Wang, K. J. *Biopolymers* **2006**, *81*, 1−7.
(29) Liu, R. H.; Orgel, L. E. *Orig. Life Evol. Biosph.* **1998**, *28*, 245−257.
(30) Luisi, P. L. *The Emergency of Life: from Chemical Origins to Synthetic Biology*; Cambridge University Press: Cambridge, U.K., 2006.
(31) Schlund, S.; Schmuck, C.; Engels, B. *J. Am. Chem. Soc.* **2005**, *127*, 11115−11124.
(32) Hill, A. R.; Orgel, L. E. *Orig. Life Evol. Biosph.* **1996**, *26*, 539−545.
(33) Wang, K. J.; Nai, P.; Li, S. S. *Anal. Biochem.* **2004**, *332*, 199−201.
(34) Koshland, D. E., Jr. *Evolution of Catalytic Function, in Cold Spring Harbor Symposia on Quantitative Biology, LII, 1987*; Cold Spring Harbor Laboratory Press: New York, 1987; pp 1−7.
(35) Leberman, R.; Soper, A. K. *Nature (London)* **1995**, *378*, 364−366.
(36) Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1990**, *216*, 1031−1044.
(37) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997−10002.
(38) Collins, K. D. *Biophys. J.* **1997**, *72*, 65−76.
(39) Parsegian, V. A. *Nature (London)* **1995**, *378*, 335−336.
(40) Vlachy, V.; Hribar-Lee, B.; Kalyuzhnyi, Y. V.; Dill, K. A. *Curr. Opin. Colloid Interface Sci.* **2004**, *9*, 128−132.
(41) Collins, K. D.; Washabaugh, M. W. *Q. Rev. Biophys.* **1985**, *18*, 323−422.
(42) Gurau, M. C.; Lim. S. M.; Castellena, E. T.; Albertrio, F.; Kataoka, S.; Cremer, P. S. *J. Am. Chem. Soc.* **2004**, *126*, 10522−10523.
(43) Kunz, W.; Nostro, P. L.; Ninham, B. W. *Curr. Opin. Colloid Interface Sci.* **2004**, *9*, 1−18, and references cited therein.
(44) To be published: this group.

MA7019286