Modelling of prebiotic synthesis and selection of peptides under isothermal conditions and thermal cycling mode

O. V. Demina,[a] A. S. Kononikhin,[a,b] A. V. Laptev,[a] A. A. Khodonov,[a] E. N. Nikolaev,[a,b] and S. D. Varfolomeev[a]*

[a]*N. M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 4 ul. Kosygina, 119334 Moscow, Russian Federation. Fax: +7 (499) 137 4101. E-mail: ovd@sky.chph.ras.ru*
[b]*Institute for Energy Problems of Chemical Physics, Russian Academy of Sciences, 38/2 Leninsky prosp., 119334 Moscow, Russian Federation. Fax: +7 (499) 939 7370*

The model peptide synthesis from mixtures of amino acids was carried out under the thermal cycling and isothermal modes. The compositions of the obtained mixtures of products and the primary amino acid sequence of the synthesized peptides were determined by Fourier transform ion cyclotron resonance mass spectrometry and tandem mass spectrometry in combination with high-performance liquid chromatography with the application of *de novo* sequencing of the synthesized products. The processes of abiogenous synthesis of peptides were shown to occur under relatively mild temperature conditions and give a substantially less number of peptides as compared with the possible statistical set. The evolution of the system takes place in the process of the synthesis in solid phase with the disappearance of a series of the most unstable peptides. The selection process with the formation of complementary peptides takes place in peptide synthesis under the thermal cyclic mode.

Key words: amino acids, biopolymers, mass spectrometry, peptide synthesis, peptides, thermal synthesis, thermal cycling.

The problem of the appearance of self-replicating polymer molecules is the key question of life origin. At present, it is assumed that multiple low-molecular organic compounds, being the starting substances for the synthesis of proteins, nucleic acids, and sugars, had arisen on the early Earth due to various chemical and radiation processes.[1—14] This direction of the search for the methods of biomacromolecules synthesis with the modelling of its possible prebiotic conditions of the synthesis is developing intensively due to new physicochemical methods of investigation.

The emergence of information and catalytically functioning macromolecular structures, such as nucleic acids and proteins, is fundamentally important. The almost infinite variety of molecules different in real structure could appear due to the spontaneous formation of phosphodiester (nucleic acids) and peptide (proteins) bonds. The huge decrease in the variants of protein structures occurred during the origin of living systems and their consequent evolution, and a very limited number of catalytic sites of enzymes has been observed in the modern biological world.[15,16] The proteins and nucleic acids presently existing on the Earth are a very small part of the molecules that could appear due to the primary processes and evolve to real biological structures. This means phenomenologically that the process of the system "ordering" had been occurred with the drastic reduction of possible variants of macromolecules.[1]

Thus, it is necessary to explain the mechanisms of molecular convergence of macromolecular structures and their selection by some properties, as well as the appearance of the mechanism of their proliferation (self-replication) for the understanding the phenomenon of the origin of life. It is important that the advanced concepts would be based on the known fundamental laws of development of natural phenomena and would not violate the known thermodynamic limitations, in particular, the second law of thermodynamics.

The great attention is given to the problems of evolution of macromolecules at the prebiotic level.[10,17—22] In particular, outstanding book written by M. Eigen[18] gives the kinetic description of selection processes of macromolecules using the "competitive advantage" concept. However, beyond the scope of these studies, the questions remain unanswered about the mechanism of macromolecule self-reproduction processes and about the driving force providing ordering, "competitive advantage," and, finally, molecular convergence of polymer variety to a restricted number of types of biomacromolecules.

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 2, pp. 419—437, February, 2012.
1066-5285/12/6102-422

We suppose that the primary mechanism of self-replicating and evolution of macromolecules is interconnected with the thermocycling conditions.[23,24] As a result of rotation, any point of the Earth surface experiences cyclic temperature fluctuations. These fluctuations had an extremely high amplitude on the primary Earth with a rather weak atmosphere. These fluctuations could be characterized by a wide spectrum of amplitudes from −100 to 200 °C in dependence of specific conditions (radiation, heat exchange, thermal conductivity). A variety of chemical reactions and phase transitions are possible under these conditions.

Modern biomacromolecules are the products of polycondensations which occurred on the early Earth and resulted in the emergence of amide (proteins and peptides), phosphodiester (nucleic acids), and acetal (polysaccharides) bonds. In the formalized form, monomeric molecules should be trifunctional including two reactive groups essential for the polymerization and substituents allowing one to distinguish chemically one molecule from another. The substituents can interact with each other supramolecular with one or other selectivity determined by the free energies of complexes formation owing to hydrogen bonds and ion or hydrophobic interactions.

Thus, the following conditions are needed for the synthesis of prebiotic macromolecules:

(1) trifunctional monomers containing several types of substituents (for example, for amino acids: amino group, carboxyl group, and a substituent containing an additional functional group);

(2) excess of monomers;

(3) thermal cycling, *i.e.*, a cyclic change of temperature providing the reversible phase transition of water from the liquid to gaseous state and back.

The kinetic theory of prebiotic evolution under the thermal cycling conditions was described in detail in publications.[1,23] It is supposed that the process of prebiotic evolution occurred successfully with very little characteristics of "competitive advantage."

The purpose of this work is the experimental study of the synthesis of peptides from mixtures of oppositely charged amino acids under two possible synthesis conditions, such as an isothermal and a thermocyclic (thermal cycling) modes.

Experimental

The following reagents were used in experiments: amino acids L-lysine hydrochloride (L-Lys, K), L-aspartic acid (L-Asp, D), L-arginine hydrochloride (L-Arg, R), and L-glutamic acid (L-Glu, E) (special purity grade, Dia-M, Russia); sulfuric acid (special purity grade), hydrochloric acid (special purity grade), and sodium hydroxide (special purity grade) (produced in Russia).

The monitoring of reactions was performed by thin layer chromatography (TLC). Plates with silica gel 60 F_{254} (Merck), a mixture of methanol, chloroform, and 10% aqueous solution of ammonia in ratio 4 : 4 : 2 as an eluent and a 5% solution of ninhydrin in methanol as a developer were used for TLC.

^{1}H NMR spectra were recorded on Bruker AD-600 and Bruker DPX-300 devices (Germany) with a working frequency of 600 and 300 MHz, respectively, in D_2O and DMSO-d_6. Chemical shifts of the signals were determined relative to acetonitrile (MeCN, δ 1.98) as an external standard. The 2D COSY ^{1}H-^{1}H NMR correlation spectra were also recorded for the assignment of signals of separate groups.

Mass spectrometry of all samples was carried out on a Finnigan LTQ FT combined mass spectrometer (Thermo Electron, Germany) consisting from an ion cyclotron resonance mass spectrometer with Fourier transform (FT ICR) equipped with a superconducting solenoid of 7 T and a highly sensitive linear quadrupole ion trap. Mass spectra were recorded in the FT ICR mode using such ionization methods as the electrospray (ESI) and the matrix-assisted laser desorption/ionization (MALDI). Mass spectrometric analysis combined with chromatography was carried using tandem mass spectrometry in combination with high-performance liquid chromatography (HPLC-MS/MS), including the preliminary fractionation of samples by the nano flow high-performance liquid chromatograph Agilent 1100 (Agilent, US) on the self-made column 75 μm×15 cm with the reversed phase ReproSil Pur C18 (a particle diameter is 3 μm, a pore diameter is 100 Å (Dr. Maisch GmbH, Germany)) using the mobile phase H_2O—MeCN with the addition of formic acid to the concentration of 0.1 vol.% (MeCN gradient in the flow from 5 to 50% within 90 min at a flow rate of 0.3 μL min^{-1}) followed by the analysis of fractions by means of the measurement of exact masses with FT ICR mass spectrometer and the acquiring collision-induced dissociation (CID) spectra in a linear quadrupole ion trap in the m/z range from 300 to 2000.

Tandem mass spectrometry with CID as a fragmentation method was used for the confirmation of peptide nature of the molecular ions and the determination of possible structures of molecular ion fragments.

The compositions of the samples and the conditions of synthesis are given in Tables 1 and 2.

Synthesis of peptides and their derivatives under the thermal cycling mode from the mixture of amino acids using different pH values of the medium (general procedure). Aqueous solutions of mixtures of amino acids L-Asp and L-Arg (see Table 1, runs *1—4*), L-Asp and L-Lys·HCl (see Table 1, runs *5—8*) in ratios of 1 : 1 and 1 : 2, respectively, were prepared with the solutions pH reduced to 4.5 and 8.5 with HCl or NaOH. The volume of the initial solutions was 15—25 mL for L-Asp (0.68 mmol) and L-Lys·HCl (0.68 mmol), as well as for L-Asp (0.68 mmol) and L-Arg·HCl (0.68 mmol), at the 1 : 1 ratio. An aliquot of the initial solution (2 mL) was placed into a thermo-resistant beaker covered with a perforated lid, and water was evaporated by slow heating to 120, 130, or 140 °C. After water evaporation, the reaction mixture was heated for 12 or 24 h at 120 or 130 °C, cooled to ambient temperature, and dissolved in a minimal volume of water (~1 mL). A sample (50 μL) was taken, and a new aliquot of the starting solution of amino acids was added. The cycle was repeated 7—9 times. After the last sample taking, the reaction mixture was heated for 36 h without the addition of the initial mixture aliquot for subsequent comparison with the penultimate probe. The synthesis of peptides from mixtures of L-Glu and L-Lys·HCl, L-Ala and L-Gly was done similarly

Table 1. Reagents, samples, and conditions of the thermal cycling synthesis

Run	Reagents and conditions	Probe 1		Probe 2		Probe 3*	
		Sample	*t*/h (*N*)	Sample	*t*/h (*N*)	Sample	*t*/h (*N*)
1	L-Arg·HCl—L-Asp (1 : 1), 120 °C, pH 4.5	**1A**	36 (3)	**1B**	85.5 (7)	**1C**	121.5 (7)
2	L-Arg·HCl—L-Asp (1 : 1), 120 °C, pH 8.5	**2A**	36 (3)	**2B**	85.5 (7)	**2C**	121.5 (7)
3	L-Arg·HCl—L-Asp (2 : 1), 120 °C, pH 4.5	**3A**	36 (3)	**3B**	85.5 (7)	**3C**	121.5 (7)
4	L-Arg·HCl—L-Asp (2 : 1), 120 °C, pH 8.5	**4A**	36 (3)	**4B**	85.5 (7)	**4C**	121.5 (7)
5	L-Lys·HCl—L-Asp (1 : 1), 120 °C, pH 4.5	**5A**	36 (3)	**5B**	108 (9)	**5C**	144 (9)
6	L-Lys·HCl—L-Asp (1 : 1), 120 °C, pH 8.5	**6A**	36 (3)	**6B**	108 (9)	**6C**	144 (9)
7	L-Lys·HCl—L-Asp (2 : 1), 120 °C, pH 4.5	**7A**	36 (3)	**7B**	108 (9)	**7C**	144 (9)
8	L-Lys·HCl—L-Asp (2 : 1), 120 °C, pH 8.5	**8A**	36 (3)	**8B**	108 (9)	**8C**	144 (9)

Note. *t* is the heating time, and *N* is the number of cycles.
* After taking probe 2, the sample was heated for 36 h more without addition of an aliquot of the initial amino acid mixture.

(the same amounts of the catalysts and ratios of the mixture components).

Synthesis of peptides and their derivatives in the solid phase from the mixture of amino acids using various catalysts (see Table 2, runs *9*—*13*). Amino acids L-Asp (3.4 mmol) and L-Lys·HCl (3.4 mmol) were put in the thermo-resistant beaker, and the catalysts were added, such as concentrated sulfuric acid (run *9*: 2 equiv.; run *10*: 3 equiv.), hydrochloric acid (run *11*: 3 equiv.), or sodium hydroxide (run *13*: 4 equiv.) (see Table 2). The mixture was mixed and heated in the drying cabinet at 130 °C. Experiment *12* was carried out as control without any catalyst but with the addition of distilled water (2 mL). Samples were taken after 104 h (probe 1) and 278 h (probe 2) of the heating. After prolonged heating (104 h and more) the reaction mixture of samples **9A**—**11A** and **9B**—**11B** became semitransparent and glassy with a slight yellowish tint. The reaction mixture of samples **12A**—**13A** and **12B**—**13B** remained visually almost unchanged.

Synthesis of peptides and their derivatives in the solid phase from a mixture of amino acids using sulfuric acid as a catalyst (see

Table 2. Reagents, samples, and conditions of the synthesis in solid phase

Run	Reagents and conditions	Probe 1		Probe 2	
		Sample	*t*/h	Sample	*t*/h
	Series I				
9	L-Lys·HCl—L-Asp (1 : 1), 130 °C, 2 equiv. H_2SO_4	**9A**	104	**9B**	278
10	L-Lys·HCl—L-Asp (1 : 1), 130 °C, 3 equiv. H_2SO_4	**10A**	104	**10B**	278
11	L-Lys·HCl—L-Asp (1 : 1), 130 °C, 3 equiv. HCl	**11A**	104	**11B**	278
12	L-Lys·HCl—L-Asp (1 : 1), 130 °C, H_2O	**12A**	104	**12B**	278
13	L-Lys·HCl—L-Asp (1 : 1), 130 °C, 4 equiv. NaOH	**13A**	104	**13B**	278
	Series II				
14	L-Lys·HCl—L-Asp (1 : 1), 130 °C, 2 equiv. H_2SO_4	**14A**	138	**14B**	300
15	L-Lys·HCl—L-Asp (1 : 2), 130 °C, 3 equiv. H_2SO_4	**15A**	138	**15B**	300
16	L-Lys·HCl—L-Asp (2 : 1), 130 °C, 3 equiv. H_2SO_4	**16A**	138	**16B**	300
17	L-Asp, 130 °C, 2 equiv. H_2SO_4	**17A**	138	**17B**	300
18	L-Asp, 130 °C, 4 equiv. H_2SO_4	**18A**	138	**18B**	300
	Series III				
19	L-Arg·HCl—L-Asp (1 : 2), 140 °C, 3 equiv. H_2SO_4	**19A**	100	**19B**	300
20	L-Arg·HCl—L-Asp (1 : 1), 140 °C, 2 equiv. H_2SO_4	**20A**	100	**20B**	300
21	L-Arg·HCl—L-Asp (2 : 1), 140 °C, 3 equiv. H_2SO_4	**21A**	100	**21B**	300
22	L-Lys·HCl—L-Glu (1 : 2), 140 °C, 3 equiv. H_2SO_4	**22A**	100	**22B**	300
23	L-Lys·HCl—L-Glu (1 : 1), 140 °C, 2 equiv. H_2SO_4	**23A**	100	**23B**	300
24	L-Lys·HCl—L-Glu (2 : 1), 140 °C, 3 equiv. H_2SO_4	**24A**	100	**24B**	300
25	L-Arg·HCl—L-Glu (1 : 2), 140 °C, 3 equiv. H_2SO_4	**25A**	100	**25B**	300
26	L-Arg·HCl—L-Glu (1 : 1), 140 °C, 2 equiv. H_2SO_4	**26A**	100	**26B**	300
27	L-Arg·HCl—L-Glu (2 : 1), 140 °C, 3 equiv. H_2SO_4	**27A**	100	**27B**	300

Note. *t* is the heating time, and *N* is the number of cycles.

Table 2, runs *14—18*). The synthesis was carried out according to the above described procedure from a mixture of L-Asp (3.4 mmol) and L-Lys · HCl (3.4 mmol) adding various amounts of concentrated sulfuric acid as the catalyst (see Table 2). Samples were taken after 138 and 300 h the heating. The ratio of amino acids was also varied, *i.e.*, 6.8 mmoles of L-Asp and 3.4 mmoles of L-Lys · HCl were taken for the preparation of samples containing L-Asp and L-Lys · HCl with the ratio equal to 2 : 1; 3.4 mmoles of L-Asp and 6.8 mmoles of L-Lys · HCl were taken for the samples L-Asp : L-Lys · HCl in a 1 : 2 ration. Control samples contained only L-Asp (7.5 mmol) and various amounts of sulfuric acid (see Table 2, runs *17* and *18*). The syntheses of peptides from L-Lys · HCl—L-Glu, L-Arg—L-Glu, and L-Arg—L-Asp mixtures were carried out using the same amounts of sulfuric acid and the same ratios of the components (see Table 2).

Alkaline hydrolysis of the obtained reaction mixtures. The hydrolysis of the reaction mixture was performed according to the modified procedure.[25] The calculated amount of sodium hydroxide solution in water (2.12 *M* NaOH, 0.425 g of sodium hydroxide per 1 g of the reaction mixture) was placed in a 25-mL round-bottom flask equipped with a magnetic stirrer. The mixture was cooled to 0 °C and then the reaction mixture of the sample was added by several portions. After the sample was completely dissolved under the stirring, the resulting mixture was heated to ambient temperature, stirred for 1 h, and then acidified to pH 6. The obtained mixture was added by portions to methanol (40 mL) cooled to 0 °C. The formed white precipitate was filtered off and dried under the reduced pressure. The quantities of polymers in the samples of runs *14—18* after the treatment were the following: in sample **14B** (2 equiv. H_2SO_4), 0.18 g of the polymer mixture was obtained from 0.51 g of the treated reaction mixture; in sample **15B** (3 equiv. H_2SO_4), 0.22 g of the polymer mixture was obtained from 0.5 g of the treating reaction mixture; in the sample **17B** (2 equiv. sulfuric acid), 0.3 g of the polymer mixture was obtained from 0.5 g of the treating reaction mixture with the total yield >50%; in the sample **18B** (4 equiv. H_2SO_4), 0.3 g of the polymer mixture from 0.65 g of the treating reaction mixture was obtained with the total yield about 50%.

Mass spectrometry of the products of amino acids polycondensation. The mass spectrometric analysis of mixtures of the obtained reaction products (see Tables 1 and 2) revealed that the reaction mixtures contained products of amino acids polycondensation (peptides and their derivatives) and unreacted initial amino acids. FT ICR mass spectrometry was used for the study of the obtained reaction mixtures compositions containing minor amounts of peptides (no more than 2—5%, according to the ^{1}H NMR spectroscopic data (see below).

The next conditions for mass spectra recording were chosen in order to determine the composition of polymers of series I and II (see Table 2): an acetonitrile—water mixture with the addition of formic acid to the concentration 0.1 vol.% as a solvent, the use of several ionization methods (electrospray and MALDI), and the detection of both positive and negative ions for the same sample. The system was washed for 15 min with 95% MeCN and then it was cleaned for 5 min with water containing 0.1% TFA. It was found that the degree of sample dilution influenced very strongly to the determination of its composition. The electrospray was the best ionization method for mass spectrometric analysis of complex mixture of polymers and peptides, since we failed to select the appropriate matrix for mixtures of so different products in the case of the MALDI-TOFF mass spectrometry usage.

The method of ion identification by tandem mass spectrometry with CID as an ionization method (see below) was used for the confirmation of peptide bonds presence and the identification of separate ions structure in the mass spectrum.

The formation of diketopiperazines for each mixture of amino acids during thermal synthesis was indicated at the decoding of the mass spectra, for example, the peak with m/z 257.18 at $z = 1$ is assigned to diketopiperazine formed during heating from two lysine molecules with the elimination of two water molecules.

Thus, it must be noted the use of both positive and negative modes, as well as complementary ionization and fragmentation methods for sample analysis are necessary in order to obtain more reliable information about the mixture composition.

The data for samples **9A** and **9B** are presented as examples of the results obtained by HPLC-MS/MS for the samples of runs *9—14* (see below).

Results and Discussion

Mixtures of oppositely charged amino acids were used as the initial reagents in the experimental study of the peptide synthesis. We used L-amino acids bearing a charge in the side chain, namely, aspartic (L-Asp, D) and glutamic (L-Glu, E) acids with the negatively charged carboxyl group of the side chain and lysine (L-Lys, K) or arginine (L-Arg, R) with the positively charged amino or guanidinium groups). The possibility of the synthesis of peptides in these model systems was studied under isothermal (thermal synthesis mode) and thermal cycling conditions (thermal cycling mode). The prepared dry mixtures were heated for a certain time under the thermal synthesis mode in solid phase. Under the thermal cycling mode, the initial cycle was multiply repeated, namely, the heating of an aliquot of solution of amino acids initial mixture — the phase transition—the polycondensation of monomers in the solid phase — the cooling — water condensation — the interaction of monomers with synthesized polymer.

The polycondensation process in the isothermal mode was carried out in a wide range of initial conditions using the catalysts (see Table 2, series I, runs *9—13*): (1) concentrated sulfuric acid, (2) concentrated hydrochloric acid, (3) sodium hydroxide, and (4) distilled water as a control. The certain amount of each catalyst was added to the thoroughly stirred mixture of amino acids. It was found that such polymerization of amino acids in the solid phase at the isothermal synthesis mode can occur in an acidic and a neutral media as well as in an alkaline medium. This process can be catalyzed by acids, bases, and metal cations. The chosen reaction conditions overlapped practically the whole range of possible acid—alkaline conditions of the prebiotic Earth[26,27] (pH from 4.2 to 8.2). It was shown that the polymerization was effective under these conditions. The reaction rate and the degree of conversion depended

on the temperature and the duration of the process. The polymerization occurred under relatively mild conditions at temperatures from 120 to 140 °C and did not achieve the high degree of polymerization, which, however, could increase under the certain conditions,[25,28] for instance, when polyphosphoric acid was used as a catalyst and the heating temperature was higher than 160 °C. The low degree of polymerization was the necessary condition for the detailed analysis of the obtained mixtures of polypeptides. The reaction conditions for the thermal cycling mode were chosen in such a way that the degree of conversion would not exceed 5% per one cycle.

The determination of polymers structures was carried out with the usage of *de novo* sequencing by means of FT ICR mass spectrometry and tandem mass spectrometry in combination with high-performance liquid chromatography (HPLC-MS/MS). Thus, all synthesized peptides in the obtained complicated mixtures of the products of amino acids polycondensations and their primary amino acid sequences were first defined by the HPLC-MS/MS methods.

Synthesis of peptides. Polymerization of amino acids in the solid phase under isothermal conditions. The conditions for the formation of polymer (peptide) chains were chosen for abiogenic thermal synthesis in model systems at the periodical mode in the presence of catalysts: mineral acids (concentrated sulfuric or hydrochloric acids) or base (sodium hydroxide) for a mixture of L-Lys and L-Asp (see Table 2, series I). Amino acids L-Lys and L-Asp were used in the ratio 1 : 1 without a solvent (water) with the addition of different equivalents of sulfuric and hydrochloric acids or sodium hydroxide. It should be mentioned that polyaspartic acid derivatives, polysuccinimides with a small amount of lysine, were mainly obtained, as well as small amounts of peptides. After 278 h of heating of sample **9B** (2 equiv. H_2SO_4), polymers containing from 2 to 12—13 residues were obtained, whereas for sample **10B** (3 equiv. H_2SO_4) the number of amino acid residues in the polymers varied from 2 to 8—11. Both polyaspartic acid derivatives and usual peptides of rather diverse composition, which underwent the cyclization under the further heating to form polyaspartic acid derivatives, were obtained in the sample **11A** (HCl as a catalyst). After 278 h of heating, predominantly polymers containing from 2 to 7—8 units (the destruction of the terminal fragment is possible) were obtained in a mixture of sample **11B**. The initial polymerization of aspartic acid was observed from 2 to 6—8 (very weak signals) residues at the noise level, and clusters of amino acids were observed too in the control sample **12B** (water). For sample **13B** (NaOH as a catalyst), the mass spectrum exhibited only peptides (di-, tri-, tetra-, and pentapeptides) in the absence of aspartic acid cyclization, and the amount of peptides did not exceed 2—5% (according to the ^{1}H NMR spectroscopic data).

The polymers, polysuccinimides with a high molecular weight, being polyaspartic acid derivatives, have been prepared earlier using the polyphosphoric acid as a catalyst at temperatures higher than 160 °C under nitrogen and a low vacuum.[25,28,29] The alkaline hydrolysis of these polysuccinimides mixtures followed by the structures determination of the obtained polyaspartic acids by NMR spectroscopy revealed the presence of various terminal fragments, some of which were not identified.[25,28,29]

We determined the structures of polymers obtained in the mixtures of amino acids and from L-Asp only and also identified the terminal fragments by FT ICR mass spectrometry.

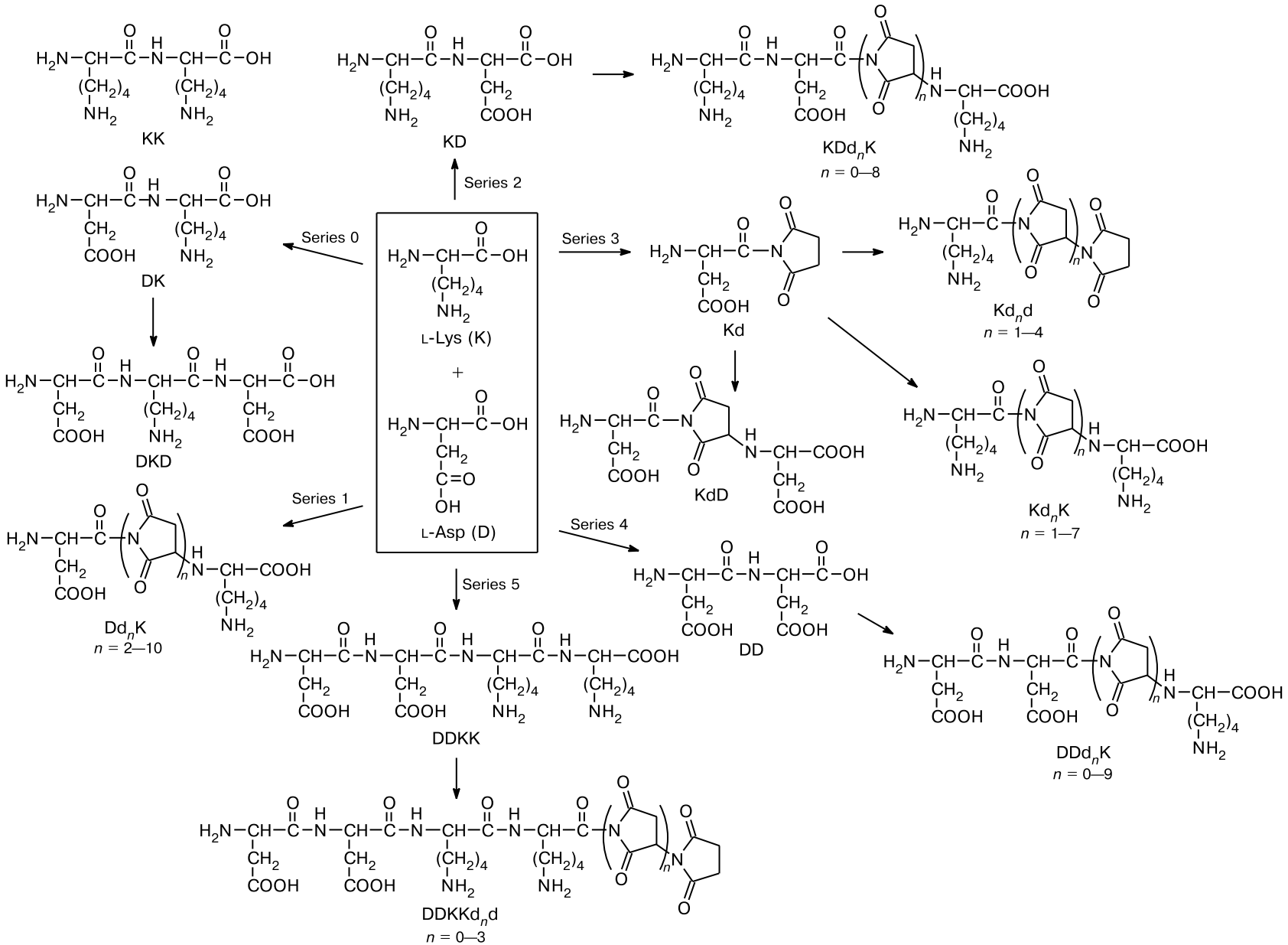
The duration of amino acid polycondensation determined the composition of the polymerization products. The system evolved with time towards a decrease in the variety and the number of possible structures.

The compositions of products mixtures for **9A** and **9B** samples obtained by the FT ICR and HPLC-MS/MS mass spectra interpretation are presented in Table 3 as the examples (see also Table 2, series I, run *9*; Schemes 1 and 2). Probe 1 and probe 2 were different only by the heating time, *i.e.*, probe 1 was heated for 104 h and probe 2 was heated for 278 h.

The comparison of the products obtained in samples of run *9* (see Table 2) in dependence of the heating time at the same temperature shows that the increase in the reaction time leads to the chain elongation in products of series 3 and 1. The decomposition processes of unstable structures and the formation of new structures occur in the solid phase under the permanent heating. It was observed that the variety of products becomes narrower under the increase in heating time, *viz.*, series 5 disappears (see Table 3, Schemes 1 and 2). The peptides obtained experimentally in very small and even trace amounts were revealed and identified by HPLC-MS/MS (see Figs 1 and 2, respectively). This shows the efficiency of the technique used by us for the identification of obtained products in complex mixtures. After incubation of the system during 104 h, we unambiguously identified the peptide derivatives KDddddddddK and the series of products: DDKK, DDKKd, DDKKdd, DDKKddd, and DDKKdddd (d is the succinimide residue), which were not revealed after 278 h of heating.

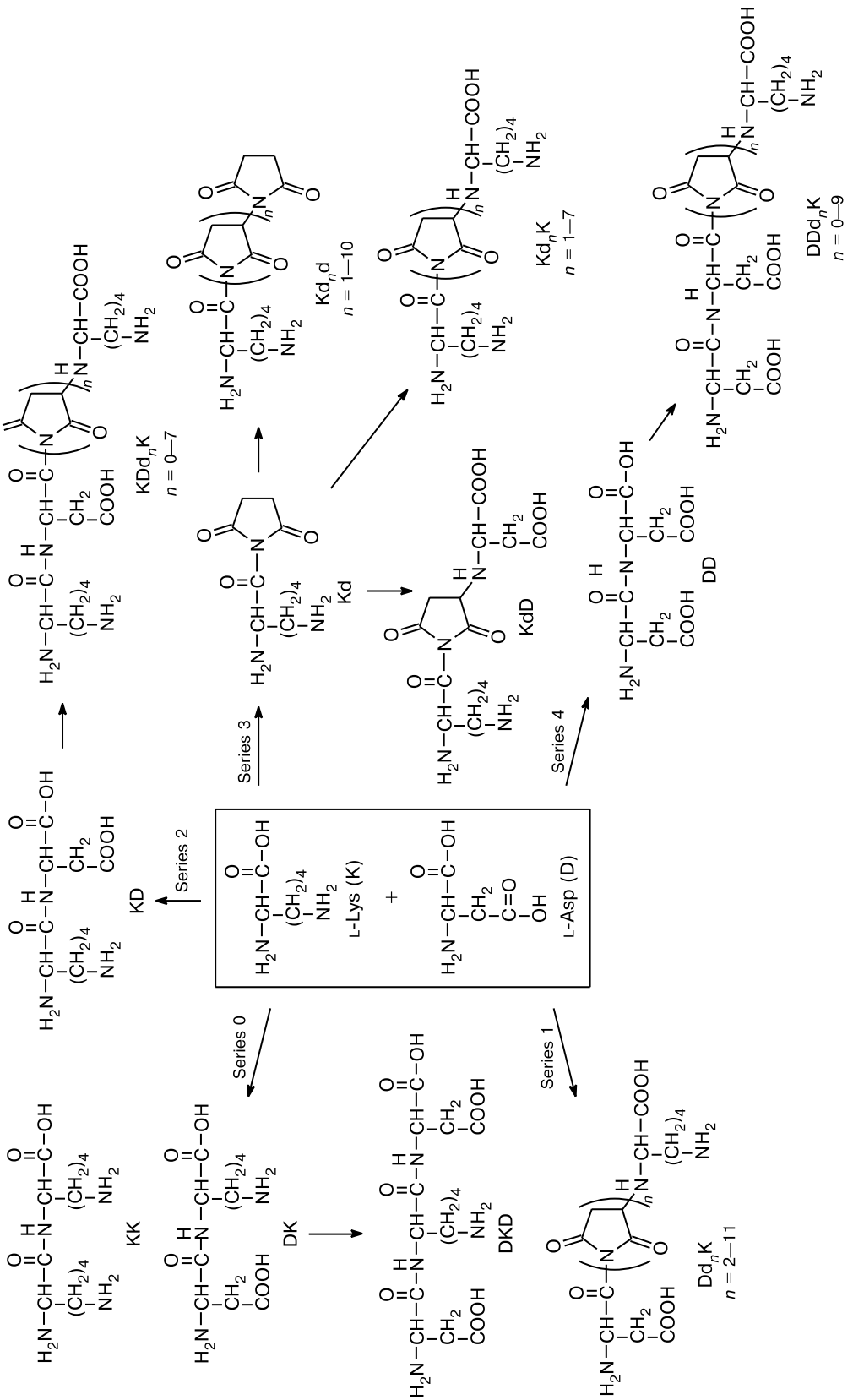
The sequences of amino acids residues or their derivatives were determined for the structures of the same series. The *de novo* sequencing and the determination of relative probabilities were performed for all structures presented in Schemes 1 and 2 by means of the processing programs for mass spectra. The peptide nature was confirmed for ions with m/z 501.1 and 271.2, *i.e.*, the molecular ion with m/z 501.1 corresponds to the ion $[NH_2 - Asp - Asp - Asp - Asp - OH + Na]^+$, the molecular ion with m/z 386.08 corresponds to the ion $[NH_2 - Asp - Asp - Asp - OH + Na]^+$, the molecular ion with m/z 271.18 corresponds to the ion $[NH_2 - Asp - Asp - OH + Na]^+$, and the molecular ion with m/z 156.0 corresponds to the ion $[Asp + Na]^+$ at $z = 1$ for all molecular ions (Fig. 3).

Scheme 1



Reagents and conditions: L-Lys : L-Asp = 1 : 1, 130 °C, 104 h (see Table 2, run *9*, sample **9A**). Here and in Schemes 2 and 3, d is the succinimide residue.

Scheme 2



Reagents and conditions: L-Lys : L-Asp = 1 : 1, 130 °C, 278 h (see Table 2, run *9*, sample **9B**).

Table 3. Composition of the mixture of products determined by the HPLC-MS/MS method for samples **9A** and **9B**

Number of residues	Formula*,**	Molecular ion $[M + H]^+$, m/z
	Series 0	
2	DK	262.1409
2	KK	275.2089
3	DKD	377.16783
	Серия 1	
4	DddK	456.1736
5	ДdddK	553.19
6	DddddK	650.2064
7	DdddddK	747.2228
8	DddddddK	844.2391
9	DdddddddK	941.2555
10	DddddddddK	1038.272
11	DdddddddddK	1135.288
12	DddddddddddK	1232.305
13	DdddddddddddK#	1329.31
	Series 2	
2	KD	262.1409
3	KDK	390.23585
4	KDdK	487.25223
5	KDddK	584.2686
6	KDdddK	681.28498
7	KDddddK	778.30135
8	KDdddddK	875.31773
9	KDddddddK	972.3341
10	KDdddddddK	1069.3505
11	KDddddddddK##	1166.3669
	Series 3	
2	Kd	244.1303
3	KdK	372.2253
3	Kdd	341.1467
3	KdD	359.1573
4	KddK	469.2417
4	Kddd	438.1631
5	Kdddd	535.1795
5	KdddK	566.258
6	Kddddd	632.1958
6	KddddK	663.2744
7	Kdddddd#	729.2122
7	KdddddK	760.2908
8	Kddddddd#	826.2286
8	KddddddK	857.3072
9	Kdddddddd#	923.245
9	KddddddddK	954.3235
10	Kddddddddd#	1020.261
11	Kdddddddddd#	1117.278
12	Kddddddddddd#	1214.294
	Series 4	
2	DD	249.0644403
3	DDK	377.16783
4	DDdK	474.18421
5	DDddK	571.20058
6	DDdddK	668.21696
7	DDddddK	765.23333
8	DDdddddK	862.24971
9	DDddddddK	959.26608
10	DDdddddddK	1056.2825
11	DDddddddddK	1153.2988
12	DDdddddddddK	1250.3152
	Series 5	
4	DDKK##	505.2628
5	DDKKd##	602.2792
6	DDKKdd##	699.2955
7	DDKKddd##	796.3119
8	DDKKdddd##	893.3283

* d is the succinimide residue.

** Products marked by index # were not identified for the reaction time of 104 h but were identified for the reaction time of 278 h; products marked by ## were identified for the reaction time of 104 h and were not identified after 278 h.

The peptides NH_2—Lys—Asp—COOH (KD) and NH_2—Asp—Asp—COOH (DD) may be initial compounds for a number of presented series: NH_2—Lys—Asp—COOH for series 2 and 3, and NH_2—Asp—Asp—COOH for series 1 and 4. Possibly, the peptides of series 5 are formed from dipeptides NH_2—Asp—Asp—COOH and NH_2—Lys—Lys—COOH (KK), without the exclusion the possibility of consequent growth of the peptide chain (see Scheme 1 and Table 2, run *9*, sample **9A**). Probably, the number of peptides and their derivatives included the Lys—Lys fragment is too insignificant, and we did not observe the structures with this fragment by means of HPLC-MS/MS, although similar structures were revealed by the preliminary interpretation of the mass spectra without HPLC.

The analysis of the obtained reaction products in experiments of series I (see Table 2) indicated that concentrated sulfuric acid was the most efficient catalyst of the process.

Based on the results of studies for the choice of conditions for the formation of polymer structures with masses more than 1000 Da (see Table 2, series I) in order to increase the yield of the polycondensation products, the influence of concentrated sulfuric acid amounts on the polypeptides and polymers synthesis in the Lys—Asp system was investigated with the addition of 2, 3, and 4 equiv. H_2SO_4 under the heating at 130 °C without a solvent for a long time (138—300 h) (see Table 2, series II, runs *14—18*, and Experimental). Dry amino acids, L-Lys

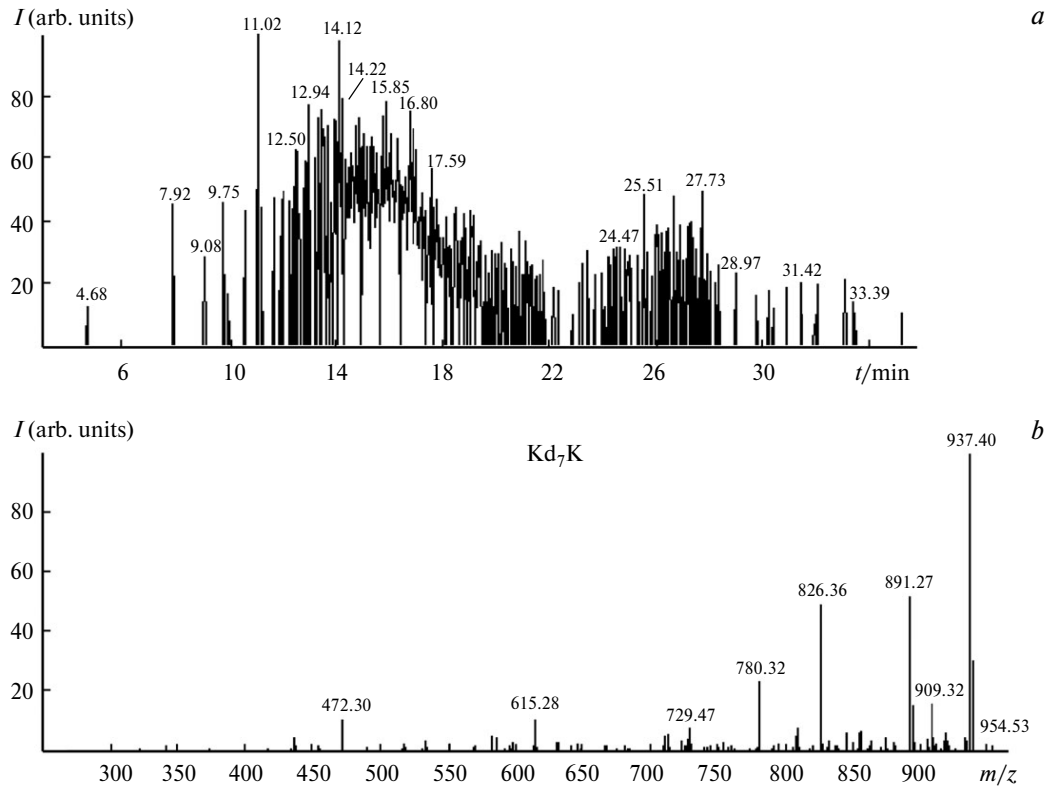


Fig. 1. The HPLC-MS/MS data for the sample **9B**: *a*, a peak in the mass chromatogram corresponding to the fraction containing peptide with a relatively higher molecular weight; *b*, the mass spectrum of the CID fragmentation products obtained with the linear quadrupole ion trap. The amino acid sequence of the peptide (Kd_7K, where d is the succinimide residue) was calculated manually using the MS/MS data.

and L-Asp, were taken in ratios of 1 : 1 and 1 : 2 and L-Asp only was used for the comparison and control.

According to TLC data, the reaction mixtures of samples **14A**, **14B**, **15A**, and **15B** contained Asp, Lys, polysuccinimide, and various polymer products of reaction, whereas Asp and polysuccinimides were observed in the samples of experiments *17* and *18*. According to the mass spectrometric analysis data, the reaction mixture obtained from a mixture of Lys and Asp even after 300 h of heating contained the structures with the aspartic acid residue instead of the succinimide fragment. This was not observed for the samples consisting of aspartic acid only. The obtained experimental data allow one to suggest the spontaneous polymerization of aspartic acid at the beginning followed by the cyclization of the polyaspartic acid to polysuccinimide.

Mass spectra of samples **14B** and **18B** of series II are presented in Fig. 4 (see Table 2). The period equal to 97 and corresponding to the mass of the succinimide residue is well seen in the mass spectrum of sample **14B** (see Fig. 4, *a*). Periods of 97 characteristic of succinimide residues in the polymer chain were revealed in the mass spectra of sample **18B** (see Fig. 4, *b* and *c*) in the final reaction mixture (before alkaline hydrolysis). It should be mentioned that the mass of these polysuccinimide derivatives increases under the registration of negative ions in comparison with that under the resitration of positive ions. The mass spectrum of a polymer mixture of sample **18B** after the alkaline treatment and acidification with positive ion registration is shown in Fig. 4, *c*, and a period of 115 corresponds to the mass of the aspartic acid residue and unambiguously indicates the presence of Asp residues as polymer chain units, but we did not determine the ratio of α- and β-isomers. Since the alkaline hydrolysis is not mild, the racemization of the final product occurs in these samples.

It was established that under the interpretation of the mass spectra the polyaspartic acid containing from 20 to 23 units (when recording negative ions) or 17—18 units (when recording positive ions) was formed for the same sample after alkaline hydrolysis and acidification of the dried precipitate from samples **17B** and **18B** of series II (see Table 2). This means that the detection of both positive and negative ions is necessary when recording similar mass spectra.

Using the mass spectra of the final reaction mixtures of samples **14B**, **15B**, and **16B** (before alkaline hydrolysis), we obtained the data confirming the formation of peptide

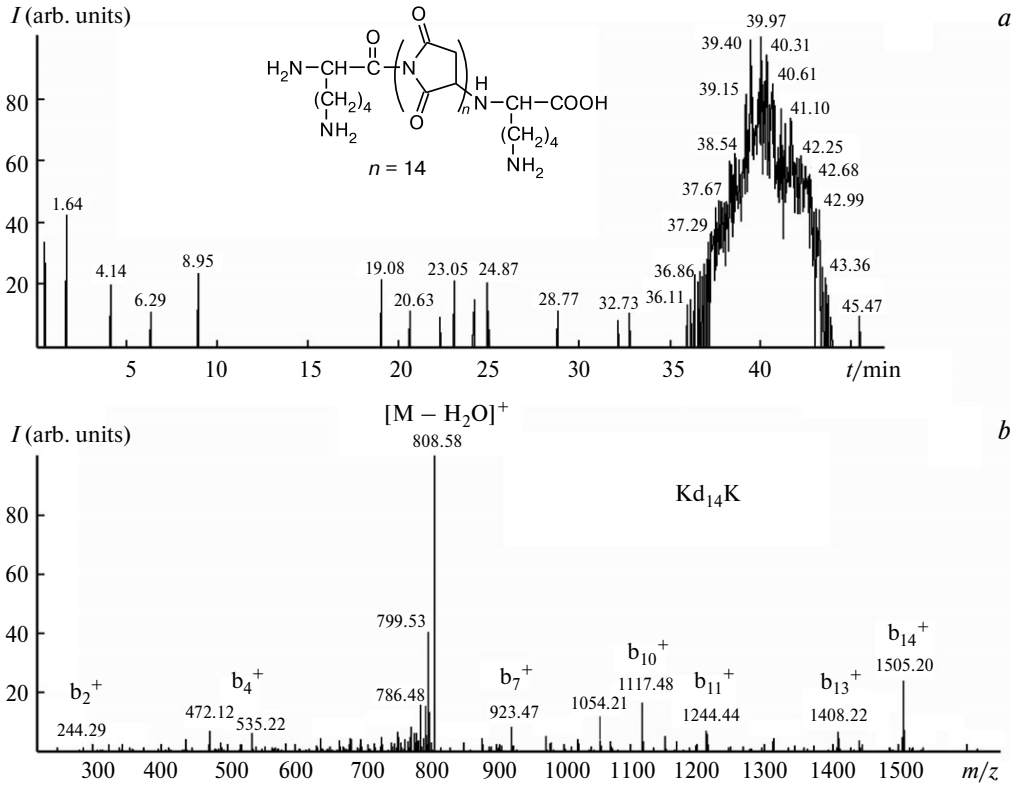


Fig. 2. The HPLC-MS/MS data for the sample **9B**: *a*, a peak in the mass chromatogram corresponding to the fraction containing a trace amount of a high-molecular-weight peptide; *b*, the mass spectrum of the CID fragmentation products obtained with a linear quadrupole ion trap. The amino acid sequence of the peptide was calculated manually using the MS/MS data.

bonds under the heating of amino acids mixtures. The samples of this series are characterized by the following components: the succinimide residue with a molecular weight of 97, the Asp residue with a molecular weight of 115, and the Lys residue with a molecular weight of 128, water molecules and Na ions. The peak with m/z 469.4 at $z = 1$ corresponds to the cation $[NH_2 - Lys - bis(succinimidyl) - Lys - OH + H]^+$.

Under polymerization conditions used by us, a series of peptide derivatives are formed: polysuccinimides for L-Asp and polysuccinimide fragments in the products of an L-Asp—L-Lys mixture. The formation of diketopiperazines and trace amounts of ethylamine is also observed. The alkaline hydrolysis of samples **14B**—**16B** gave polyaspartic acids of variable composition with different terminal groups, such as residues of lysine, aspargine, and aspartic acid and, possibly, ethylamine residue (see Experimental) whereas before the treatment of these samples, the succinimide and 3-aminosuccinimide fragments and residues of lysine and aspartic acid are observed. It follows from the results of hydrolysis that the amount of the mixture of obtained polymer products was greater when using 3 equiv. H_2SO_4 for a mixture of amino acids (sample **15B**) than that when 2 equiv. H_2SO_4 were used (sample **14B**): 0.22 g and 0.18 g, respectively, from 0.5 g of the initial mixture. For control samples containing only Asp, the amount of the mixture of obtained polymer products was greater for the use of 2 equiv. H_2SO_4 as

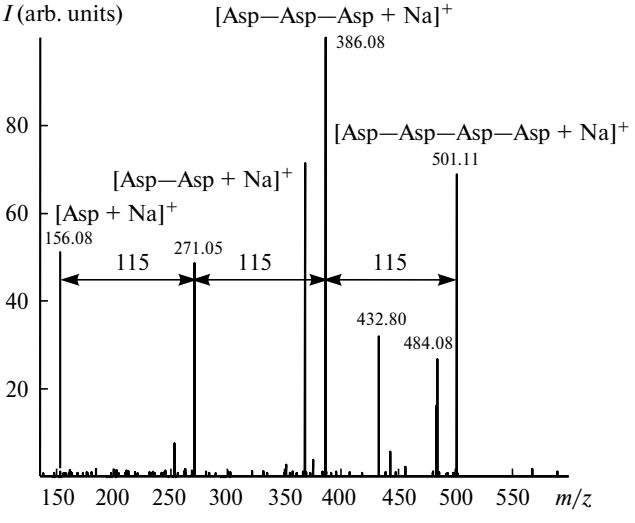


Fig. 3. The use of tandem mass spectrometry to confirm the peptide nature of the peak with m/z 501.11 in the mass spectrum of sample **18B** after 300-h heating and the treatment of the reaction mixture with detection of positive ions.

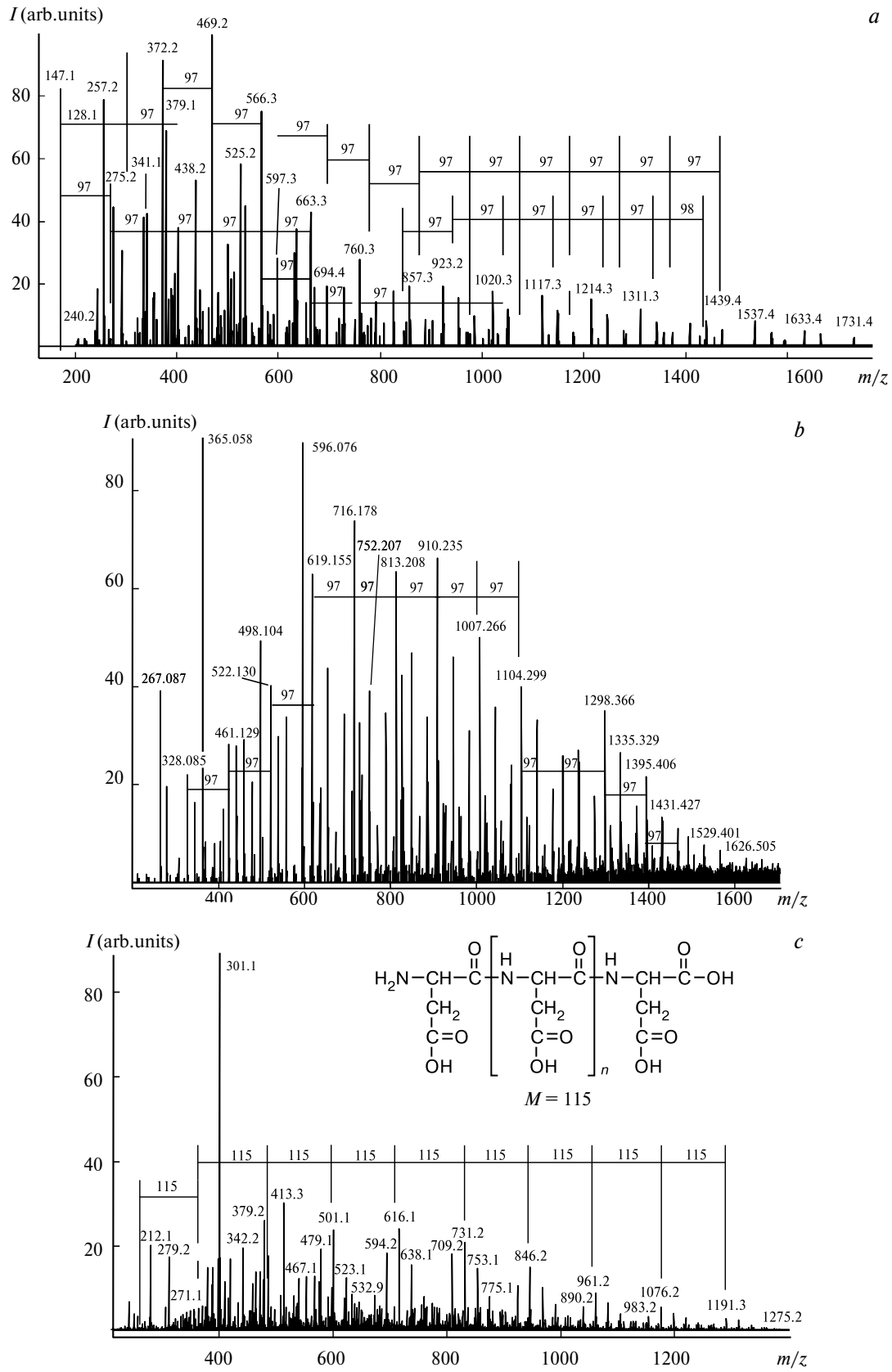


Fig. 4. Mass spectra of sample **14B** (see Table 2, run *14*) before the alkaline treatment using FT ICR with MALDI for the detection of positive ions (*a*) and the reaction mixture of sample **18B** before (*b*) and after the alkaline treatment of the reaction mixture (*c*) with detection of positive ions. The samples were heated for 300 h.

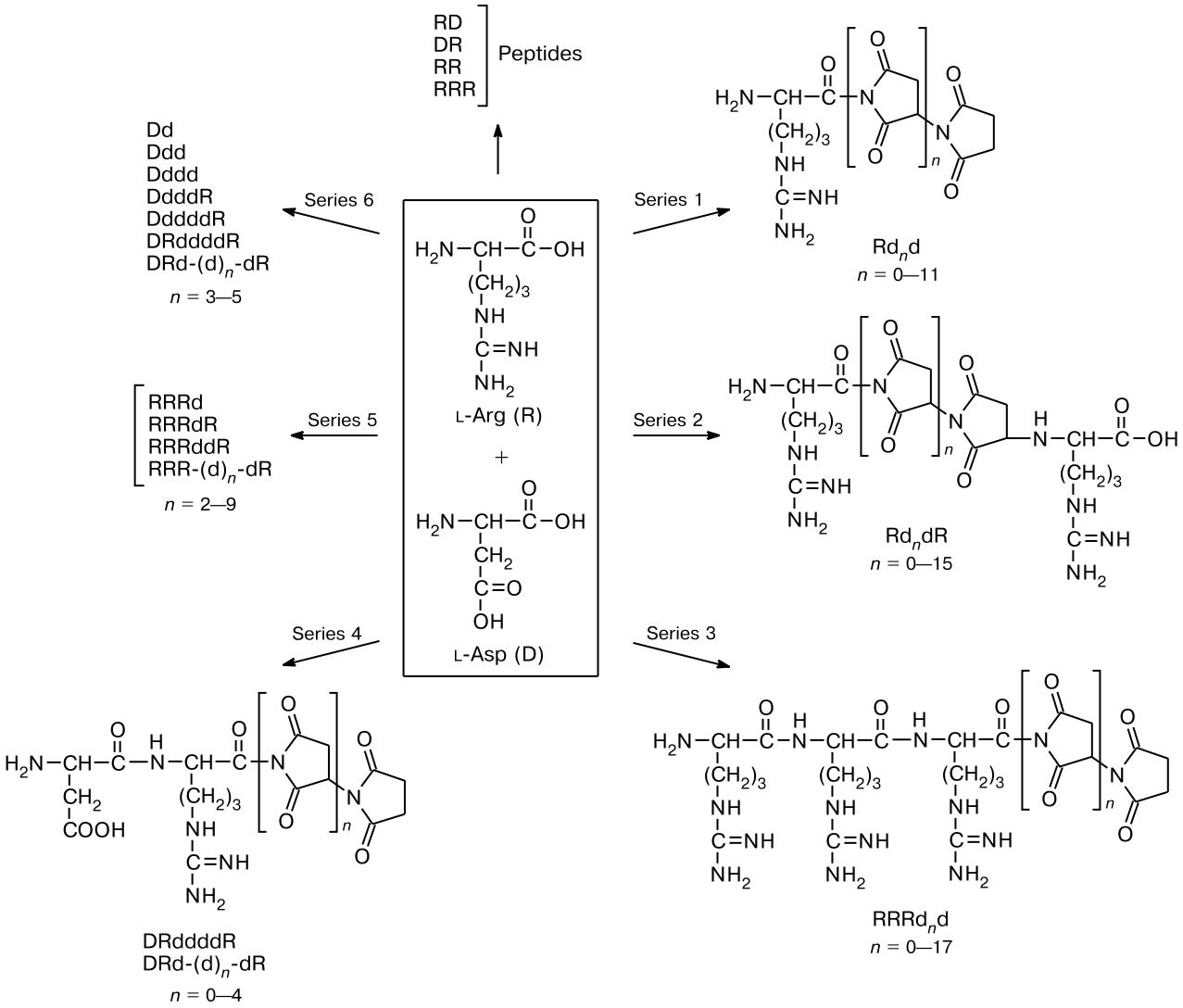
a catalyst (sample **17B**) than that for the use of 4 equiv. H_2SO_4 (sample **18B**), *i.e.*, 0.3 g from 0.5 g and 0.3 g from 0.65 g of the initial mixture, respectively.

When 2 equiv. H_2SO_4 were used as a catalyst, the number of residues in peptides attained 21—23 for Asp (sample **17B**) and 14—20 for a mixture of amino acids (sample **14B**), decreasing to 10—12 residues when 3 equiv. H_2SO_4 were used (sample **15B**). It should be noted that a decrease in the amount of amino acid containing an additional amino group in a mixture of amino acids at the same amount of sulfuric acid results in polymers with a smaller number of residues in the chain, which is characteristic of all pairs of amino acids used by us.

We analyzed the samples of experiments *19—27* of series III: samples **19A—27A** after 138 h of heating (probe 1) and samples **19B—27B** after 300 h of heating (probe 2) (see Table 2).

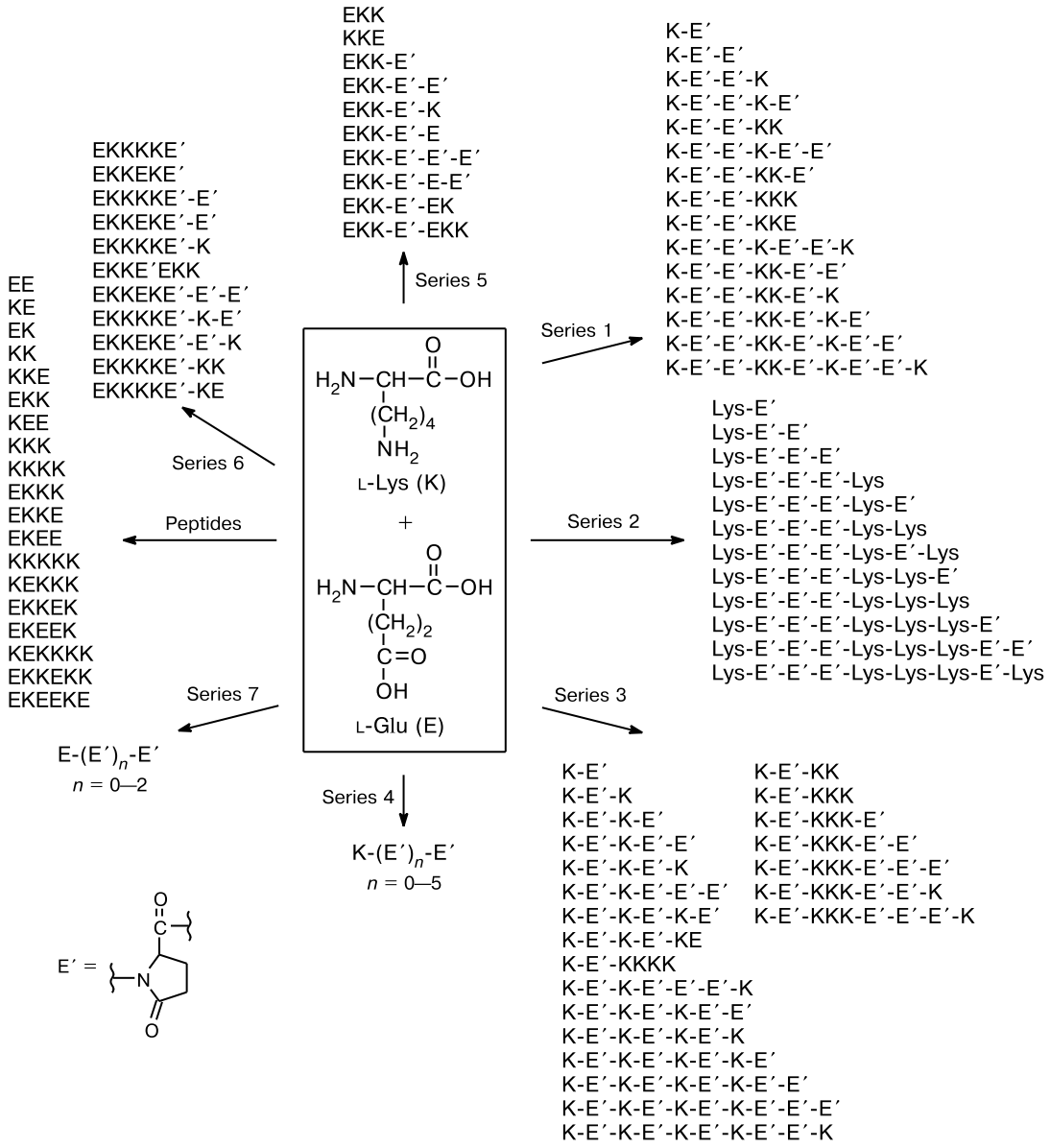
The synthesis of peptides from mixtures of L-Asp with L-Arg (Scheme 3), L-Glu with L-Lys, or L-Glu with L-Arg under the same reaction conditions gave mixtures of peptides and biopolymers, *i.e.*, peptide derivatives containing the pyroglutamic acid fragment. The composition of the mixture of obtained products for the ratio of amino acids L-Arg : L-Asp = 1 : 1 is shown in Scheme 3. According to the obtained experimental data, the oligomerization rate of L-Glu was lower than that of L-Asp and, hence, the composition of the final products of mixtures L-Glu—L-Lys or L-Glu—L-Arg is more diverse than that in the case of mixtures L-Asp—L-Lys or L-Asp—L-Arg (Schemes 4 and 5). The glutamic acid residue is transformed into the pyro-

Scheme 3



Reagents and conditions: L-Arg : L-Asp = 1 : 1, 2 equiv. H_2SO_4, 140 °C, 300 h (see Table 2, run *20*, sample **20B**).

Scheme 4

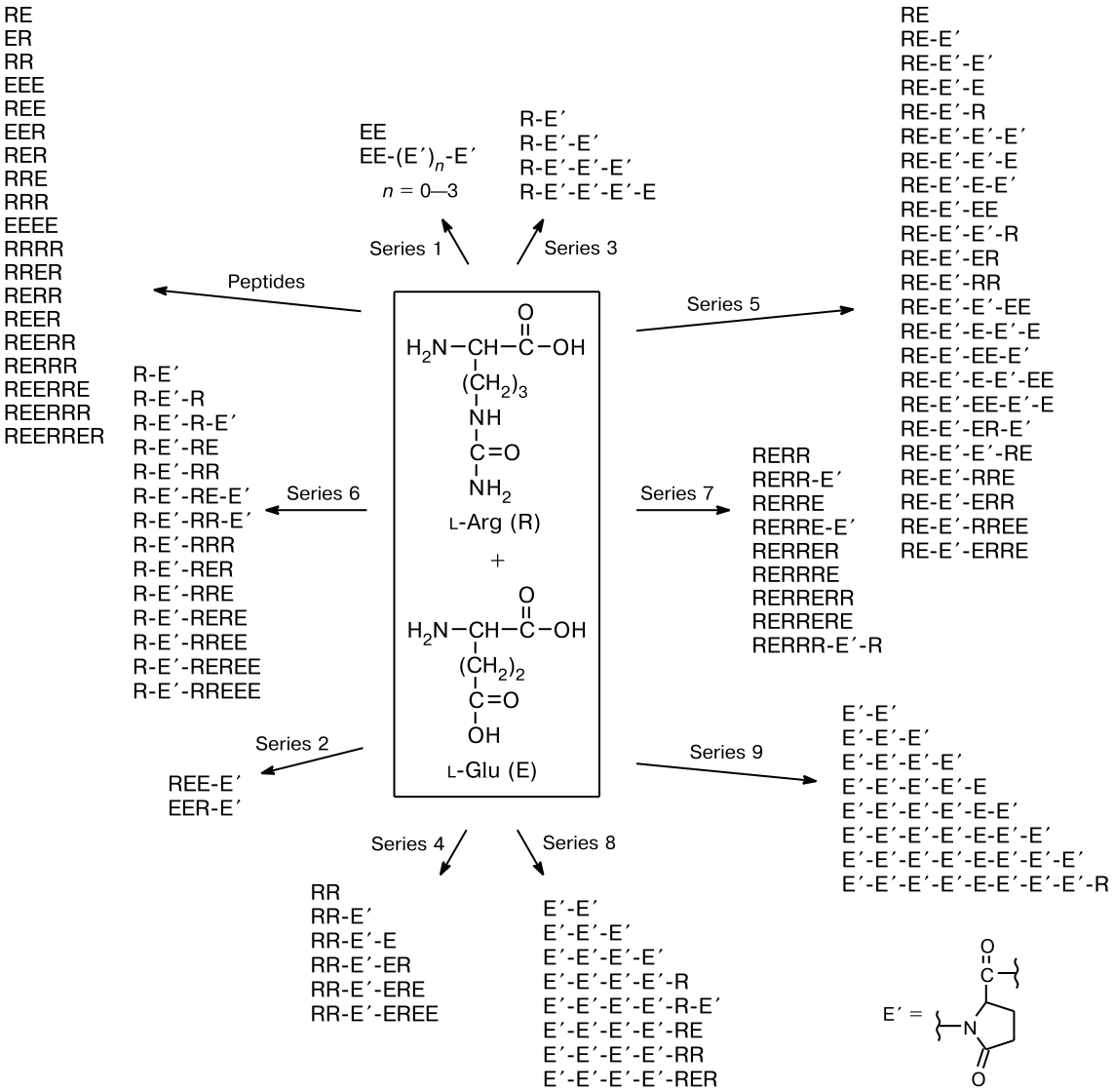


Reagents and conditions: L-Lys : L-Glu = 1 : 1, 2 equiv. H_2SO_4, 140 °C, 300 h (see Table 2, run *23*, sample **23B**).

glutamic acid residue to form peptide derivatives under the heating.

The data for the exact structures of two series of biopolymers for sample **20B** are presented in Table 4 (L-Asp : L-Arg = = 1 : 1, see Table 2, series III, run *20*). Both single- and double-charged ions correspond to these structures in the mass spectra. It must be noted that all series of derivatives shown in Scheme 3 are characteristic for the samples **19B**—**21B** (runs *19*—*21*, probe 2) and different by the ratio of amino acids only (see Table 2), and the ratio of amino acids determined the number of amino acid residues in the obtained products. For the ratio Asp : Arg = 2 : 1 in series 1, the maximum number of succinimide residues (n_{max}) is 18 and trace amounts of polysuccinimides with n_{max} equal to 19 and 20 were also observed, whereas n_{max} = 11 for Asp : Arg = 1 : 1. In series 2 for the ratio Asp : Arg = 2 : 1, the number of succinimide residues n_{max} is 19 (n_{max} = 15 for Asp : Arg = 1 : 1), in series 3 n = 21 (n_{max} = 17 for Asp : Arg = 1 : 1), and in series 4 and 5 n_{max} = 8 and 9 (n_{max} = 4 and 9 for Asp : Arg = 1 : 1), respectively. For the ratio Asp : Arg = 1 : 2, the number of succinimide residues n_{max} is 9 in series 1, n_{max} = 13 in series 2, n_{max} = 14 in

Scheme 5



Reagents and conditions: L-Arg : L-Glu = 1 : 1, 2 equiv. H_2SO_4, 140 °C, 300 h (see Table 2, run *26*, sample **26B**).

series 3, and $n_{max} = 2$ and 8 in series 4 and 5, respectively. Thus, the increase of the initial amount of Asp in sample **19B** results in the formation of peptide derivatives with a higher weight in series 1—4 compared to samples **20B** and **21B**, and a decrease of the initial amount of Asp in sample **21B** reduces the amount and weight of polymers in all series of the obtained reaction products. Both single- and double-charged ions (series 1—4) and also four-charged ions (series 5) corresponded to these products in the mass spectra. The formation of diketopiperazines was characteristic of all mixtures of series III studied by us.

The general scheme of peptide fragmentation when the collision-induced dissociation (CID) method is used and the mass spectra confirmed the structures of biopolymers and peptides in samples **20B**, **23B**, and **26B** are shown in Fig. 5 (see Table 2, runs *20*, *23*, and *26*). The mass spectra of CID fragmentation products of molecules AspArgddddddArg, Glu—Lys—Glu—Lys, and Arg—Glu—Glu—Arg were recorded in a linear ion trap (see Fig. 5).

Thus, we established that both peptides and their derivatives containing the succinimide or pyroglutamic fragments can be obtained from a mixture of two oppositely charged amino acids by means of the isothermal peptide synthesis. The composition of products depends on the amount of sulfuric acid used as a catalyst, the properties of the initial amino acids, and their ratio. It was mentioned

Table 4. Compositions of several series of the products for the sample **20B** (Arg—Asp mixture) according to the HPLC-MS/MS data

Number of residues in peptide or its derivative	Formula*	Molecular ion	
		$[M + H]^+$, m/z ($z = +1$)	$[M + 2H]^{2+}$, m/z ($z = +2$)
	Series1		
2	Rd	272.1365	—
3	Rdd	369.1529	—
4	Rddd	466.1692	—
5	Rdddd	563.1856	—
6	Rddddd	660.202	—
7	Rdddddd	757.2184	—
8	Rddddddd	854.2347	—
9	Rdddddddd	951.2511	—
10	Rd_9	1048.267	—
11	Rd_{10}	1145.284	—
12	Rd_{11}	1242.3	—
13	Rd_{12}	1339.317	—
	Series 2		
3	RdR	428.2376	—
4	RddR	525.254	—
5	RdddR	622.2703	—
6	RddddR	719.2867	360.1474
7	RdddddR	816.3031	408.6555
8	RddddddR	913.3195	457.1637
9	Rd_7R	1010.336	505.6719
10	Rd_8R	—	554.1801
11	Rd_9R	—	602.6883
12	$Rd_{10}R$	—	651.1965
13	$Rd_{11}R$	—	699.7047
14	$Rd_{12}R$	—	748.2129
15	$Rd_{13}R$	—	796.721
16	$Rd_{14}R$	—	845.2292
17	$Rd_{15}R$	—	893.7374
18	$Rd_{16}R$	—	942.2456

* d is the succinimide residue.

that the number of polysuccinimides and peptide derivatives containing the pyroglutamic acid fragments increases for the excess of amino acid with the additional carboxyl group (Asp, Glu). It was shown that a decrease in the polymerization rate of glutamic acid compared to aspartic acid results in the greater variety of the mixture of obtained products. It was found that the presence of diketopiperazines and amines in the final mixtures of products is characteristic of all mixtures of amino acids used by us.

Synthesis of peptides under the thermal cycling mode. Complementarity. The investigation of the abiogenic thermal synthesis in model systems under the cyclic mode was done by us. The process includes the following stages:

(1) the preparation of aqueous solutions of the initial mixtures of amino acids (L-Lys—L-Asp and L-Arg—L-Asp) with pH 4.5 and 8.5;

(2) the sampling of an aliquot of the solution into a glass beaker with the heating in a thermostat to 120—130 °C, the evaporation of water, and the formation of the solid phase;

(3) the heating of the solid phase during 12—24 h and the polycondensation in the solid phase;

(4) the cooling to ambient temperature, the addition of an aliquot of the initial solution of amino acids.

Then this cycle is multiply repeated to the complete consumption of the initial mixture of amino acids.

The polycondensation of amino acids occurred with the formation of a series of polypeptides from dipeptides to pentapeptides in the solid-phase polymerization mode at relatively low temperatures and short incubation times. The by-products formed under the thermal synthesis conditions in the solid phase, such as diketopiperazines, succinimide fragments, amines, and others (see above), were practically absent in this case. The compositions of mixtures of peptides obtained under the thermal cycling conditions and the amino acid sequences of individual peptides were determined by chromatography—mass spectrometric analysis (Figs 6—8). The determination of the amino acids primary sequence for tetrapeptide NH_2—Lys—Asp—Lys—Lys—COOH (KDKK) by tandem mass spectrometry in combination with HPLC using CID fragmentation is shown in Fig. 7.

The reproducibility of the obtained results is of note: several identical experiments were carried out for mixtures L-Asp—L-Lys and L-Asp—L-Arg with a time difference of 1 year. A comparison of the compositions of obtained samples presented in Table 5 shows that the number of peptides does not almost increase in time. The reproducibility of the results was confirmed by the comparison of the sample **3A** (see Table 1, run *3*) with the sample **28** (see Table 5). Sample **28** was obtained and analyzed 1 year earlier than sample **3A**. The composition of the both samples is the same; *i.e.*, the final mixtures of products with the same composition were obtained experimentally with the difference of 1 year (see Table 5).

It must be noted that 4 dipeptides, 8 tripeptides, 16 tetrapeptides, and 32 pentapeptides, *i.e.*, 60 variants of

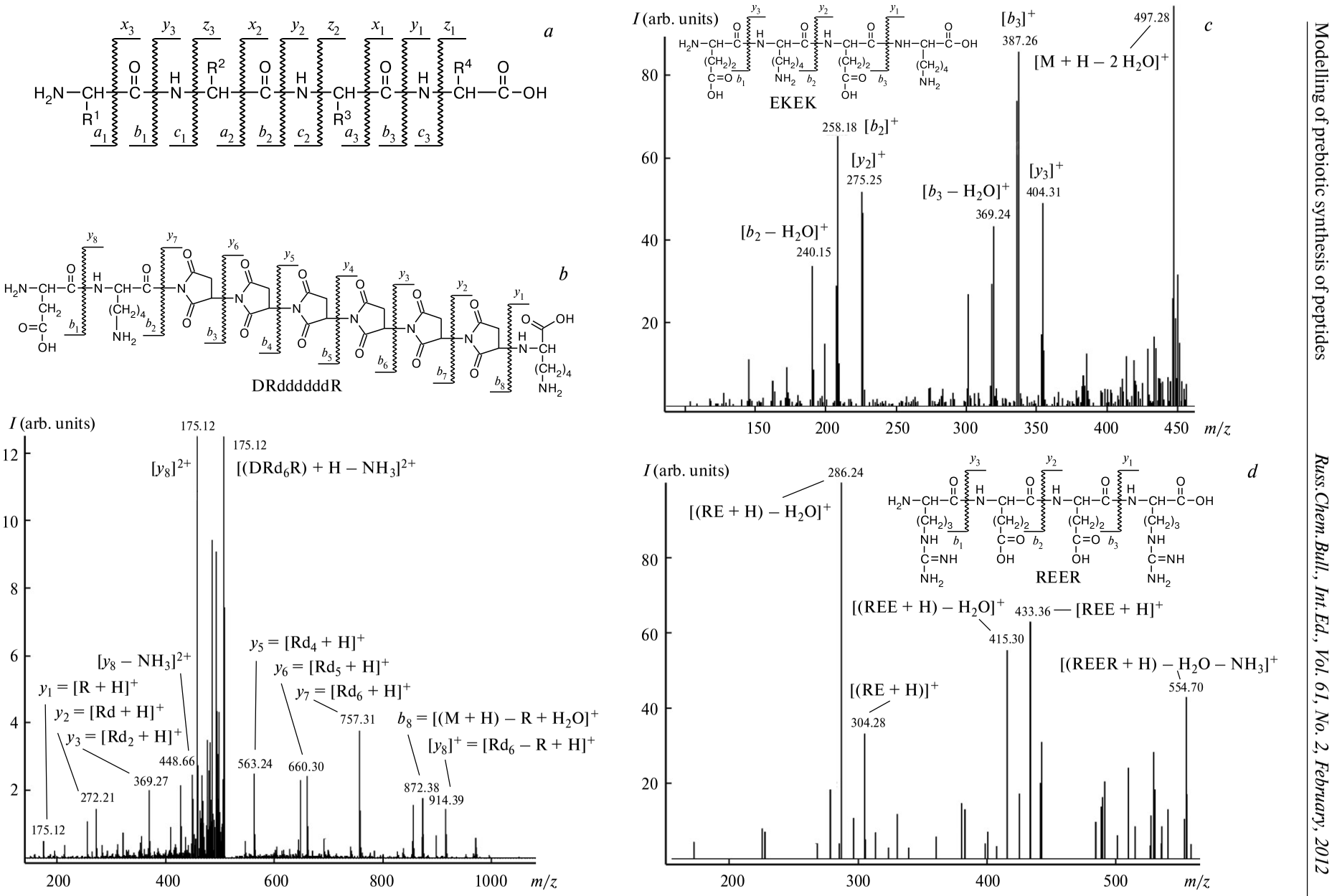


Fig. 5. The determination of the amino acid sequence of the thermal synthesis products in samples of mixtures Arg—Asp (sample **20B**), Lys—Glu (sample **23B**), and Arg—Glu (sample **26B**) by HPLC-MS/MS: *a*, the scheme of bond cleavage in the molecule; *b*, the mass spectrum of the CID fragmentation of DRd_6R (sample **20B**); *c* and *d*, the mass spectra of the CID fragmentation of EKEK (sample **23B**) (*c*) and REER (sample **26B**) (*d*). The mass spectra of the CID fragmentation products were detected in the linear ion trap.

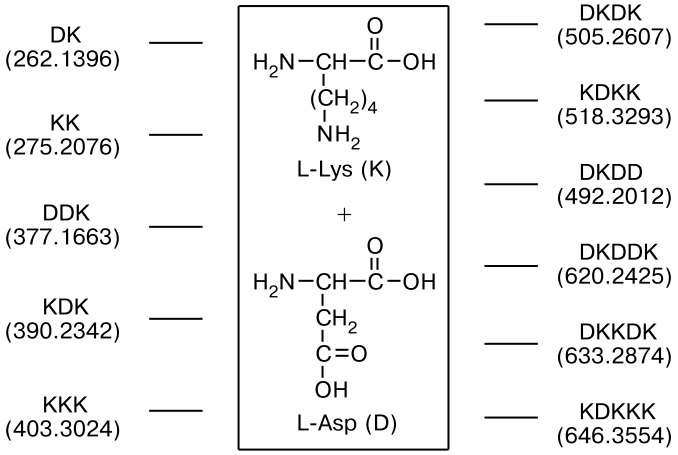


Fig. 6. The composition of a peptide mixture obtained from an L-Lys—L-Asp (1 : 1) mixture in the thermal cycling mode (see Table 1, run *6*, sample **6C**). The mass spectrum of the CID fragmentation products was detected in the linear ion trap. The reaction conditions: the heating temperature is 120 °C, the heating time is 144 h, 9 cycles, and pH 8.5. The experimentally determined masses of peptides are given in parentheses.

Table 5. Composition of products for samples **6A** and **6C** (L-Asp—L-Lys), **3A** and **3C** (L-Arg—L-Asp), and **28** (L-Arg—L-Asp) obtained in the thermal cycling mode at different numbers of cycles and heating times (the peptide structure was determined by HPLC-MS/MS)

Number of residues in peptide	Primary amino acid sequence	Molecular ion	
		$[M_{calc} + H]^+$	$[M_{exp} + H]^+$
	Sample **6A**		
2	KK	275.2089	275.2076
2	DK	262.1409	262.1396
3	DDK	377.1678	377.1663
3	KDK	390.2358	390.2342
3	KKK*	403.3039	403.3024
4	DKDK*	505.2628	505.2607
4	KDKK*	518.3308	518.3293
4	DKDD	492.1948	492.2012
5	DKDDK*	620.2897	620.2425
5	DKKDK*	633.3577	633.2874
5	KDKKK*	646.4258	646.3554
	Sample **6C**		
2	KK	275.2089	275.2076
2	DK	262.1409	262.1396
3	DDK	377.1678	377.1663
3	KDK	390.2358	390.2342
3	KKK	403.3039	403.3024
4	DKDK	505.2628	505.2607
4	KDKK	518.3308	518.3293
4	DKDD	492.1948	492.2012
5	DKDDK	620.2897	620.2425
5	DKKDK	633.3577	633.2874
5	KDKKK	646.4258	646.3554

Number of residues in peptide	Primary amino acid sequence	Molecular ion	
		$[M_{calc} + H]^+$	$[M_{exp} + H]^+$
	Sample **3A**		
2	RR	331.2212	331.2201
2	DR	290.147	290.1459
3	DDR	405.174	405.1728
3	RDR	446.2481	446.2470
4	DRDR	561.2751	561.2892
4	DDDR	520.2009	520.2151
	Sample **3C**		
2	RR	331.2212	331.2201
2	DR	290.147	290.1459
3	DDR	405.174	405.1728
3	RDR	446.2481	446.2470
4	DRDR	561.2751	561.2892
4	DDDR	520.2009	520.2151
5	DDDDR*	635.2279	635.2422
	Sample **28****		
2	RR*	331.2212	331.2201
2	DR	290.147	290.1459
3	DDR*	405.174	405.1728
3	RDR*	446.2481	446.2470
4	DRDR*	561.2751	561.2892
4	DDDR*	520.2009	520.2151

* Trace amounts.

** L-Arg—L-Asp (1 : 1), pH 4.5, 120 °C, 52 h, 4 cycles.

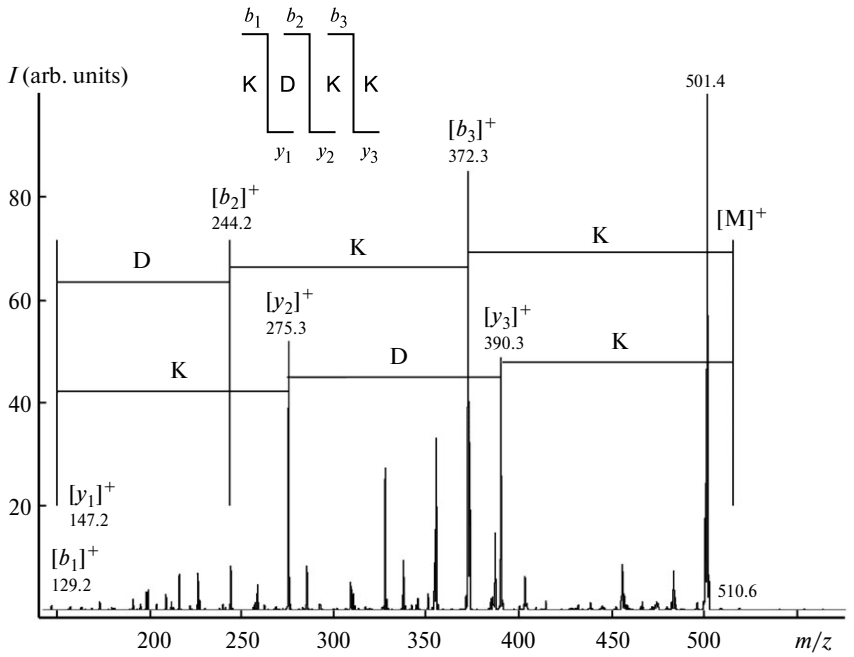


Fig. 7. The determination of the amino acid sequence of tetrapeptide KDKK synthesized from a L-Lys—L-Asp (1 : 1) mixture in the thermal cycling mode (see Table 1, run *6*, sample **6C**) by HPLC-MS/MS. The mass spectrum of the CID fragmentation products was detected in the linear ion trap. The reaction conditions: the heating temperature is 120 °C, the heating is time 144 h, 9 cycles, and pH 8.5.

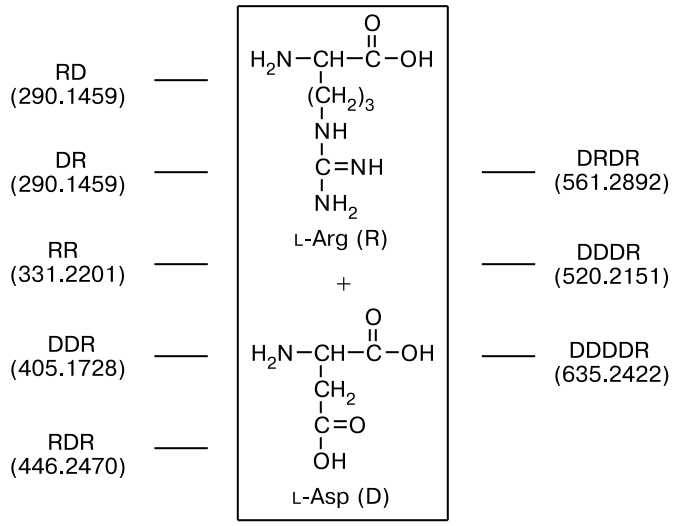
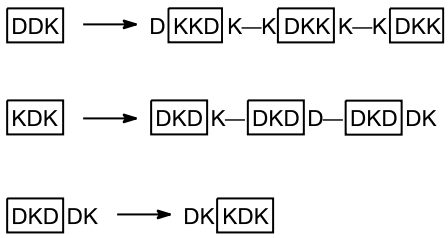


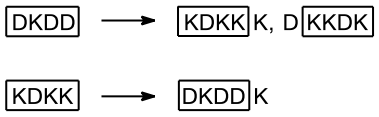
Fig. 8. The composition of the peptide mixture obtained from a L-Arg—L-Asp (1 : 1) mixture in the thermal cycling mode (see Table 1, run *3*, sample **3C**). The reaction conditions: the heating temperature is 120 °C, the heating time is 121.5 h, 7 cycles, and pH 4.5. The experimentally determined masses of peptides are given in parentheses.

structures, can be formed under the statistically equipropable synthesis modes from two amino acids. In the case of the L-Lys—L-Asp mixture, 11 peptides were identified, whereas 7 peptides were observed in the case of the L-Arg—L-Asp mixture. The products obtained from the L-Lys—L-Asp mixture contain no "independent" peptides, since 10 peptides of 11 have complementary the affinity to each other.

Complementary affinity for the "triples"



Complementary affinity for the "quadruples"



We assume that the above mentioned peptides can form complexes with each other due to the complementary salt bridges.

Thus, only peptides were prepared from the mixture of two oppositely charged amino acids in the peptide synthesis under the thermal cycling mode, and any polysuccinimides or peptide derivatives containing succinimide residues were not formed.

The experimental study of the peptide synthesis based on the polycondensation of trifunctional monomers under the thermal cycling and isothermal synthesis modes made it possible to draw several conclusions.

1. The processes of polycondensation with the peptide bonds formation in the studied mixtures of amino acids forming occur under relatively mild conditions at 120—140 °C and a normal pressure.

2. The use of FT ICR mass spectrometry and HPLC-MS/MS allowed us to identify almost all components of the complicated mixtures without a preliminary sample preparation.

3. The set of obtained products depends considerably on the conditions of the process, namely, when using sulfuric acid as a catalyst in the isothermal synthesis, aspartic acid undergoes a cyclization into the polymer chain to form succinimide fragments or polysuccinimide regions of the chain upon the heating time prolongation. The process of peptide bond formation under the action of strong acids, such as H_2SO_4, runs mainly on the α-carboxyl group with the appearance of cyclic imide after the formation of the "regular" peptide bond, which has been described for the conditions of classical peptide synthesis.[30] Under alkaline conditions, the proceses occur with the formation of normal peptides containing aspartic acid with the free γ-carboxyl group. The final set of synthesis products is more diverse when glutamic acid is used instead of aspartic acid.

4. No statistic inclusion of amino acids into polypeptide chains is observed for the peptide synthesis under both isothermal and thermal cycling modes. The number of different synthesized chains is much smaller than the statistically possible one. If to accept that the number of possible variants of the primary structures is defined by the proportion α^{β}, where α is the number of the code letters, and β is the chain length (some peptides are composed of the two-letter code, for example, K and D or K and d, while some others consist of the three-letter code, namely, K, D, d, where d is the succinimide), then the total number of potentially possible variants of structures for the samples presented in Table 3 is 3 472 980. In fact, only 58 structures were determined after the synthesis termination.

5. The evolution of the system takes place with the disappearance of some structures through the process prolongation even when the reaction is running in isothermal mode. For instance, after the incubation of the system for 104 h, the product KDddddddddK and the whole series of products DDKK, DDKKd, DDKKdd, DDKKddd, and DDKKdddd were unambiguously identified and disappeared after 278 h of incubation.

6. The synthesis of peptides with the complementary sequence of amino acids is observed under the thermal cycling mode. The synthesis of peptides through the thermal cycling mode showed the selection process with the formation of complementary chain, which agrees with the notions of the kinetic theory of the prebiotic evolution of macromolecules advanced by us.[23]

This work was partially supported by the Presidium of the Russian Academy of Sciences (Programs Nos 18 and 25).

References

1. E. M. Galimov, *Fenomen zhizni* [*Phenomenon of Life*], Editorial URSS, Moscow, 2001, 256 pp. (in Russian).
2. J. P. Ferris, Jr. A. R. Hill, R. Liu, L. E. Orgel, *Nature*, 1996, **381**, 59.
3. S. W. Fox, *J. Chem. Educ.*, 1957, **34**, 472.
4. S. W. Fox, *The Emergence of Life: Darwinian Evolution from the Inside*, Basic Books, New York, 1988, 224 pp.
5. C. Huber, G. Wächterhäuser, *Science*, 1998, **281**, 670.
6. S. A. Kauffman, *J. Theor. Biol.*, 1986, **119**, 1.
7. S. A. Kauffman, *The Origin of Order Self Organization and Selection in Evolution*, Oxford Univ. Press, Oxford, 1993, 184 p.
8. S. L. Miller, L. E. Orgel, *The Origins of Life on the Earth*, Prentice-Hall, Englewood Cliffs, New York, 1974, 229 pp.
9. A. I. Oparin, *The Chemical Origin of Life*, Charles C. Thomas Publ., Springfield, Illinois, 1964, 124 p.
10. L. E. Orgel, *The Origins of Life: Molecules and Natural Selection*, John Wiley, New York, 1973, 189 c.
11. L. E. Orgel, *Orig. Life Evol. Biosph.*, 1992, **28**, 227.
12. G. Wächtershäuser, *J. Theor. Biol.*, 1997, **187**, 483.
13. G. Wächtershäuser, *Microbiol. Rev.*, 1988, **52**, 452.
14. C. Chiarabelli, D. De Lucrezia, *Orig. Life Evol. Biosph.*, 2007, **37**, 357.
15. S. D. Varfolomeev, K. G. Gurevich, *Izv. Akad. Nauk, Ser. Khim.*, 2001, 1629 [*Russ. Chem. Bull., Int. Ed.*, 2001, **50**, 1709].

16. S. D. Varfolomeev, I. V. Uporov, I. A. Gariev, *Usp. Khim.*, 2005, **74**, 67 [*Russ. Chem. Rev.* (*Engl. Transl.*), 2005, **74**, 61].
17. M. Eigen, P. Shuster, *The Hypercycle. A Principle of Natural Self-Organization*, Springer-Verlag, Berlin—Heidelberg—New York, 1979, p. 92.
18. M. Eigen, *Self Organization of Matter and the Evolution of Biological Macromolecules*, Springer, Berlin, 1971, p. 108.
19. L. E. Orgel, *Nature*, 1992, **358**, 203.
20. L. E. Orgel, *Orig. Life Evol. Biosph.*, 1998, **28**, 91.
21. W. Szaflarski, K. H. Nierhaus, *Orig. Life Evol. Biosph.*, 2007, **37**, 423.
22. R. Popa, *Between Necessity and Probability: Searching for the Definition and Origin of Life*, in *Adv. Astrobiol. Biogeophys.*, Springer—Verlag, Berlin—Heidelberg, 2004, 253 pp.
23. S. D. Varfolomeev, *Mendeleev Commun.*, 2007, **17**, 7.
24. S. D. Varfolomeev, O. V. Demina, A. A. Khodonov, A. V. Laptev, E. N. Nikolaev, A. S. Kononikhin, in *Problemy zarozhdeniya i evolyutsii biosfery* [*Problems of Biosphere Origination and Evolution*], Ed. E. M. Galimov, Knizhnyi dom "LIBROKOM," Moscow, 2008, p. 57 (in Russian).
25. K. Matsubara, T. Nakato, M. Tomida, *Macromolecules*, 1998, **31**, 1466.
26. F. Chen, D. Yang, *Orig. Life Evol. Biosph.*, 2007, **37**, 47.
27. A. Commeyras, H. Collet, L. Boiteau, J. Taillades, O. Vandenabeele-Trambouze, H. Cottet, J.-P. Biron, R. Plasson, L. Mion, O. Lagrille, H. Martin, F. Selsis, M. Dobrijevic, *Polym. Int.*, 2002, **51**, 661.
28. K. Matsubara, T. Nakato, M. Tomida, *Macromolecules*, 1997, **30**, 2305.
29. T. Nakato, M. Yoshitake, K. Matsubara, M. Tomida, *Macromolecules*, 1998, **31**, 2107.
30. N. L. Benoiton, in *Chemistry of Peptide Synthesis*, Ed. N. L. Benoiton, Taylor and Francis Group LLC, 2006, p. 172.

Received June 22, 2011;
in revised form January 27, 2011