Orientation of the peptide formation of *N*-phosphoryl amino acids in solution

CHEN Zhongzhou[1], TONG Yufeng[2], CHEN Shuibing[1], LI Yanmei[1], CHEN Yi[1], ZHAO Yufen[1] & WANG Jinfeng[2]

1. Bioorganic Phosphorus Chemistry Laboratory, Department of Chemistry, School of Life Science and Engineering, Tsinghua University, Beijing 100084, China;
2. National Laboratory of Biomacromolecules, Center for Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Correspondence should be addressed to Li Yanmei.

Abstract **The peptide formation of *N*-phosphoryl amino acids with amino acids proceeds in aqueous solution without any coupling reagents. After being separated in sephadex gel column, the phosphoryl dipeptides were analyzed by the electrospray ionization tandem mass spectrometry (ESIMS/MS). The result demonstrates that phosphoryl dipeptides were detected in all the reaction systems. It is found that the formation of *N*-phosphoryl dipeptides is oriented: the N-terminal amino acid residues of the *N*-phosphoryl dipeptides are from *N*-phosphoryl amino acids, and the peptide elongation happened at the C-terminal. Only α-dipeptide, no β-dipeptide, is formed in the *N*-phosphoryl dipeptides, showing that α-carboxylic group is activated selectively by *N*-phosphorylation. Theoretical calculation shows that the peptide formation of *N*-phosphoryl amino acids might happen through a penta-coordinate carboxylic-phosphoric intermediate in solution. These results might give some clues to the study on the origin of proteins and protein biosynthesis.**

Keywords: ***N*-phosphoryl amino acids, orientation of peptide formation, solution, *N*-phosphoryl dipeptides.**

The biosynthesis and chemical synthesis of peptides and proteins are very important in many fields. Why did nature choose α-amino acids as the protein building block? Why is the peptide elongation direction in the biosynthesis of protein from N-terminal to C-terminal? Which role does phosphorus play in the biosynthesis of proteins? It was found in our laboratory that *N*-phosphoryl amino acids could be self-activated to form peptides without any catalyst, coupling reagent or energy carrier such as ATP[1—3]. Nucleotides could be formed when *N*-phos-phoryl amino acids react with nucleosides[4]. It was shown recently that the peptide formation has high selectivity[5]: peptides are found only in *N*-phosphoryl α-amino acids with coupling reagents, rather than *N*-phosphoryl β-amino acids. In this note, the reaction of *N*-phosphoryl polar amino acids with amino acids is carried out in aqueous solution, and the products are analyzed and identified by the electrospray ionization (ESI) mass spectrometry and NMR techniques. The orientation of the peptide formation is studied systematically. It might give a clue to the investigation of the role of phosphorus in the origin of life and in many biochemical processes[6].

1 Experimental

(ⅰ) Materials and apparatus. Authentic serylhistidine and histidylserine were purchased from Sigma Co. Ltd. *L*-serine, *L*-alanine, *L*-histidine, *L*-aspartic acid, other reagents and solvents were analytical grade. Serylalanine and alanylserine were synthesized by the DCC method[7].

The mass spectra were obtained using a Bruker ESQUIRE~LC^{TM} electrospray ionization ion trap spectrometer. The ^{31}P NMR spectra were acquired on a Bruker ACP 200 MHz spectrometer with 85% phosphoric acid as the external reference. The ^{1}H, ^{13}C NMR, COSY and HMBC spectra of DIPP-Asp-Ala-OEt were acquired on a Bruker Avance DMX600 spectrometer. Proton chemical shifts were referenced relative to that of the internal 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) at δ = 0.0. The ^{13}C chemical shifts were referenced indirectly using the consensus Ξ ratio[8].

(ⅱ) Synthesis of *N*-diisopropyl phosphoryl (DIPP) amino acids and dipeptides. DIPP and dipeptides were synthesized using diisopropyl phosphite according to ref. [9].

(ⅲ) Peptide formation of *N*-phosphoryl amino acids in solution. 100 mg *N*-phosphoryl amino acids and 100 mg amino acid were dissolved in aqueous solutions (pH 8.0). The mixture was incubated at 37℃ for 2 h. The reaction was monitored by ^{31}P NMR, and the yield of *N*-phosphoryl dipeptides was 5%—6% according to the ^{31}P NMR integration. The reaction mixture was concentrated, then separated on a 1 cm × 25 cm column of Sephadex gel LH20, using methanol as eluant. The purified products were analyzed by the ESI mass spectrometer in the positive-ion and negative-ion modes.

Peptide formation of DIPP-Asp with HCl • Ala-OEt in $CHCl_3$: 1.00 g of DIPP-Asp was dissolved in 16 mL $CHCl_3$, 0.40 g triethylamine and 0.57 g HCl • Ala-OEt were added. The mixture was incubated at 40℃ for 1 h. The reaction mixture was washed with ethyl acetate to remove the by-products. The solution was then acidified to pH=3.0 with dilute HCl and extracted with a mixed solvent of *tert*-butanol and ethyl acetate (volume ratio 2∶3). After being dried with $MgSO_4$, the solvent was evaporated to give the crude product. The phosphoryl dipeptide was isolated by gradient elution with the silica gel column. The eluting solvents consisted of chloroform and methanol. Total yield was 5%—6%.

2 Results and discussion

Serine, histidine and aspartic acid are the three conserved amino acid residues in the active sites of various

enzymes[10]. It is found that serylhistidine and related oligo-peptides could cleave DNA, RNA and proteins[11]. This is important for the study on the cleavage mechanism of nucleotides, the development for site-selective DNA cleavage reagents and gene cure. However, the traditional synthesis of serylhistidine was very complicated[12]. It was found in our laboratory that *N*-phosphoryl amino acids can form a peptide library in the presence of coupling reagents[5]. The peptide sequences and synthesis methods are very important in combination chemistry. Therefore, the peptide formation of *N*-phosphoryl amino acids might be useful for the synthesis of peptides and peptide libraries.

For the reaction of *N*-diisopropyl phosphoryl serine (DIPP-Ser) with histidine in aqueous solution, the products were analyzed by the positive-ion and negative-ion ESI mass spectra (fig. 1). The positive-ion ESI mass spectrum is very complex for the existence of cluster ions (fig. 1(a)), such as the cluster ions $(\text{DIPP-Ser})_2+\text{Na}^+$ (*m/z*: 561) and DIPP-Ser+His+H^+ (*m/z*: 425). On the contrary, the negative-ion mass spectrum (fig. 1(b)) is simple. In the positive-ion and negative-ion ESI mass spectra, two phosphoryl dipeptides are found. One is *N*-phosphoryl homodipeptide DIPP-Ser-Ser (M_w=356), formed by the self-assembly of DIPP-Ser. The other is *N*-phosphoryl heterodipeptide (M_w = 406), formed by the reaction of DIPP-Ser with histidine. The *N*-phosphoryl hetero-dipeptide is identified by ESIMS/MS mass spectra. The $[\text{His}+\text{H}]^+$ ion (*m/z*: 156, y_1 ion, the nomenclature is according to the reference[13], fig. 2(a) and table 1) and $[\text{DIPP-Ser}+\text{H}]^+$ ion (*m/z*: 270, $b_1+\text{H}_2\text{O}$[13], fig. 2(a) and table 1) are observed in the positive-ion ESIMS/MS mass spectrum. Moreover, the $[\text{His}-\text{H}]^-$ ion (*m/z*: 154, y_1 ion) is observed in the negative-ion ESIMS[3] mass spectra. Therefore, the *N*-phosphoryl hetero-dipeptide is identified as DIPP-Ser-His. For the reaction of *N*-diisopropyl phosphoryl his-

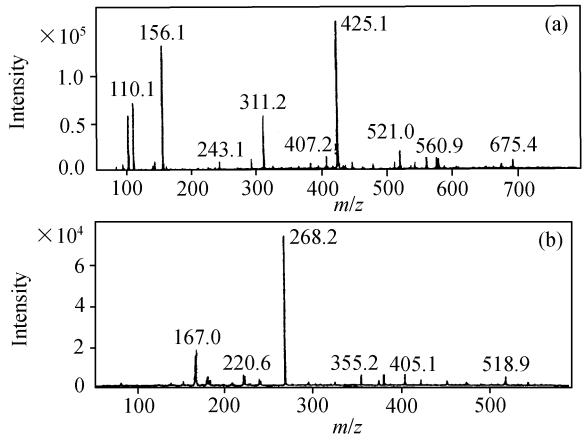

Fig. 1. (a) Positive-ion and (b) negative-ion ESIMS spectra of the reaction products of DIPP-*L*-Ser and *L*-His in water.

Table 1 The ESIMS/MS data of *N*-phosphodipeptides in the reactions of *N*-phosphoryl amino acid and amino acid in aqueous solution

Reaction	Compound	Precursor ion	Fragment ions and relative intensity percentage (% , in bracket)
DIPPSer + His	DIPP-Ser-His (M_w = 406)	$[\text{M+H}]^+$	389(84)[a], 365(100)[b], 347(60), 305(20), 280(19), 270(18), 243(20), 225(95), 208(18), 196(19), 180(19), 156(17)
		$[\text{M–H}]^-$	375(100)[c]
	DIPP-Ser-Ser (M_w = 356)	$[\text{M+Na}]^+$	361(50)[a], 337(42)[b], 319(8), 265(6), 295(9), 292(7), 277(9), 197(100), 163(22)
		$[\text{M–H}]^-$	325(100)[c], 295(31), 268(17), 265(6), 251(3), 238(6), 235(8), 181(7)[d], 173(7), 139(4)
DIPPHis + Ser	DIPP-His-Ser (M_w = 406)	$[\text{M+H}]^+$	389(100)[a], 365(8)[b], 347(12), 305(5), 274(18), 225(21)
		$[\text{M–H}]^-$	387(9)[a], 375(75)[c], 345(5), 318(100), 316(20), 303(60), 260(6)
	DIPP-His-His (M_w = 456)	$[\text{M+H}]^+$	439(35)[a],415(32)[b],411(27),373(6),320(27),293(7), 275(100), 274(12), 258(30), 247(7), 156(18)
		$[\text{M–H}]^-$	413(16)[b],395(100),357(10),353(30),335(10),333(7),309(9), 291(39), 273(8), 256(27), 230(13), 212(14), 176(18), 154(42)
DIPPSer + Ala	DIPP-Ser-Ala (M_w = 340)	$[\text{M+H}]^+$	323(34)[a], 305(14), 299(100)[b], 281(10), 257(21), 239(8),159(50)
		$[\text{M–H}]^-$	309(100)[c], 279(10), 249(5), 181(10)[d], 157(4), 139(3)
DIPPSer + Gly	DIPP-Ser-Gly (M_w = 326)	$[\text{M+H}]^+$	309(3)[a], 285(100)[b], 243(21), 145(39)
		$[\text{M–H}]^-$	295(100)[c], 265(35), 235(7), 181(9)[d], 139(2)
DIPPAla + Ser	DIPP-Ala-Ser (M_w = 340)	$[\text{M+Na}]^+$	345(100)[a], 321(53)[b], 303(5), 279(6), 276(6), 199(8), 181(47), 163(73), 138(8)
		$[\text{M–H}]^-$	321(30)[a], 309(100)[c], 297(4)[b], 279(17), 261(13), 252(40), 249(82), 192(7)
DIPPHis + Ala	DIPP-His-Ala (M_w = 390)	$[\text{M+H}]^+$	373(51)[a], 349(70)[b], 331(21), 307(21), 274(26), 227(44), 209(100), 110(22)
		$[\text{M–H}]^-$	329(37), 306(100), 302(5), 300(7), 272(4), 269(5), 251(8), 225(5), 93(19)
DIPPAsp + His	DIPP-Asp-His (M_w = 434)	$[\text{M+H}]^+$	417(100)[a], 393(30)[b], 375(21), 351(10), 274(26), 295(10), 253(35), 156(12)
		$[\text{M–H}]^-$	415(100)[a], 399(20), 372(8), 278(7), 198(10)

a) The loss of one water, b) the loss of one propylene, c) the loss of CH_2O, d) the characteristic fragmentation peak, corresponding to $(\text{Pr}^i\text{O})_2\text{P(O)O}^-$.

tidine (DIPP-His) with serine in aqueous solution, two phosphoryl dipeptides were detected too (table 1). One was DIPP-His-His (M_w = 456), the other was *N*-phosphoryl hetero-dipeptide (M_w = 406) from the reaction of DIPP-His with serine. The hetero-dipeptide is identified as DIPP-His-Ser by ESIMS/MS mass spectra (fig. 2(b), table 1), because the a_1 ion (*m/z*: 274, fig. 2(b)) and DIPP-His anion (*m/z*: 318, $b_1 + H_2O$[14], table 1) are observed in the positive-ion and negative-ion ESIMS/MS mass spectra, respectively. The characteristic fragmentation patterns of the two phosphoryl hetero-dipeptides, DIPP-Ser-His and DIPP-His-Ser are consistent with the authentic samples. For the reaction of DIPP-His with alanine in aqueous solution, two phosphoryl dipeptides are identified as DIPP-His-His (M_w = 456) and DIPP-His-Ala (M_w = 390) (table 1), because of the similar fragmentation patterns in the ESIMS/MS mass spectra (table 1), such as the presence of a_1 ion (*m/z*: 274)[13].

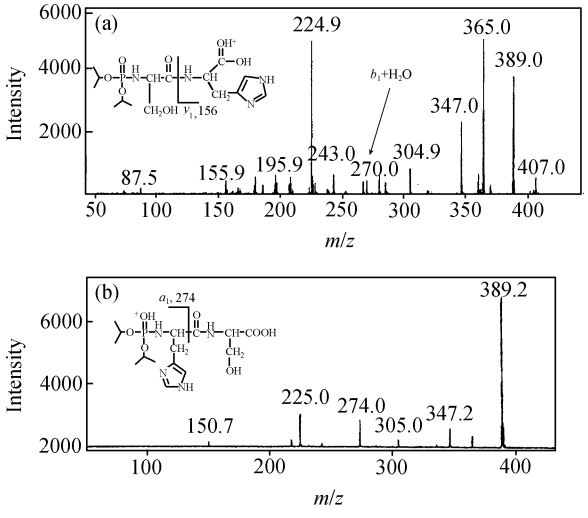

Fig. 2. Positive-ion ESIMS/MS spectra of (a) DIPP-Ser-His and (b) DIPP-His-Ser.

For the reaction of DIPP-Ser with alanine in aqueous solution, two phosphoryl dipeptides are found by ESI mass spectra. One is DIPP-Ser-Ser (M_w = 356), the other is identified as DIPP-Ser-Ala (M_w = 339), for the rearrangement ion $(Pr^iO)_2P(O)O^-$ (*m/z*: 181)[15] (fig. 3(a) and table 1) is observed in the negative-ion mass spectra. On the contrary, DIPP-Ala-Ala (M_w = 321) and DIPP-Ala- Ser (M_w = 339) are found in the reaction of DIPP-Ala with serine in aqueous solution. Because no rearrangement ion $(Pr^iO)_2P(O)O^-$ is found in the negative-ion mass spectra (fig. 3(b) and table 1), and the characteristic ions b_1+H_2O (*m/z*: 276 sodium adduct and *m/z* 252, table 1) are found in the positive-ion and negative-ion ESIMS/MS mass spectra, respectively. The fragmentation is consistent with the observation that the rearrangement ion $(Pr^iO)_2P(O)O^-$ is found only in *N*-phosphoryl amino acids or *N*-phosphoryl dipeptides with a free β-OH or CO_2H group on the side chain[15].

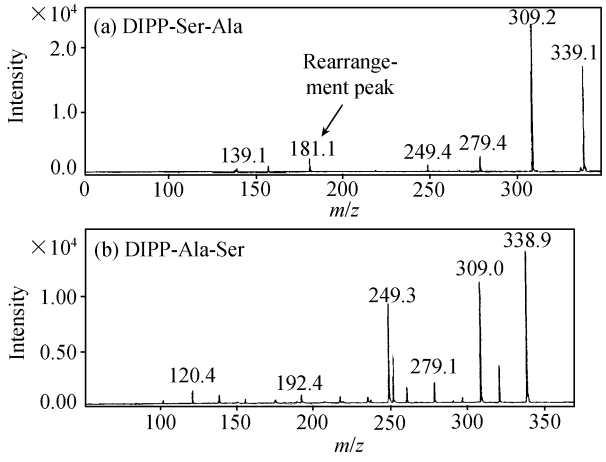

Fig. 3. Negative-ion ESIMS/MS spectra of (a) DIPP-Ser-Ala and (b) DIPP-Ala-Ser.

For the reaction of DIPP-Asp with histidine in aqueous solution, *N*-phosphoryl hetero-dipeptide was found too. The $[His+H]^+$ ion (*m/z*: 156, table 1) is observed in the positive-ion ESIMS/MS mass spectrum. So the *N*-phosphoryl hetero-dipeptide is identified as DIPP-Asp-His. However, phosphoryl aspartic acid had both α-COOH and β-COOH in one molecule. The phosphoryl hetero-dipeptide might be α-dipeptide, β-dipeptide, or the mixture. It is important to know which carboxylic group is involved in the peptide bond. To understand the structure of the dipeptide, we separated the corresponding hetero-dipeptide and identified it with two-dimensional NMR techniques.

To improve the reaction yield and separate easily, the reaction of DIPP-Asp with Ala-OEt was performed in chloroform. After being separated by silica column, the product DIPP-Asp-Ala-OEt (M_w = 396) was obtained. The ^1H NMR, ^{13}C NMR, COSY and HMBC spectra of the DIPP-Asp-Ala-OEt were acquired on a Bruker Avance DMX600 spectrometer[1)]. In the ^{13}C NMR and COSY spectra, it is found that the ^{13}C peak of the α-carbonyl group in Asp residue is split by the J-coupling from the phosphorus atom, with ^{13}C chemical shifts at 171.91 and 171.99 (fig. 4(a)). The ^{13}C peak of the β-carbonyl group is a singlet, with ^{13}C chemical shift at 175.1. Compared with phosphoryl aspartic acid, only ^{13}C chemical shift of the α-carbonyl group in aspartic residue decreases greatly. In the HMBC spectra, only the α carbonyl group of Asp residue has a cross peak with NH(Ala), while no cross peak is found for β carbonyl group of Asp residue with NH(Ala) (fig. 4(b)). This means that α-carbonyl group is

1) Chen, Z. Z., Tong, Y. F., Tan, B. et al., Comparing the activities of α-COOH and β-COOH in *N*-phosphoryl aspartic acids by NMR and theoretical study, J. Am. Chem. Soc., submitted.

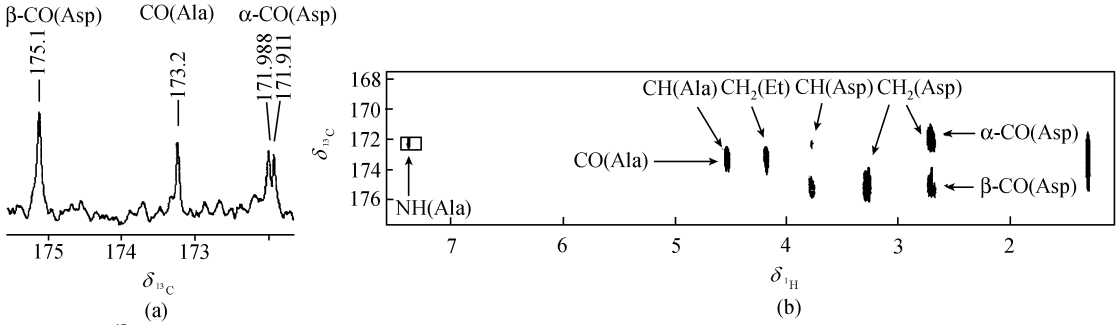

Fig. 4. (a) 1D-^{13}C spectrum of the carbonyl groups in DIPP-Asp-Ala-OEt, (b) 2D-HMBC spectrum of the carbonyl group chemical shift region in DIPP-Asp-Ala-OEt, the cross peak of α-CO(Asp) and NH(Ala) is boxed.

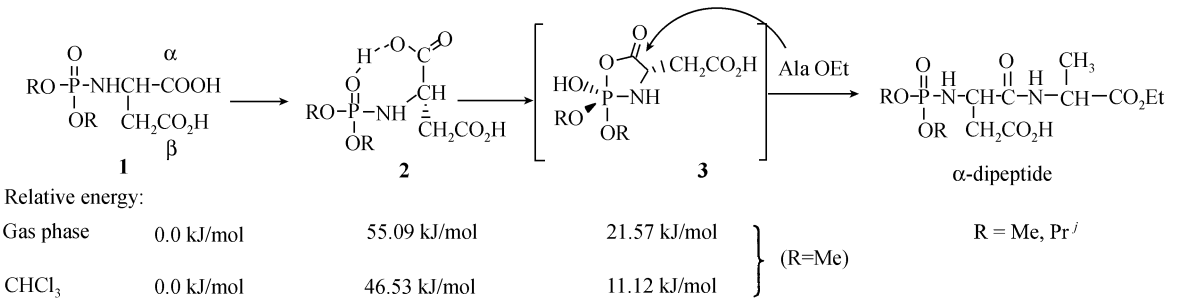

Fig. 5. The reaction mechanism and the relative energies of the intermediates and transition states in the peptide formation process from *N*-phosphoryl amino acids. Zero-point vibrational energies are considered in the energies at the b3LYP/6-31g** level.

linked to NH(Ala). Therefore, the phosphoryl aspartyl dipeptide ester must be DIPP-Asp-(α)-Ala-OEt, rather than DIPP-Asp-(β)-Ala-OEt. This result is consistent with the fact that ester exchange on phosphorus atom and ester formation are found only in phosphoryl amino acids with a free α-COOH group[16].

As shown in fig. 5, it is proposed that *N*-phosphoryl amino acid can form a penta-coordinate mixed anhydride intermediate (**3**), and *N*-phosphoryl dipeptide is formed by the attack of a free amino group to the α-carbonyl group in the penta-coordinate intermediate (**3**)[17–19]. To save the CPU time, the isopropyl groups in the phosphoryl group were simplified to methyl groups. By theoretical calculations, a hydrogen-bridge bond was found in the transition state (**2**)[20]. Compared with *N*-dimethyl phosphoryl aspartic acid (**1**) at the b3lyp/6-31G** level, the relative energies of the transition state (**2**) and intermediate (**3**) are 55.09 kJ/mol and 21.57 kJ/mol, respectively. If the solvent effects are considered, their relative energies in chloroform will be lower, with the energies as 46.53 kJ/mol and 11.12 kJ/mol (fig. 5), respectively. Because the relative energies of the transition state (**2**) and intermediate (**3**) are low, the formation of the peptide will proceed under mild conditions (room temperature) without any coupling reagents.

3 Conclusions

From analysis of the positive-ion and negative-ion ESI mass spectra, the peptide formation of *N*-phosphoryl polar amino acids with amino acids can proceed without any coupling reagents in aqueous solution. The formation of *N*-phosphoryl dipeptides is orientated: the *N*-terminal amino acid residues of *N*-phosphoryl dipeptides are from *N*-phosphoryl amino acids, and the peptide elongation happens at the C-terminal. The results are consistent with that from other *N*-phosphoryl amino acids[21]. For the peptide formation of *N*-phosphoryl aspartic acid with both α-COOH and β-COOH, only phosphoryl α-dipeptide, no β-dipeptide, was found in the *N*-phosphoryl dipeptides. These results might be helpful to the study on the origin of proteins, protein biosynthesis and the question as to why nature chose α-amino acids as protein backbone.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 20072023 and 20175026) and the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institute of Ministry of Education of China.

References

1. Li, Y. M., Yin, Y. W., Zhao, Y. F., Phosphoryl group participation leads to peptide formation from *N*-phosphoryl amino acids, Int. J. Pept. Protein Res., 1992, 39: 375.
2. Ju, Y., Zhao, Y. F., Sha, Y. W. et al., Phosphoryl promotion and differentiation effect on amino acids and prebiotic synthesis of

protein, Phosphorus, Sulfur Silicon Relat. Elem., 1995, 101: 117.
3. Zhao, Y. F., Ju, Y., Li, Y. M. et al., Self-activation of *N*-phosphoamino acids and *N*-phosphodipeptides in oligopeptide formation, Int. J. Pept. Protein Res., 1995, 45: 514.
4. Zhou, W. H., Ju, Y., Zhao, Y. F. et al., Simultaneous formation of peptides and nucleotides from *N*–phosphothreonine, Orig. Life Evol. Biosph., 1996, 26: 547.
5. Fu, H., Li, Z. L., Zhao, Y. F. et al., Oligomerization of N,O-bis (trimethylsilyl)-alpha-amino acids into peptides mediated by *o*-phenylene phosphorochloridate, J. Am. Chem. Soc., 1999, 121: 291.
6. Zhao, Y. F., Cao, P. S., Phosphoryl amino-acids——common origin for nucleic-acids and protein, J. Biol. Phys., 1994, 20: 283.
7. Bodanszky, M., Bodanszky, A., Protecting Groups and Activation and Coupling: in The Practice of Peptide Synthesis, Berlin: Springer-Verlag, 1984: 18—110.
8. Markley, J. L., Bax, A., Arata, Y. et al., Recommendations for the presentation of NMR structures of proteins and nucleic acids, Pure Appl. Chem., 1998, 70: 117.
9. Ji, G. J., Xue, C. B., Zeng, J. N. et al., Synthesis of *N*-O,O- diisopropyl phosphoryl amino acids and dipeptides, Synthesis, 1988: 444.
10. Nardini, M., Dijkstra, B. W., α/β hydrolase fold enzymes: the family keeps growing, Curr. Opin. Struc. Biol., 1999, 9: 732.
11. Li, Y. S., Zhao, Y. F., Hatfield, S. et al., Dipeptide seryl-histidine and related oligopeptides cleave DNA, protein, and a carboxyl ester, Bioorg. Med. Chem., 2000, 8: 2675.
12. Okawa, K., Studies on serine peptides, Ⅲ. Synthesis of *L*-seryl-*L*-histidine, *L*-seryl-*L*-histidine-L-leucine and carbobenzyloxy-glycyl-o-benzyle-*DL*-serine ethyl ester, Bull. Chem. Soc. Jpn., 1957, 30: 976.
13. Johnson, R. S., Martin, S. A., Biemann, K. et al., Fragmentation of protonated peptides: surface-induced dissociation in conjunction with a quantum mechanical approach, Anal. Chem., 1987, 59: 2621.
14. Sadagopan, N., Watson, J. T., Mass spectrometric evidence for mechanisms of fragmentation of charge-derivatized peptides, J. Am. Soc. Mass Spectrom., 2001, 12: 399.
15. Chen, Z. Z., Chen, S. B., Chen, Y. et al., Negative-ion electrospray ionization tandem mass spectrometry of *N*-phosphoryl amino acids and dipeptides, Rapid Commun. Mass Spectrom., 2002, 16 (8): 790.
16. Li, Y. M., Zhang, D. Q., Zhang, H. W. et al., β-Carboxyl catalytic effect of *N*-phosphoryl aspartic acid, Bioorg. Chem., 1992, 20: 285.
17. Chen, Z. Z., Tan, B., Li, Y. M. et al., Differentiation of alpha-COOH from beta-COOH in aspartic acids by *N*-phosphorylation, I. J. Quan. Chem., 2001, 83: 41.
18. Zhong, R. G., Zhao, Y. F., Dai, Q. H., Mechanism of intramolecular mixed anhydride formation as intermediates in the reactions of *N*-phosphorylamino-acids——An MNDO research, Chinese Sci. Bull., 1998, 43: 917.
19. Chen, Z. Z., Tan, B., Li, Y. M. et al., Differentiation of alpha-COOH from gamma-COOH in glutamic acid by *N*-phosphorylation, J. Mol. Struc-Theo. Chem., 2001, 574: 163.
20. Chen, Z. Z., Tan, B., Li, Y. M. et al., Theoretical study on the activity of α-COOH and β-COOH of *N*-phosphoryl aspartic acids, Chinese Chem. Lett., 2001, 12: 1093.
21. Jiang, Y., Fu, H., Xu, L. et al., Identification of self-assembly products from *N*-phosphoamino acids by electrospray ionization mass spectrometry, Rapid Commun. Mass Spectrom., 2000, 14: 1491.

(Received May 13, 2002)