CHEMISTRY LETTERS, pp. 1885-1888, 1993.

Solution Structure of a Peptide Designed on Plastocyanin Metal Center in Trifluoroethanol and Incorporation of Cobalt(II) into the Peptide

Hiroaki ZAIMA,* Norikazu UEYAMA, Akira NAKAMURA, and Saburo AIMOTO†
Department of Macromolecular Science, Faculty of Science, Osaka University, Toyonaka, Osaka 560
†Institute for Protein Research, Osaka University, Suita, Osaka 565

A novel sixteen amino acid peptide which has a portion of the conserved amino acid sequence of potato plastocyanin was synthesized. Its solution structure in trifluoroethanol (TFE) was determined by ^{1}H NMR spectroscopy. Incorporation of cobalt(II) ion into the peptide having a turn structure was examined. Cys6, His9, and Met14 of the peptide in TFE are located close to each other and suitable to bind a metal ion. Incorporation of Co(II) ion to the peptide results in formation of tetrahedral coordination site.

A major challenge in protein design is to create stable scaffolds into which tailored functions can be introduced. Here we report the design, a solution structure in TFE of a novel sixteen amino acid peptide (**16P1**) which has a portion of a conserved amino acid sequence of potato plastocyanin metal center, and a specific incorporation of cobalt(II) ion into the **16P1**. The novel peptide (**16P1**) consists of an artificial sequence part (1-4 residues) which has one histidine as a ligand and a conserved sequence part (5-16 residues) of potato plastocyanin.

Ac-Ile-His-Ala-Gly-Phe-Cys-Ala-Pro-His-Leu-Gly-Ala-Gly-Met-Val-Ala-NH_2 (**16P1**)

Plastocyanin has a distorted tetrahedral geometry of Cu(II), being coordinated by the N(imidazole) atoms of two histidine residues(His37, His87), the S(thiolate) atom of a cysteine residue(Cys84) and the S(thioether) atom of a methionine residue(Met92). These residues and other residues are conserved in other plastocyanins. We chose the conserved sequence of plastocyanin metal center as the part of **16P1** because: (a) plastocyanin has been well studied about the structure and the function [1]; (b) the conserved sequence, which should be important for its function, has three metal binding residues (Cys84, His87, and Met92); and (c) the conserved sequence containing **16P1** and its metal complex can be studied structurally and functionally compared to the plastocyanin. To create a tetrahedral metal coordination site in **16P1**, the artificial sequence part was designed by computer simulation using a BIOGRAF program. Co(II) was adopted for the incorporation study because using of Cu(II) results in a decomposition of its complex at room temperature. Fortunately, many studies of Co(II)-substituted type 1 copper proteins have been reported.[2-4]

This peptide was synthesized by a conventional solid phase method and purified with HPLC. **16P1** is soluble in N,N'-dimethylformamide (DMF) and TFE and insoluble to water. The CD spectrum of **16P1** in TFE

indicates that **16P1** has one bent or turn structure in TFE.[5] To elucidate the solution structure, H-H COSY and NOESY spectra were measured in 10% H_2O-TFE-d_4 and a distance geometry calculation was performed with a MolSkop (JEOL) program using the NOESY data.

Fifty NOE cross-peaks were assigned in the NOESY spectra of **16P1**. The NOESY results gave 30 independent structures by starting from either β-strand or random chain. A superposition of 5-structures is shown in Fig.1-(a). The structures are flexible at the first 5 N-terminal residues and from residue 14-16. Figure 1-(b) shows a minimum R.M.S.D structure as a tube model. Structural features of this calculated model are as follows; **16P1** has one turn structure at Cys-Ala-Pro-His position as shown in CD spectra. Cys6, His9, and Met14 are close to each other and His2 is away from these residues. Especially, it is interesting that the solution structure of the conserved sequence part in TFE is similar to the crystallographically determined structure of the native poplar plastocyanin.[6]

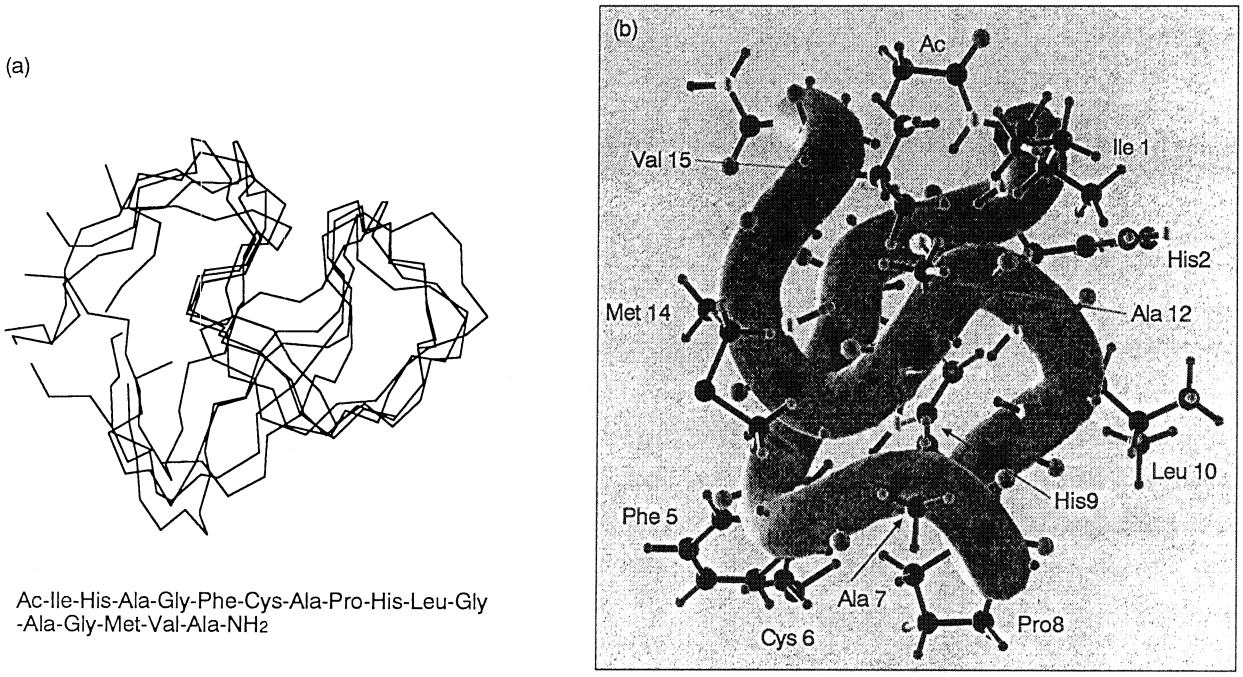

Fig.1. Superposition of the backbones of 5 structures (a) and Averaged solution structure of **16P1** illustrared by tube model which are obtained from restrained distance geometry calculation (b).

Cysteine thiolate, histidine imidazole, and methionine thioether of **16P1** may coordinate to the same metal ion to form a chelate of which the coordination structure is similar to the metal site of plastocyanin, because Cys6, His9, and Met14 are located close to each other in TFE. We tried to incorporate a Co(II) ion into **16P1** with the turn structure kept in TFE to obtain a complex of which the structure is similar to the metal site of plastocyanin. A TFE solution of $Co(ClO_4)_2 \cdot 6H_2O$ was used for the incorporation of Co(II) into **16P1**. The result of Co(II) titration monitored at 300 nm indicates that 0.6 equivalent of Co(II) reacts with **16P1** (formation constant 570 M^{-1}). This complex (**16P1-Co**) is unstable to dioxygen in solution. Figure 2 shows the CD spectra of **16P1** and **16P1-Co** in the region of 200-300 nm in TFE assignable to amide n-π^* transition. The CD data shows that the peptide chain of **16P1-Co** has one turn structure in TFE. Therefore, the major backbone structure of **16P1-Co**

in TFE solution is similar to that of **16P1**.

Figure 3 shows the UV-visible, CD, and MCD spectra of **16P1-Co** in the region of 300 - 700 nm in TFE. **16P1-Co** has a shoulder at 300 - 350 nm in the absorption spectra and the corresponding peaks in the CD and MCD spectra at 300-400 nm which is assigned to S→Co(II) charge transfer band on basis of the spectra of the reported Co(II) complexes[7] and Co-substituted metalloproteins.[8-10,11] The Co(II) titration experiment shows the formation of Co-S bond in 60% yield. Forty of percents **16P1** may have unfavorable solution structures for the formation of Co-S bond. Intensities and regions of *d-d* transition band of absorption spectra and MCD spectra show similarities to those of $Co(DVPM)_2$, $Co(OH)_4^{2-}$, $Co(hisH_{-1})_2^{2-}$,[12] cobalt carboxypeptitase A, cobalt thermolysin,[13] and cobalt angiotensin-inhibitor complexes,[14] whose metal-sites are known to have a distorted tetrahedral geometry. However, MCD maximum at 580 nm due to *d-d* transition of **16P1-Co** is different from those at 600 - 700 nm in Co-substituted blue copper proteins.[11]

^1H NMR spectrum measurement shows signals in the region of 100 - -50 ppm that are assignable to isotropically shifted protons of amino acid residues due to the high-spin cobalt(II) (Fig. 4). Ten broad signals having small T_1 values (<5 ms) at 93, 69, 63, 59, 54, 49, 42, -28, -33 and -45 ppm are assignable to the protons of the amino acid residues close to the cobalt(II) ion. The signal at 93 ppm having twofold intensity relative to the signals at 69, 63, -28 and -33 ppm. These signals are assignable to two set of protons of two histidine imidazole $C_\varepsilon H$, NH, and $C_\delta H$ (His2, His9), respectively.[4,15,16] The signals at 59, 54 and 49 are assignable to three protons of the histidine imidazole when only the imidazole groups coordinate to cobalt(II) ion. Two remained signals were tentatively assigned to Met $C_\gamma H$ and Met $C_\delta H$ protons. Cysteine thiolate coordinated cobalt(II) complexes show broad signals of cysteine $C_\beta H$ in the range 300 - 100 ppm.[4,15] Obvious signals which are assignable to cysteine $C_\beta H$ of **16P1-**

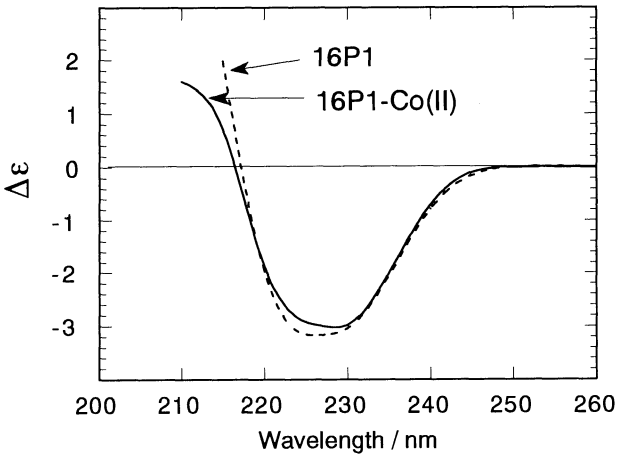

Fig.2. The high energy part of CD spectra of **16P1** and **16P1-Co** in TFE.

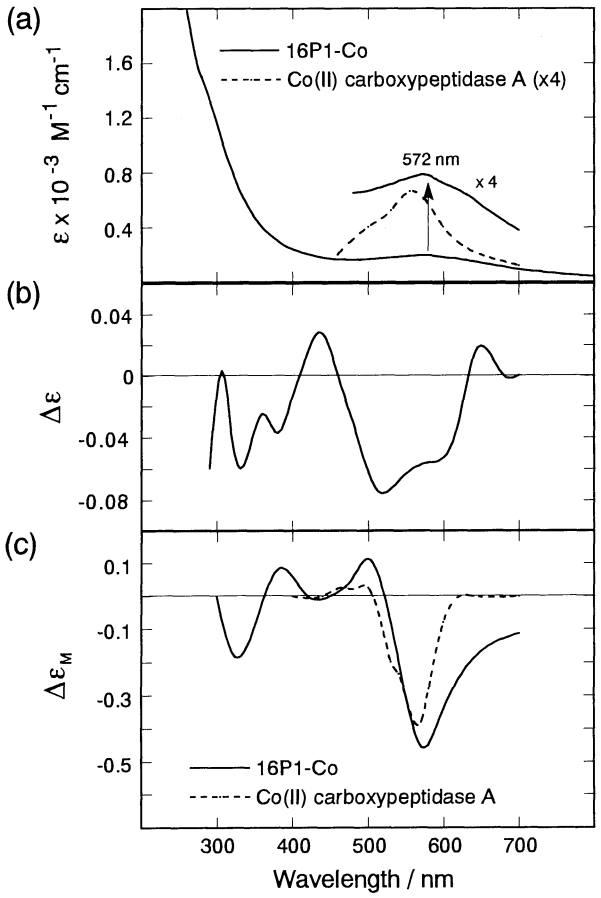

Fig.3. UV-visible(a), CD(b), and MCD(c) spectra of 16P1-Co in TFE(——) and Co(II) substituted carboxypeptidase A (------).

Co are not observed in the range 300 - 100 ppm in this spectrum measurement condition because of the extremely short T_1 values of cysteine CβH protons of **16P1-Co**.

The UV-visible, CD and MCD spectra reveal that **16P1-Co** has a tetrahedral coordination sphere and suggest that cysteine thiolate coordinate to the cobalt ion. ^{1}H NMR spectrum reveals that two histidine imidazoles coordinate to the cobalt ion. There are not enough data to decide the methionine as the fourth ligand.

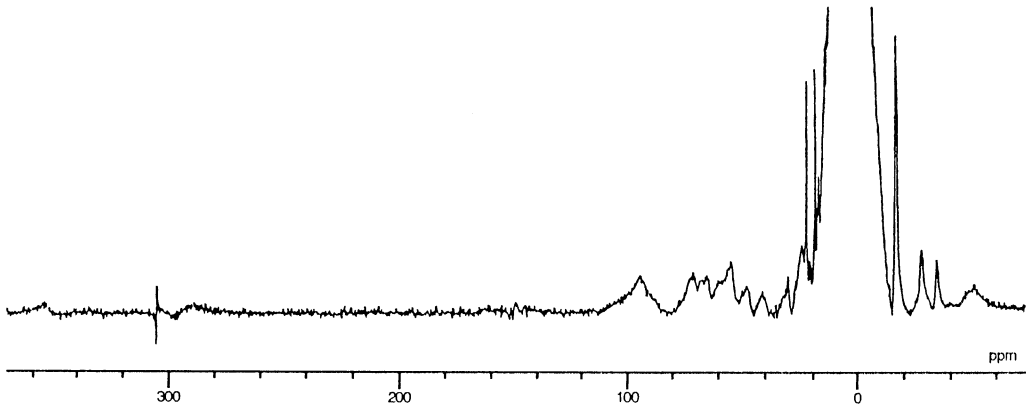

Fig.4. ^{1}H NMR (270 MHz) spectrum of **16P1-Co** in DMF-d_7.

The authors gratefully acknowledge Ken-ichi Fujita in JEOL Ltd. for the analysis of solution structure using a MolSkop program.

References

1) J.M.Guss, P.R.Harrowell, M.Murata, V.A.Norris, and H.C.Freeman, *J. Mol. Biol.*, **192**, 361 (1986).
2) E.I.Solomon, J.Rawlings, D.R.Mcmillin, P.J.Stephens, and H.B.Gray, *J. Am. Chem. Soc.*, **98**, 8046 (1976).
3) H.A.O.Hill, B.E.Smith, and C.B.Storm, *Biochem. Biophys. Res. Commun.*, **70**, 783 (1976).
4) S.Dahlin, B.Reinhammer, and J.Ångström, *Biochemistry* , **28**, 7224 (1989).
5) D.Greff, S.Fermandjian, P.Fromageot, M.C.Khosla, R.R.Smeby, and M.Bumpus, *Eur. J. Biochem.*, **61**, 297 (1976).
6) J.M.Guss and H.Freeman, *J. Mol. Biol.*, **169**, 521 (1983).
7) D.T.Corwin,Jr., R.Fikar, and S.A.Koch, *Inorg. Chem.*, **26**, 3079 (1987).
8) A.J.Sytkowski and B.L.Vallee, *Biochemistry* , **17**, 2850 (1978).
9) R.Bicknell, A.Schäffer, S.G.Waley, and D.S.Auld, *Biochemistry* , **25**, 7208 (1986).
10) D.W.Fitzgerald and J.E.Coleman, *Biochemistry* , **30**, 5195 (1991).
11) S.Suzuki, T.Sakurai, S.Shidara, and H.Iwasaki, *Inorg. Chem.*, **28**, 802 (1989).
12) T.A.Kaden, B.Holmquist, and B.L.Vallee, *Inorg. Chem.*, **13**, 2585 (1974).
13) B.Holmquist, T.A.Kaden, and B.L.Vallee, *Biochemistry* , **14**, 1454 (1975).
14) R.Bicknell, B.Holmquist, F.S.Lee, M.T.Martin, and J.F.Riordan, *Biochemistry* , **26**, 7291 (1987).
15) I.Bertini, M.Gerber, G.Lanini, C.Luchinat, W.Maret, S.Rawer, and M.Zeppezauer, *J. Am. Chem. Soc.*, **106**, 1826 (1984).
16) H.Adachi, Ph.D.Thesis, Osaka University, Osaka, Japan (1992).

(Received July 28, 1993)