



HOKKAIDO UNIVERSITY

Title	Structure and Thermodynamic Studies of a Ca ²⁺ -Binding Lysozyme : Biophysical Characterization of Canine Milk Lysozyme
Author(s)	Koshiba, Takumi
Degree Grantor	北海道大学
Degree Name	博士(理学)
Dissertation Number	甲第5381号
Issue Date	2001-03-23
DOI	https://doi.org/10.11501/3182075
Doc URL	https://hdl.handle.net/2115/51495
Type	doctoral thesis
File Information	000000399621.pdf

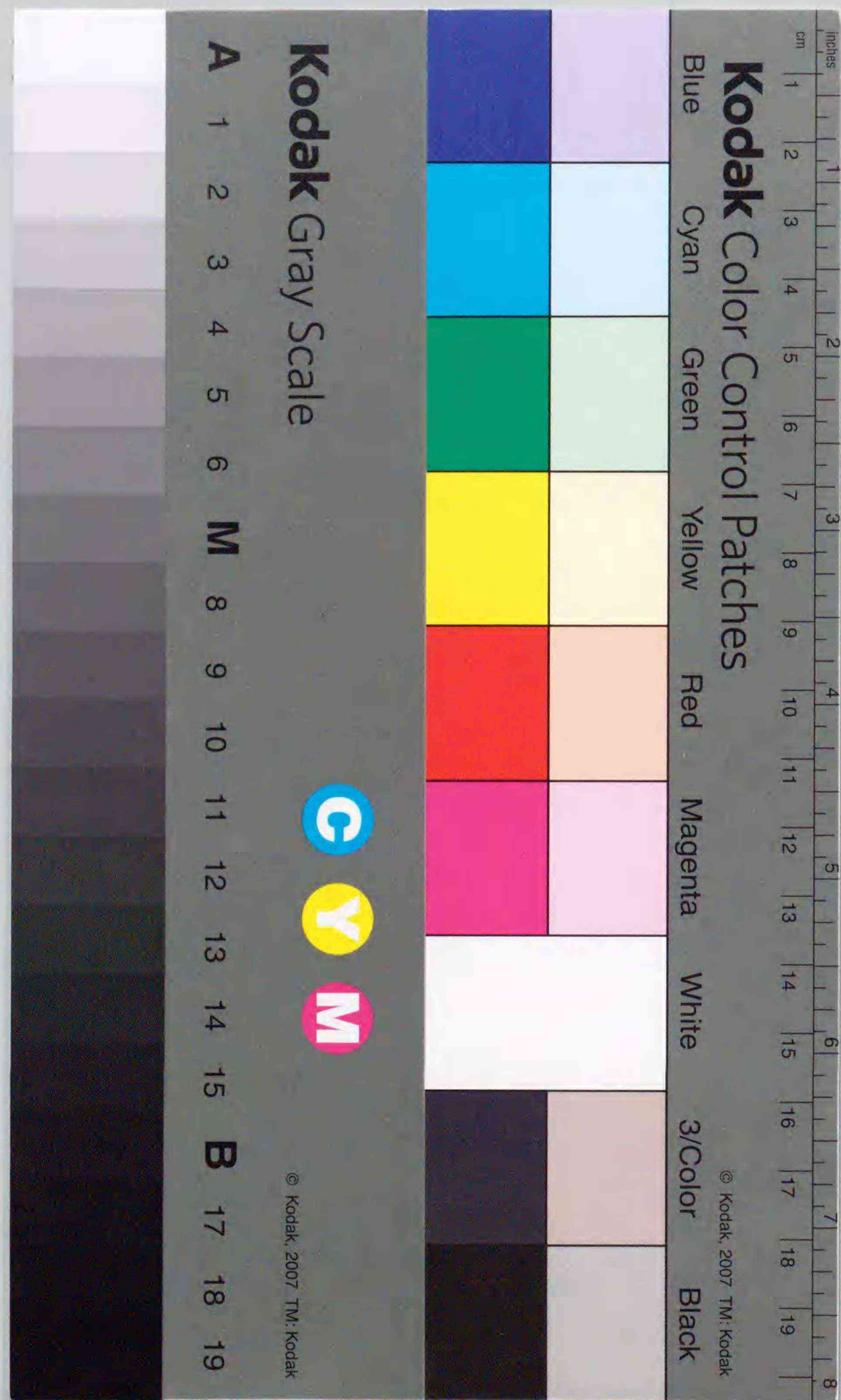


Structure and Thermodynamic Studies of a
 Ca^{2+} -Binding Lysozyme: Biophysical
Characterization of Canine Milk Lysozyme

Takumi KOSHIBA

*Division of Biological Sciences
Graduate School of Science
Hokkaido University*

The Year of 2001



①

Structure and Thermodynamic Studies of a
 Ca^{2+} -Binding Lysozyme: Biophysical
Characterization of Canine Milk Lysozyme

Takumi KOSHIBA

*Division of Biological Sciences
Graduate School of Science
Hokkaido University*

The Year of 2001

CONTENTS

GENERAL INTRODUCTION

3-6

CHAPTER I

**Expression of a Synthetic Gene Encoding Canine Milk
Lysozyme in *Escherichia coli* and Characterization of the
Expression Protein**

7-31

CHAPTER II

**Structure and Thermodynamics of the Extraordinarily Stable
Molten Globule State of Canine Milk Lysozyme**

32-65

CHAPTER III

**Energetics of Three-state Unfolding of a Protein: Canine Milk
Lysozyme**

66-96

CHAPTER IV

**Calcium-Induced Conformational Changes of Canine Milk
Lysozyme Revealed by Structural and Thermodynamic
Evidence**

97-121

ACKNOWLEDGEMENTS

122-123

GENERAL INTRODUCTION

Proteins are linear polymers assembled from some 20 different types of amino acids by polypeptide linkages. In general, they can be divided into two categories. One is fibrous protein, and it has a static function in providing one of the structural elements of living matter. The other is globular protein, and has a dynamic function that is highly related to their three-dimensional structure. The globular proteins exhibit a highly convoluted structure with one or more polypeptide chains within an overall form. Therefore, the mechanism by which an unstructured polypeptide chain folds to its final three-dimensional structure is one of the major problems in molecular structural biology (1).

Lysozyme- α -lactalbumin superfamily is one of the most investigated protein superfamilies with biophysical and biochemical methodologies (2-4). Lysozyme and α -lactalbumin are homologous to each other and have undoubtedly diverged from a common ancestral protein. Lysozyme is distributed widely in animal kingdom and especially occurs in some avian egg white. It catalyzes the hydrolysis of glycosidic bonds between *N*-acetylmuramic acid and *N*-acetylglucosamine residues of peptidoglycans in bacterial cell walls. In contrast, α -lactalbumin is found in the milk of lactating mammals, and its gene is expressed only in the mammary gland. The role of α -lactalbumin is to modify the enzymatic specificity of galactosyl transferase; it forms a complex with the enzyme and converts the complex into lactose synthase (5-7). Thus, the functions of these two proteins are quite different in spite of similarities in their amino acid sequences (8), the intron-exon distribution of their genes (9), and three-dimensional structures (10, 11). The calcium-binding properties of the two proteins are also known to be different. All the findings to date show that α -lactalbumin strongly binds a calcium ion (12, 13), whereas chicken type (c-type) lysozyme does not specifically bind calcium ions in most cases. The previous studies of our laboratory have also revealed, however, that there is a group of calcium-binding lysozymes, although only several such examples have been identified at present (14, 15). These

calcium-binding lysozymes are especially important in understanding the evolutionary relationship between lysozyme and α -lactalbumin, because they seem to provide an evolutionary link between the two proteins (15).

Canine milk lysozyme (CML) belongs to the family of calcium-binding lysozymes (16, 17), and its physicochemical properties are known to be considerably different from those of non-calcium-binding *c*-type lysozyme (conventional lysozyme). A notable difference has been observed in their unfolding profiles (17). A stable equilibrium intermediate has been shown to accumulate during the guanidine hydrochloride-induced unfolding of CML, although conventional lysozymes usually unfold in a cooperative two-state manner without the accumulation of an intermediate. Thus, the unfolding behavior of this protein is more similar to that of α -lactalbumin, which has been shown to exhibit a well-characterized molten globule state in its equilibrium unfolding (2, 4, 18, 19), and the unfolding intermediate of CML may also be in the molten globule state. Among the known examples of calcium-binding lysozyme, only CML and equine milk lysozyme (EML) are known to show a stable equilibrium molten globule state (20-22). In α -lactalbumin and EML, the molten globule state has been shown to be identical to an early transient intermediate in kinetic folding from the fully unfolded state (2, 4, 23-27). Interestingly, the molten globule state of EML is stabilized, at least in part, by native-like specific interactions and is more stable than that of α -lactalbumin (27).

In this thesis, I have demonstrated the folding mechanism and calcium-binding property of CML by using several biophysical techniques such as calorimetry, CD spectroscopy, NMR spectroscopy, and X-ray crystallography. The thesis consists of four CHAPTERs. In CHAPTER I, I have described about the expression system of CML by using *Escherichia coli* in order to study its physicochemical properties as shown in the following CHAPTERs. In CHAPTER II and III, I have shown that an unfolding intermediate of CML is extraordinarily stable compared with that of

the other members of the lysozyme- α -lactalbumin superfamily, which has been studied previously. And the final CHAPTER, I have investigated the conformational changes of CML by inducing calcium ions. Each content in the thesis has been published or submitted in the following journals.

CHAPTER I: Koshiba *et al.*, (1999). *Protein Engineering*, **12**, 429-435.

CHAPTER II: Koshiba *et al.*, (2000). *Biochemistry*, **39**, 3248-3257.

CHAPTER III: Koshiba *et al.*, submitted.

CHAPTER IV: Koshiba *et al.*, submitted.

References

1. Kim, P. S. & Baldwin, R. L. (1990) *Annu. Rev. Biochem.*, **59**, 631-660.
2. Kuwajima, K. (1989) *Proteins: Struct. Funct. Genet.*, **6**, 87-103.
3. Ptitsyn, O. B. (1995) *Adv. Protein Chem.*, **47**, 83-229.
4. Kuwajima, K. (1996) *FASEB J.*, **10**, 102-109.
5. Ebner, K. E., Denton, W. L. & Brodbeck, U. (1966) *Biochem. Biophys. Res. Commun.*, **24**, 232-236.
6. Richardson, R. H. & Brew, K. (1980). *J. Biol. Chem.*, **255**, 3377-3385.
7. McKenzie, H. A. & White, F. H. Jr. (1991) *Adv. Protein Chem.*, **41**, 173-315.
8. Brew, K., Castellion, F. J., Vanaman, T. C. & Hill, R. L. (1970) *J. Biol. Chem.*, **245**, 4570-4582.
9. Qasba, P. K. & Safaya, S. K. (1984) *Nature*, **308**, 377-380.
10. Smith, S. G., Lewis, M., Aschaffenburg, R., Fenna, R. E., Wilson, I. A., Sundaralingam, S. D. I., Stuart, D. I. & Phillips, D. C. (1987) *Biochem. J.*, **342**, 353-360.
11. Acharya, K. R., Stuart, D. I., Walker, N. C. P., Lewis, M. & Phillips, D. C. (1989) *J. Mol. Biol.*, **208**, 99-127.
12. Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S. & Murai, N. (1980) *Biochem. Biophys. Res.*

Commun., **95**, 1098-1104.

13. Sugai, S. & Ikeguchi, M. (1994) *Adv. Biophys.*, **30**, 37-84.
14. Nitta, K., Tsuge, H., Sugai, S. & Shimazaki, K. (1987) *FEBS Lett.*, **223**, 405-408.
15. Nitta, K. & Sugai, S. (1989) *Eur. J. Biochem.*, **182**, 111-118.
16. Grobler, J. A., Rao, K. R., Pervaiz, S. & Brew, K. (1994) *Arch. Biochem. Biophys.*, **313**, 360-366.
17. Kikuchi, M., Kawano, K. & Nitta, K. (1998) *Protein Sci.*, **7**, 2150-2155.
18. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. (1976) *J. Mol. Biol.*, **106**, 359-373.
19. Uchiyama, H., Perez-Prat, E. M., Watanabe, K., Kumagai, I. & Kuwajima, K. (1995) *Protein Eng.*, **8**, 1153-1161.
20. Nitta, K., Tsuge, H. & Iwamoto, H. (1993) *Int. J. Pept. Protein Res.*, **41**, 118-123.
21. Van Dael, H., Haezebrouck, P., Morozova, L., Arico-Muendel, C. & Dobson, C. M. (1993) *Biochemistry*, **32**, 11886-11894.
22. Griko, Y. V., Freire, E., Privalov, G., Van Dael, H. & Privalov, P. L. (1995) *J. Mol. Biol.*, **252**, 447-459.
23. Kuwajima, K. (1977) *J. Mol. Biol.*, **114**, 241-258.
24. Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. (1985) *Biochemistry*, **24**, 874-881.
25. Kuwajima, K. (1977) *Methods Mol. Biol.*, **40**, 115-135.
26. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986) *Biochemistry*, **25**, 6965-6972.
27. Mizuguchi, M., Arai, M., Ke, Y., Nitta, K. & Kuwajima, K. (1998) *J. Mol. Biol.*, **283**, 262-277.

CHAPTER I

Expression of a Synthetic Gene Encoding Canine Milk Lysozyme in *Escherichia coli* and Characterization of the Expression Protein

Abstract

A high-expression plasmid of the canine milk lysozyme, which belongs to the family of calcium-binding lysozymes, has been constructed in order to study its physicochemical properties. Because the cDNA sequence of the protein has not yet been determined, a 400 base-pair gene encoding canine milk lysozyme was first designed on the basis of the known amino acid sequence. The gene was constructed by an enzymatic assembly of 21 chemically synthesized oligonucleotides and inserted into an *Escherichia coli* expression vector by stepwise ligation. The expression plasmid thus constructed was transformed into BL21(DE3)/pLysS cells. The gene product accumulated as inclusion bodies in an insoluble fraction. Recombinant canine milk lysozyme was obtained by purification and refolding of the product and showed the same characteristics in terms of bacteriolytic activity and the far- and near-UV CD spectra as the authentic protein. The NMR spectra of refolded lysozyme were also characteristic of a native globular protein. I thus concluded that recombinant canine milk lysozyme was folded into the correct native structure. Moreover, the thermal unfolding profiles of the refolded recombinant lysozyme showed a stable equilibrium intermediate, indicating that the molten globule state of this protein was extraordinarily stable. This expression system of canine milk lysozyme will enable us to extend the biophysical and structural studies of this protein.

Introduction

In the field of protein chemistry, it is desirable to obtain large amounts of desired proteins with high efficiency. Production of large amounts of desired proteins in *Escherichia coli* (*E. coli*) is a first step in the study of protein chemistry. However, many eukaryotic proteins are often produced in *E. coli* as insoluble aggregates, inclusion bodies, and these must be refolded to obtain functional proteins. In refolding process of disulfide-containing proteins, such as RNase A, in particular, disulfide bonds must be rearranged correctly. To date, the refolding of such proteins has been carried out by rapid dilution into aqueous solution containing both oxidizing and reducing reagent (1-3) or by gentle dialysis with such aqueous solution (4). In the case of several lysozymes, since it is difficult to refold from inclusion body with high efficiency by using these dilution or dialysis methods (Koshiba, unpublished results), I have carried out the refolding process using *E. coli* thioredoxin, a known catalyst for reducing protein disulfides (5).

In this CHAPTER, a high-expression plasmid of the canine milk lysozyme (CML) gene has been constructed in order to study the physicochemical properties of CML in relation to lysozyme- α -lactalbumin evolution and its folding mechanisms. Because the cDNA sequence of CML has not yet been determined, chemically synthesized deoxyoligonucleotides, the sequences of which are based on the known amino acid sequences of the protein (6), were used for constructing the plasmid. I show that the protein over-expressed as inclusion bodies in *Escherichia coli* is efficiently refolded to the native state through the use of an *E. coli* thioredoxin refolding system (7) and that the refolded protein exhibits full enzymatic activity as an authentic lysozyme. Furthermore, the thermal unfolding transitions of the protein have shown that the molten globule state of CML is extraordinarily stable, even more stable than that of equine milk lysozyme (EML) (8, 9). Therefore, CML is a useful model protein for elucidating the molecular mechanisms by which the molten globule state is stabilized and determining the role of this state in protein folding.

Materials and methods

Materials

All enzymes used for the genetic experiments were purchased from Takara Shuzo (Kyoto, Japan). The 21 deoxyoligonucleotides that were used for construction of the CML gene (Figure 1) were synthesized by RIKAKEN Co. Ltd (Tokyo, Japan). SP-Sepharose FF and Sephacryl S-200 were obtained from Pharmacia Biotech (Uppsala, Sweden). *Micrococcus lysodeikticus* for measuring the enzymatic activity of lysozymes was obtained from Sigma Chemicals (St Louis, USA). Hen egg white lysozyme and human lysozyme were from Sigma Chemicals (St Louis, USA). Authentic EML and CML have been prepared previously (10, 11). All other reagents were of biochemical research grade.

Construction of a synthetic CML gene

The CML gene was assembled from three gene segments, PS1, PS2, and PS3, and each segment was constructed from seven oligonucleotides (Figure 1A).

To prepare the PS1 segment, each of the seven oligonucleotides from PS1-1 to PS1-7 was phosphorylated in a 20 μ L reaction mixture that contained 50 pmol of the purified oligonucleotide, 50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 10 mM β -mercaptoethanol, 5 μ M ATP, and 40 units of T4 polynucleotide kinase. After being incubated for 30 minutes at 37° C, the kinase was inactivated by heating at 65° C for 15 minutes. The phosphorylated oligonucleotides were then mixed and annealed in a 10 μ L reaction mixture that contained 1 μ L of the above phosphorylation reaction mixture of each oligonucleotide and 3 μ L of water. The oligonucleotide mixture was heated at 100° C for 3 minutes and then cooled slowly to room temperature over a period of 1 hour to allow the complementary strands to anneal. The annealed segment, PS1, was ligated into a cloning vector (pMT7); pMT7 was double-digested with *Nco* I and *Bam* HI, then mixed with the annealed segment

with a molar ratio of 1:10 (vector : insert), and then ligated. The ligation was carried out using a TAKARA Version 2 ligation kit and this generated a plasmid pMT7PS1 (Figure 1B). *E. coli* strain MV1184 was transformed by the ligation mixture, and the pMT7PS1 DNA was purified from the transformed cells. The inserted segment was confirmed by DNA sequencing (Pharmacia, Sweden).

The other two segments, PS2 and PS3, were prepared in the same manner as shown above. The PS2 segment was inserted into pMT7PS1 to generate pMT7PS1/2, and the PS3 segment was inserted into pMT7PS1/2 to generate pMT7PS1/2/3, which contained the entire canine milk lysozyme gene (Figure 1B).

Construction of a versatile expression plasmid (pHK7CML)

To construct a versatile expression plasmid of CML, the entire gene segment excised from pMT7PS1/2/3 was cloned into a plasmid pSCREEN 1-b(+) (Novagen Inc., Madison, USA). The pMT7PS1/2/3 was digested with *Pvu* II and *Xba* I, and the CML gene segment obtained was ligated into the *Eco* RV-*Xba* I sites of pSCREEN 1-b(+). The resulting plasmid was called pHK7CML. The pHK7CML contains T7 transcription and translation signals, both of which result in the high expression of CML, and the *f1* origin of replication, which allowed us to prepare the single-stranded plasmid DNA.

Expression and partial purification of recombinant CML

Recombinant CML was overexpressed in *E. coli* (BL21(DE3)/pLysS) that had been transformed by pHK7CML. The protein was expressed as inclusion bodies and was solubilized in 6M guanidine hydrochloride containing 50 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 650 mM β -mercaptoethanol. The solubilized protein was partially purified by gel filtration on a column of Sephacryl S-200 as described previously (7).

Refolding and isolation of recombinant CML

The refolding of recombinant CML that had been partially purified by gel filtration (see above) was carried out by the method of Koshihara *et al.* (7) with slight modifications. Thioredoxin (20 mL; 50 μ M) was first preincubated with 10 μ L of DTT (10 mM for 30 minutes, and the resultant reduced thioredoxin solution was mixed with 20 mL of oxygenated thioredoxin (50 μ M) to make a refolding buffer solution. After mixing the refolding buffer solution, 4.1 mL of denatured recombinant CML (25 μ M) and 90 μ L of CaCl_2 (0.5 M) were immediately added to this refolding buffer solution. The reaction mixture was incubated for 4 days at 25° C, then dialyzed against 20 mM Tris-HCl buffer (pH 8.0) containing 1 mM CaCl_2 and applied on a SP-Sepharose FF column (8 $\times$ 50 mm, Pharmacia). Refolded recombinant CML was eluted from the column by a linear NaCl gradient from 0 to 0.3 M. The eluate containing lysozyme was collected, dialyzed against 0.01 M HCl, and lyophilized.

Measurements of lysozyme activity

The bacteriolytic activity of lysozymes was assayed by the method of Kumagai *et al.* (12).

Preparation of Apo- and Holo-proteins

Preparation of apo- and holo-proteins were performed as described previously (7).

Measurements of CD spectra

The CD spectra of recombinant CML (27.5 μ M) were measured on a Jasco J-500 spectropolarimeter (Tokyo, Japan). Far- and near-UV CD spectra were measured using a cell with 0.1- and 1-cm optical path lengths, respectively, at pH 4.5 and 25° C.

Differential scanning calorimetry (DSC)

Differential scanning calorimetric (DSC) measurements of recombinant CML were carried out with an MC-2 microcalorimeter (MicroCal Inc., Northampton, USA) at a scanning rate of 1.0 K/min. The sample solution for DSC measurement was prepared by dissolving lyophilized lysozyme in a 50 mM sodium acetate buffer (pH 4.5). The lysozyme concentration used was 1.5 mg/mL. The pH of the sample solution was checked before and after each measurement. The calorimetric enthalpy (ΔH_{cal}) was evaluated using Origin computer software (Microcal Inc.).

One-dimensional ^1H NMR spectra

One-dimensional ^1H NMR spectra were obtained on a JEOL α -400 spectrometer operating at 400 MHz. The sample solution for ^1H NMR measurement was prepared by dissolving lyophilized lysozyme in D_2O and the pH was adjusted with small quantities of NaOD or DCl to pH 4.5. In all experiments, the sweep width was 15.0 ppm, and the digital resolution 0.36 Hz/point. The water signal was suppressed by DANTE pulse-sequence (13).

SDS-polyacrylamide gel electrophoresis

The expression level of recombinant CML in *E. coli* was analyzed by SDS-polyacrylamide gel electrophoresis according to the method of Laemmli (14).

Estimation of protein concentration

The protein concentrations were estimated spectrophotometrically using the following extinction coefficients at 280 nm: $E^{1\%}_{1\text{cm}} = 23.2$ for authentic and recombinant CML (10) and $E^{1\%}_{1\text{cm}} = 11.4$ for *E. coli* thioredoxin (15).

Results

Design of the synthetic CML gene

The cDNA sequence of CML has not yet been determined, I therefore designed a synthetic gene that encodes CML on the basis of its known amino acid sequence (6). The design of the synthetic gene progressed as follows. First, the amino acid sequence of the lysozyme was reverse-translated into a DNA sequence through the use of the triplets that occur most frequently in *E. coli* (16). Second, the sequence was checked by a computer program (GENETYX) to ensure that there were no excessively stable secondary structures in the encoded mRNA, especially in the upstream region of the gene. Third, for obvious convenience, unique restriction endonuclease recognition sites were introduced into the sequence without changing the encode amino acid sequence.

Construction of the synthesis gene and the expression plasmid

The three segments (PS1, PS2, and PS3) that made up the designed lysozyme gene were constructed from 21 oligonucleotides ranging in size from 20 to 56 (Figure 1A). The three segments were: (i) the 152-bp *Nco* I/*Bam* HI fragment composed of seven oligonucleotides (PS1), (ii) the 133-bp *Bam* HI/*Sph* I fragment composed of seven oligonucleotides (PS2), and (iii) the 115-bp *Sph* I/*Pvu* II fragment composed of seven oligonucleotides (PS3). In order to use the *Nco* I site of the cloning vector pMT7, PS1 segment had a *Nco* I cohesive end, and the initiation codon (ATG) was added at the 5'-end of the synthetic lysozyme gene. Thus, the recombinant protein expressed in *E. coli* had an extra methionine residue at the N-terminus.

The assembly of the synthetic gene was carried out by a stepwise ligation of the three gene segments into pMT7 that had been used for the overexpression of staphylococcal nuclease in *E. coli* (17), to create the plasmid pMT7PS1/2/3 (Figure 1B). The insert of the assembled gene into pMT7 was confirmed by DNA sequencing.

To construct an expression plasmid of CML, pMT7PS1/2/3 was digested with *Pvu* II and *Xba* I, and the lysozyme gene fragment was ligated directly into an expression vector, pSCREEN-1b(+) (Novagen), which had been digested with *Eco* RV and *Xba* I (Figure 1B). The map of the expression plasmid thus obtained, pHK7CML, is shown in Figure 2.

Expression and purification of recombinant CML

The expression plasmid pHK7CML was transformed into competent BL21(DE3)/pLysS cells, and the protein was expressed as inclusion bodies (Figure 3). The expression level of the protein was 40-50 mg/L of culture as estimated by SDS-PAGE. The protein expressed as inclusion bodies was solubilized in 6M guanidine hydrochloride solution and purified by gel filtration using Sephacryl S-200. Denatured recombinant canine milk lysozyme with more than 90% purity was obtained, and this sample was used for refolding (see below).

Refolding of recombinant CML

My previous study has shown that the use of a refolding system that utilizes *E. coli* thioredoxin is very effective in completing the refolding of denatured and disulfide-reduced human lysozyme (7). This refolding system was therefore also used in the present study in the refolding of the recombinant CML, with the following modification; 1 mM CaCl_2 was added to the refolding reaction mixture. Because CML is a calcium-binding protein, calcium ions are expected to promote the refolding. In fact, when CaCl_2 was present, a significant decrease in the turbidity of the refolding mixture of CML was observed as compared with the refolding mixture prepared in the absence of calcium ions.

Figure 4 shows the refolding process of CML as monitored by the lysozyme activity as a function of the refolding time. Recombinant CML fully refolded at 100 hours after the refolding

started. The refolding of recombinant CML was found to be faster than that of the human lysozyme previously studied (7).

Characterization of refolded recombinant CML

I measured the bacteriolytic lysozyme activity, far (210-250 nm) and near (250-310 nm) UV CD spectra, and 400-MHz ^1H NMR spectra of refolded recombinant CML. The thermal transition of the protein measured by DSC was also investigated.

Figure 5 shows the bacteriolytic activities of various lysozymes including refolded CML assayed in 50 mM sodium phosphate buffer (pH 6.2). The bacteriolytic activity of recombinant CML, which was found to be the same as the activity of authentic CML, was two-fold lower than the activity of hen egg white lysozyme.

Figure 6 shows the CD spectra of recombinant apo- and holo-CML measured in a 50 mM sodium acetate buffer (pH 4.5) containing 1 mM EDTA or 1 mM CaCl_2 , respectively. Both the spectra of recombinant apo- and holo-CML in the region from 210 to 250 nm and 250 to 310 nm are very similar to the spectra of the authentic CML studied previously (10), suggesting that recombinant CML has folded into the same secondary and tertiary structures as the authentic lysozyme. The native structures of recombinant lysozyme in the apo- and holo-states were also characterized by ^1H NMR spectra (Figure 8A and B). Both the apo- and holo-state spectra show characteristic chemical-shift dispersions, including upfield shifts of aliphatic protons between 0 and -1 ppm, for a globular protein having the folded native structure.

DSC measurements of the thermal unfolding of recombinant CML were carried out in 50 mM sodium acetate buffer (pH 4.5) containing 10 mM CaCl_2 . Under the condition, α -lactalbumin and EML are known to show a single heat absorption peak caused by a cooperative transition from the native to the thermally unfolding state (8, 9, 18, 19). Figure 7A shows a DSC curve of the

thermal transition of recombinant CML. The results indicate that the unfolding transition of recombinant CML occurs around 68° C in a highly cooperative manner, which is characteristic of a folded globular protein. However, unlike α -lactalbumin and EML, CML shows the second heat absorption peak centered at 90° C, indicating that there is a stable unfolding intermediate in the thermal transition (see below).

The molten globule state of CML

The molten globule state of CML is stably populated at pH 2.0 in the absence of denaturants and at an intermediate concentration (~2 M) of a denaturant, guanidine hydrochloride, at neutral pH. Figure 8C shows a ¹H NMR spectrum of recombinant CML in the molten globule state at pH 2.0, and it is compared with the native spectra (Figure 8A and B) and the spectrum in the fully unfolded state at 6 M guanidine hydrochloride (Figure 8D). As previously observed in α -lactalbumin and EML, the spectrum in the molten globule state is more similar to the spectrum in the unfolded state, and the characteristic chemical shift dispersion is mostly disappeared. The spectral width is broader than the spectra in the native and the fully unfolded states, and this is also a characteristic of the molten globule state (8, 20, 21). However, it should also be noted that small resonance peaks are still present in the upfield region from 0 to -1 ppm, which arise from ring-current shifts of aliphatic protons and are due to specific packing interactions of the aliphatic side chains around an aromatic ring. Whether these ring current shifts are due to the specific interactions in the molten globule state or the incomplete unfolding of the native state, has not yet been determined.

In order to investigate the stability of the molten globule state, the thermal transition of recombinant CML in the apo state at pH 4.5 was investigated by the DSC measurement (Figure 7B). Under the condition, EML is known to show two heat absorption peaks, one caused by unfolding from the native to the molten globule state and the other caused by unfolding from the molten

globule to the thermally unfolded state (8, 9). This behavior has also been observed in the DSC curve of CML (Figure 7B), but the second transition of the protein occurs at a much higher temperature than the corresponding transition of EML (70° C) (9); the two peaks of CML are located at 45 and 90° C (Figure 7B). The results indicate that the molten globule state of CML is much more stable. Comparison of the DSC curves in the holo and apo states of CML also indicates that the extraordinary stability of the molten globule state of this protein has led to the appearance of the second heat absorption peak in the holo-protein observed above (Figure 7A).

Discussion

The purpose of this CHAPTER is to build up an expression system of CML in *E. coli* and to characterize the expressed protein. Because the cDNA sequence of the protein has not yet been determined, I have designed and constructed a synthetic gene based on the reported amino acid sequence (6). The protein, which was overexpressed from the synthetic gene in *E. coli*, formed insoluble inclusion bodies, but it was efficiently refolded through the use of the *E. coli* thioredoxin refolding system developed by Koshiba *et al.* (7). The refolded recombinant protein showed the same enzymatic activity in the lysis of *Micrococcus lysodeikticus* cells as the authentic lysozyme (Figure 5), suggesting that the protein correctly folded into the native and active conformation.

The results of the CD, the ¹H NMR spectra, and the DSC measurements (Figure 6, 7 and 8) showed that the recombinant CML refolded from the inclusion bodies have the correct native structure. It also has secondary and tertiary structures, as indicated by the CD spectra, very similar to the spectra of the authentic protein in both the far- and near-UV regions. The typical excess heat capacity curve of the thermal unfolding of recombinant CML suggests that the unfolding transition of this recombinant lysozyme occurs highly cooperatively although it shows the two different thermal transitions (Figure 7). The cooperative thermal transition is characteristic of a globular

protein having a specific tertiary structure in the native state (18, 22-24).

The ^1H NMR spectra at pH 2.0 show that the spectral features of the molten globule state are very similar to those found in α -lactalbumin and EML (Figure 8). The thermal unfolding curves measured by the DSC, however, show that the unfolding of the molten globule state occurs around 90° C, so that it is remarkably more stable than the molten globule state of EML, which shows the unfolding around 70° C (9). The cooperative heat absorption from the molten globule also suggests that the molten globule state of this protein is stabilized, at least in part, by native-like specific interactions as suggested in EML. Therefore, recombinant CML is a useful model protein for elucidating the molecular mechanisms by which the molten globule state is stabilized and determining its role in protein folding.

My preliminary study has, however, also shown that the thermal stability of the recombinant protein is significantly lower than that of authentic CML; the melting temperatures in 3.0 M guanidine hydrochloride in the presence of 10 mM CaCl_2 (pH 4.5) are 30.2 and 32.4° C for the recombinant and authentic proteins, respectively (Koshiba *et al.*, unpublished data). Automated N-terminal amino acid (five residues) sequencing showed that the recombinant CML expressed in *E. coli* possesses an extra methionine residue at the N-terminal (data not shown). This was made clear because the methionine aminopeptidase does not remove the N-terminal methionine residue when the mature sequence begins with lysine, which does not occur with the sequence of the CML (25). It is very likely that the decrease in the thermal stability of the recombinant protein compared with the authentic protein arises from the presence of the extra methionine residue at the protein N-terminus. Mine *et al.* (26) have reported that the presence of the extra methionine residue at the N-terminus in recombinant hen egg-white lysozyme brings about a decreased refolding yield and a decrease in solubility in comparison with the authentic protein (27). Recently, Ishikawa *et al.* (28) and Chaudhuri *et al.* (29) have also shown that the extraneous N-terminal methionine residue in

recombinant α -lactalbumin destabilizes the protein through an entropic effect and that the entropic destabilization arises from an increase in the conformational entropy of the recombinant protein in the unfolded state.

In conclusion, the expression system of the synthetic gene that encodes CML should enable us to extend the biophysical and structural studies of this protein and its molten globule state. The high level of recombinant CML expression from the expression plasmid (pHK7CML) will facilitate the purification of large amounts of mutant proteins required for such studies. As mentioned above, however, the recombinant protein possesses an extra N-terminal methionine residue that affects the protein stability. Therefore, it may still be necessary to modify this expression system if complete removal of the extra methionine residue is required.

References

1. Saxena, V. P. & Wetlaufer, D. B. (1970) *Biochemistry*, **9**, 5015-5023.
2. Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol.*, **33**, 231-297.
3. Maeda, Y., Ueda, T., Yamada, H. & Imoto, T. (1994) *Protein Eng.*, **7**, 1249-1254.
4. Maeda, Y., Koga, H., Yamada, H., Ueda, T. & Imoto, T. (1994) *Protein Eng.*, **8**, 201-205.
5. Pigiet, V. P. & Schuster, B. J. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 7643-7647.
6. Grobler, J. A., Rao, K. R., Pervaiz, S. & Brew, K. (1994) *Arch. Biochem. Biophys.*, **313**, 360-366.
7. Koshiba, T., Tsumoto, K., Masaki, K., Kawano, K., Nitta, K. & Kumagai, I. (1998) *Protein Eng.*, **11**, 683-690.
8. Van Dael, H., Haezebrouck, P., Morozova, L., Arico-Muendel, C. & Dobson, C. M. (1993) *Biochemistry*, **32**, 11886-11894.
9. Griko, Y. V., Freire, E., Privalov, G., Van Dael, H. & Privalov, P. L. (1995) *J. Mol. Biol.*, **252**,

447-459.

10. Kikuchi, M., Kawano, K. & Nitta, K. (1998) *Protein Sci.*, **7**, 2150-2155.
11. Mizuguchi, M., Arai, M., Ke, Y., Nitta, K. & Kuwajima, K. (1998) *J. Mol. Biol.*, **283**, 262-277.
12. Kumagai, I., Sunada, F., Takeda, S. & Miura, K. (1992) *J. Biol. Chem.*, **267**, 4608-4612.
13. Morris, G. A. & Freeman, R. (1978) *J. Magn. Reson.*, **29**, 433-470.
14. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
15. Pigiet, V. P. & Conley, R. R. (1978) *J. Biol. Chem.*, **253**, 1910-1920.
16. Sharp, P. M. & Li, W. H. (1987) *Nucleic Acids Res.*, **15**, 1281-1295.
17. Ikura, T., Tsurupa, G. P. & Kuwajima, K. (1997) *Biochemistry*, **36**, 6529-6538.
18. Yutani, K., Ogasahara, K. & Kuwajima, K. (1992) *J. Mol. Biol.*, **228**, 347-350.
19. Griko, Y. V., Freire, E. & Privalov, P. L. (1994) *Biochemistry*, **33**, 1889-1899.
20. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986) *Biochemistry*, **25**, 6965-6972.
21. Nitta, K., Tsuge, H. & Iwamoto, H. (1993) *Int. J. Pept. Protein Res.*, **41**, 118-123.
22. Khechinashvili, N. N., Privalov, P. L. & Tiktopulo, E. I. (1973) *FEBS Letters*, **30**, 57-60.
23. Privalov, P. L. & Khechinashvili, N. N. (1974) *J. Mol. Biol.*, **86**, 665-684.
24. Pfeil, W. & Privalov, P. L. (1976) *Biophys. Chem.*, **4**, 23-50.
25. Moerschell, R. P., Hosokawa, Y., Tsunasawa, S. & Sherman, F. (1990) *J. Biol. Chem.*, **265**, 19638-19644.
26. Mine, S., Ueda, T., Hashimoto, Y. & Imoto, T. (1997) *Protein Eng.*, **10**, 1333-1338.
27. Imoto, T., Yamada, H., Yasukochi, T., Yamada, E., Ito, Y., Ueda, T., Nagatani, H., Miki, T. & Horiuchi, T. (1987) *Protein Eng.*, **4**, 333-338.
28. Ishikawa, N., Chiba, T., Chen, L. T., Shimizu, A., Ikeguchi, M. & Sugai, S. (1998) *Protein Eng.*, **11**, 333-335.
29. Chaudhuri, T. P., Horii, K., Yoda, T., Arai, M., Nagata, S., Terada, T. P., Uchiyama, H., Ikura, T.,

Tsumoto, K., Kataoka, H., Matsushima, M., Kuwajima, K. & Kumagai, I. (1998) *J. Mol. Biol.*, **285**, 1179-1194.

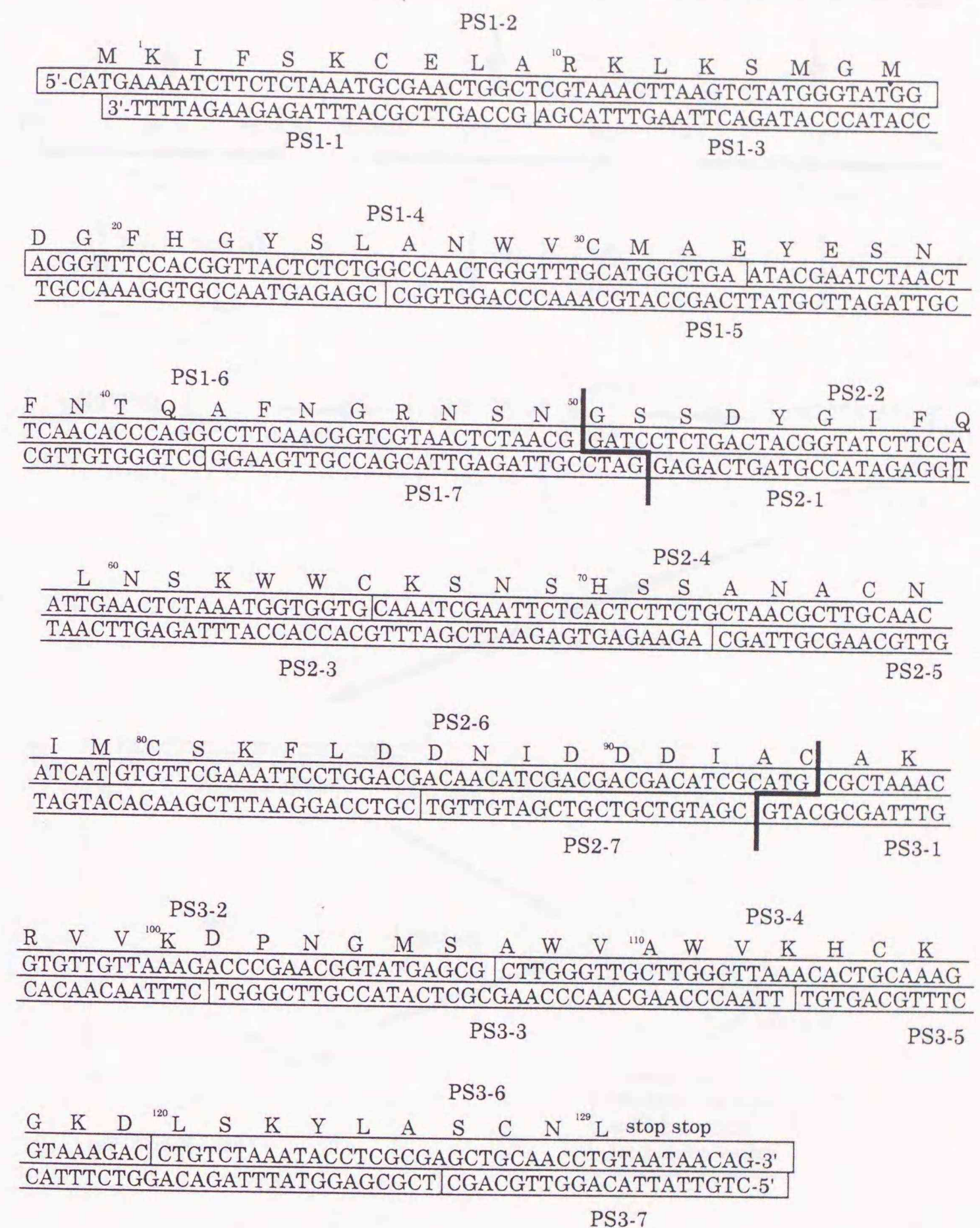


Figure 1A. The DNA sequence of the synthetic CML gene. Each box indicates each of 21 oligonucleotides used to assemble the lysozyme gene. The thick lines indicate the ends of three gene segments (PS1, PS2, and PS3) (see text). The amino acid sequence is shown above the DNA sequence.

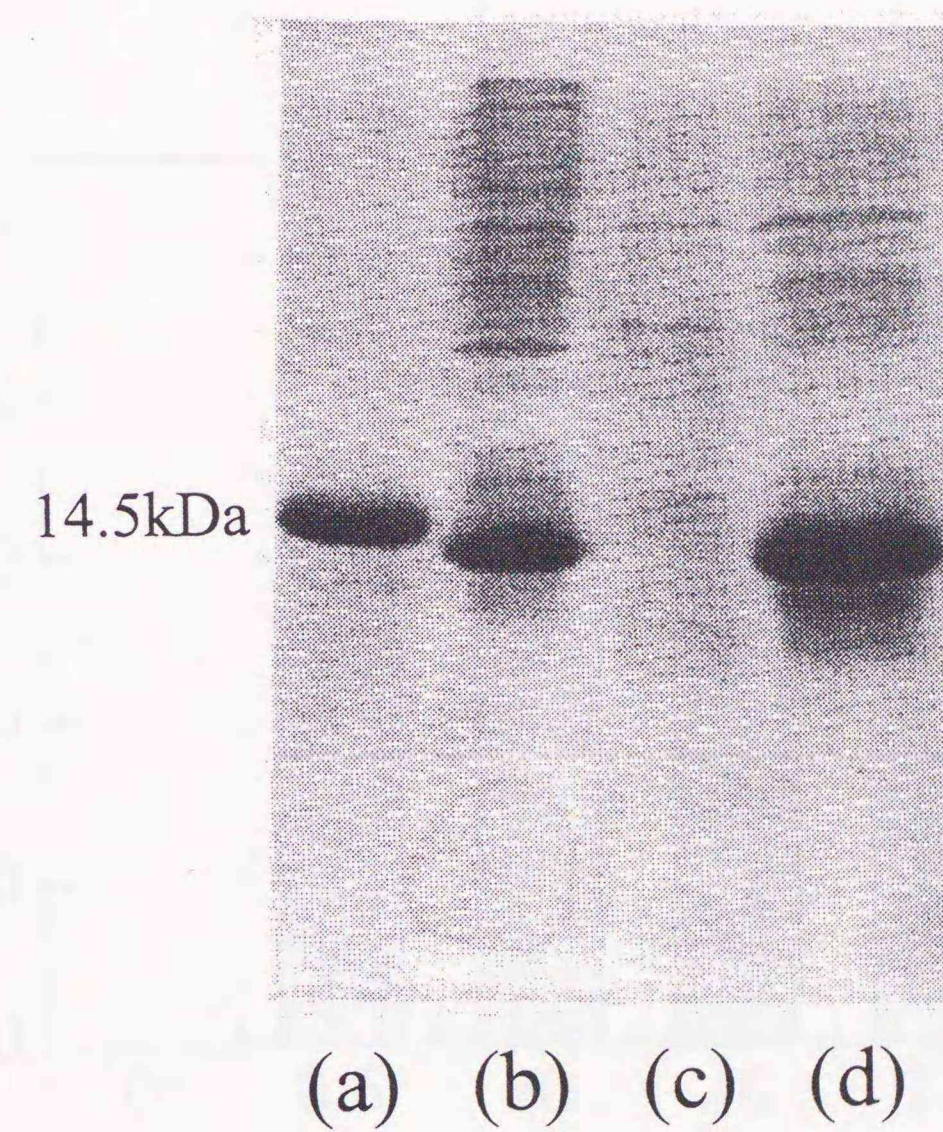


Figure 3. The SDS-gel electrophoretic pattern of recombinant CML expressed in *E. coli* BL21(DE3)/pLysS cells. The lanes shown are: (a) native human lysozyme; (b) the whole cells grown in LB media, after 3 hours of IPTG induction; (c) supernatant of the ultrasonicated cells; (d) precipitate of the ultrasonicated cells that includes inclusion bodies. The concentration of polyacrylamide gel was 18.8%.

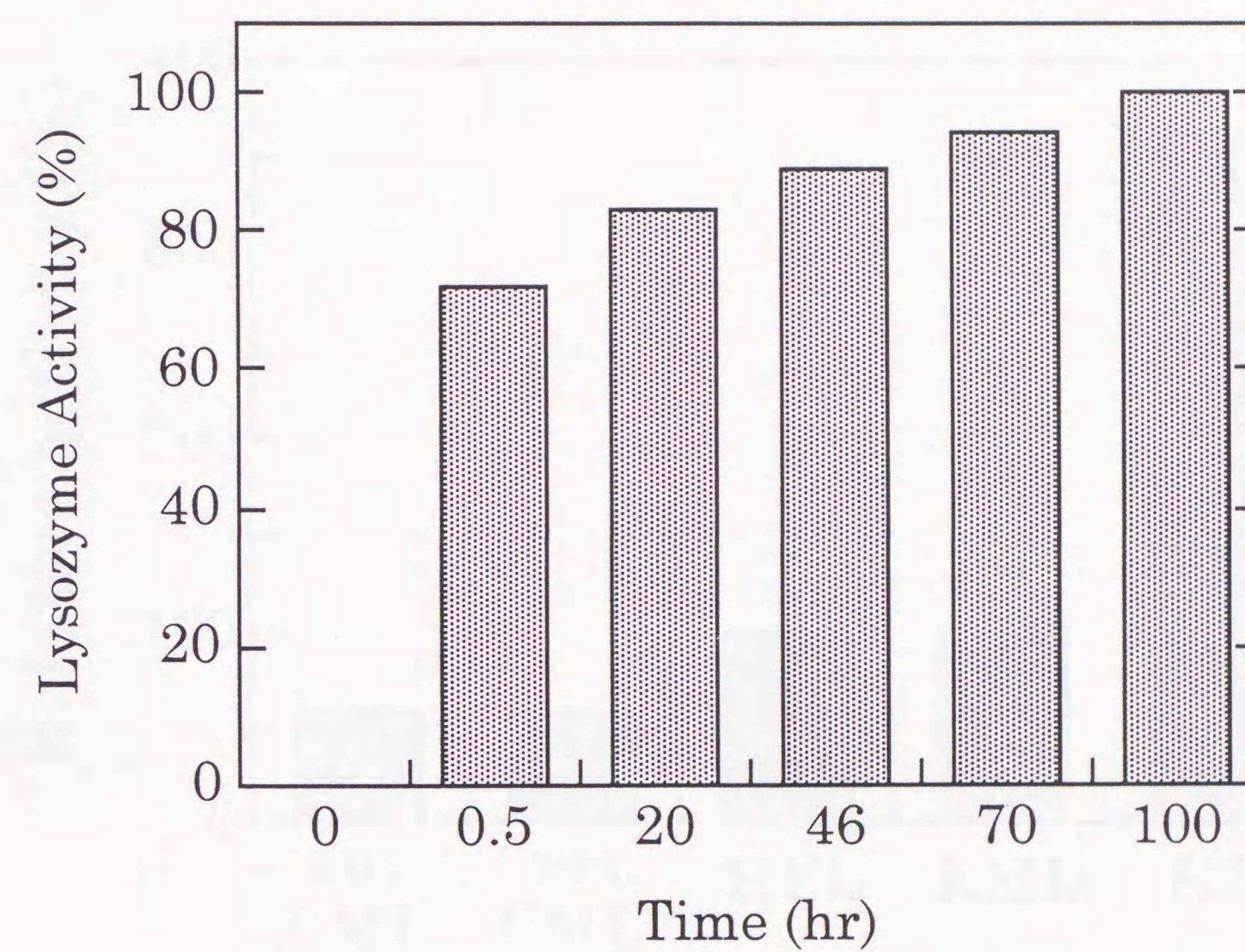


Figure 4. Kinetics of the refolding of denatured recombinant CML in the thioredoxin refolding system containing 1 mM CaCl_2 . A 20 mL of 50 μM thioredoxin was preincubated with 10 μL of 10 mM DTT for 30 min and added to 20 mL of 50 μM thioredoxin that was bubbled with oxygen immediately prior to the addition of 4.1 mL of 25 μM denatured recombinant CML. The refolding reaction was carried out in 20 mM Tris-HCl buffer (pH8.0) containing 1 mM CaCl_2 and 0.2 M NaCl at 25° C.

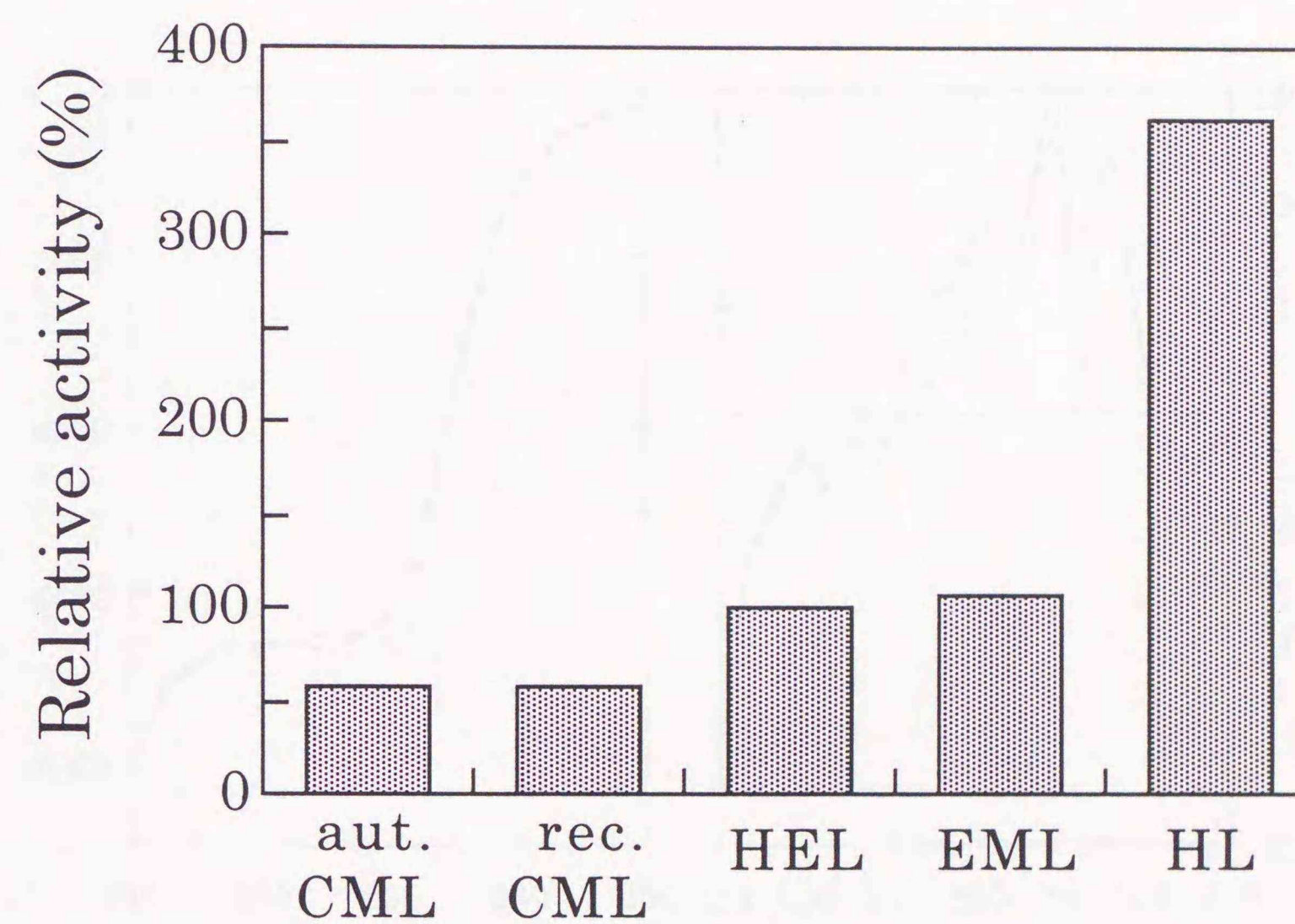


Figure 5. Bacteriolytic activities of various lysozymes and refolded recombinant CML. The relative activities are expressed by taking the activity of hen egg white lysozyme as 100%. Aut. CML, rec. CML, HEL, EML, and HL represent authentic canine milk lysozyme, recombinant canine milk lysozyme, hen egg white lysozyme, equine milk lysozyme, and human lysozyme, respectively.

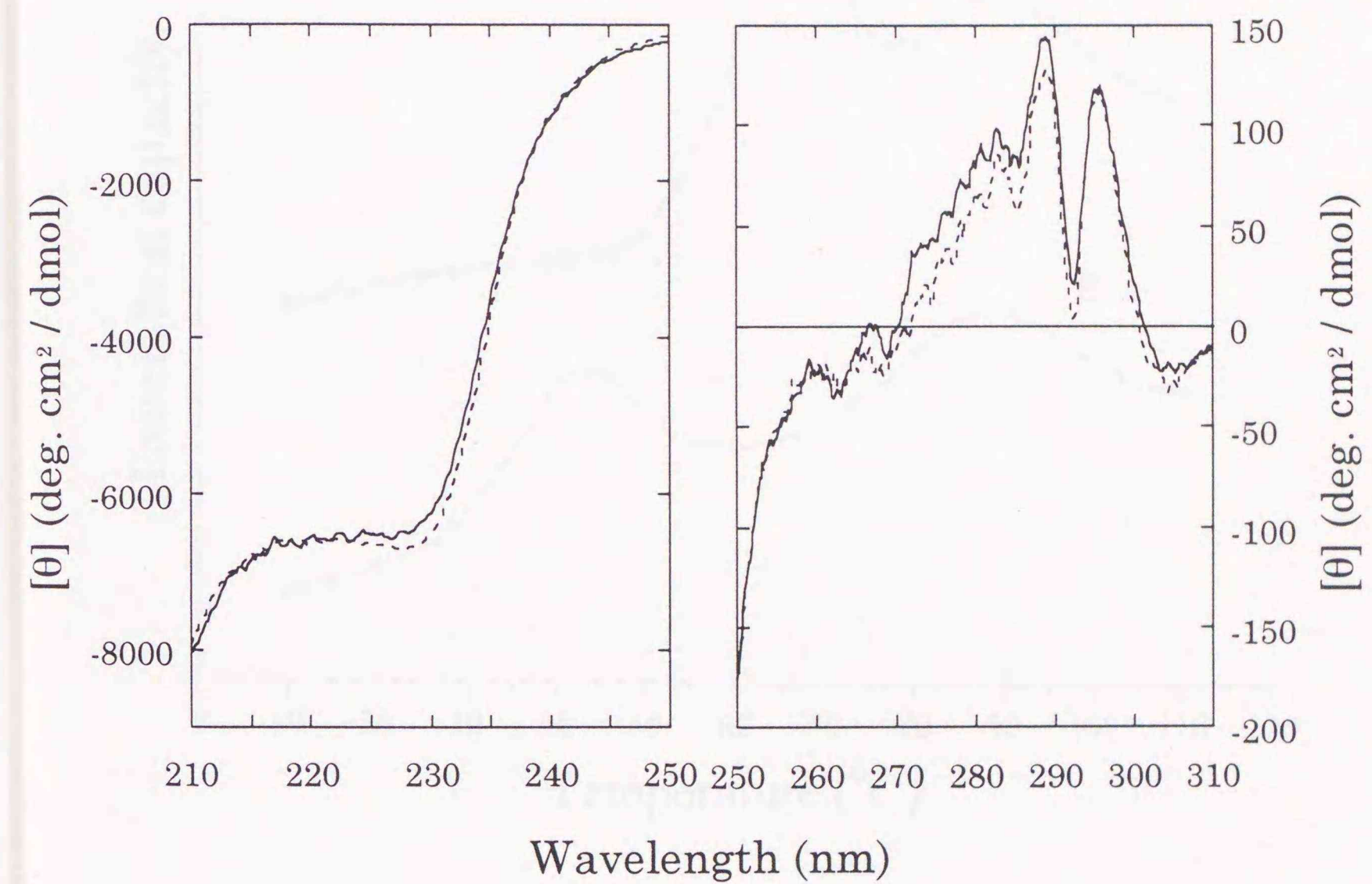


Figure 6. CD spectra in the near and far UV regions of refolded recombinant apo- and holo-CML in the presence of 1 mM EDTA or CaCl₂ at 25° C and pH 4.5. Solid line and dashed line represent apo- and holo-CML, respectively.

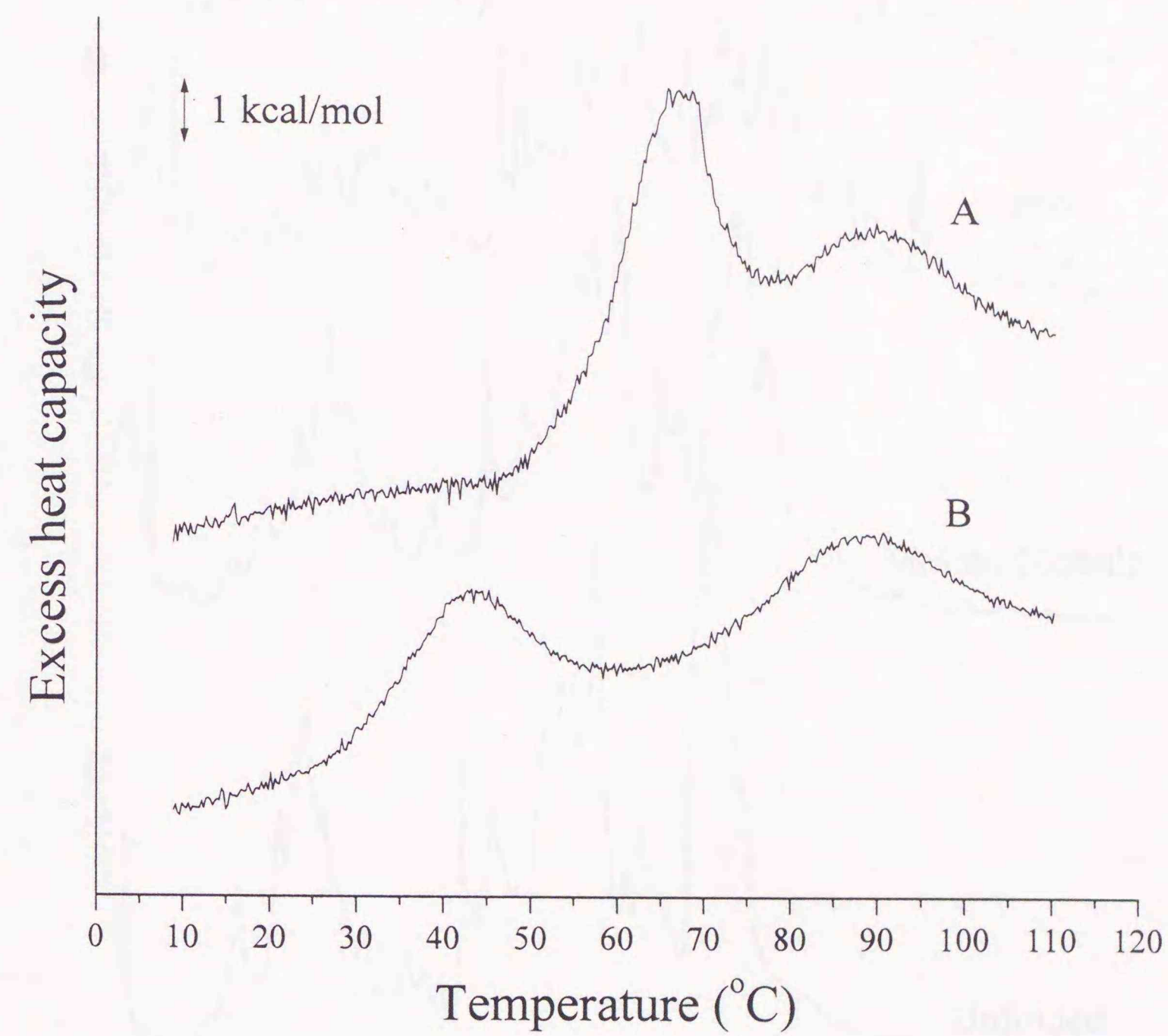


Figure 7. The thermal unfolding transition of refolded recombinant apo- and holo-CML at pH 4.5. The unfolding transition is expressed by excess heat capacity as a function of temperature. (A): Holo-lysozyme in the presence of 10 mM CaCl_2 (B): Apo-lysozyme in the absence of CaCl_2

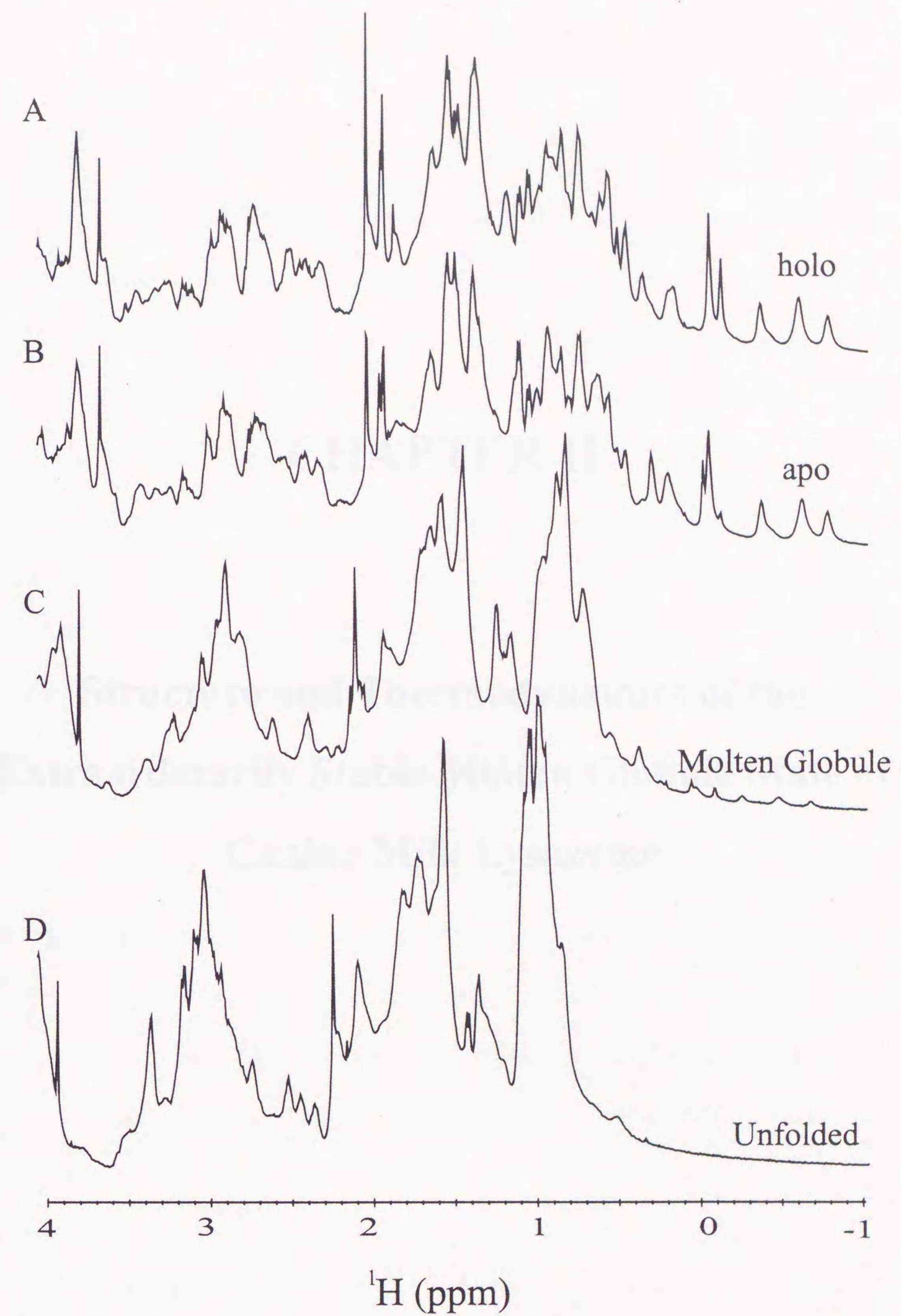


Figure 8. ^1H NMR spectra of recombinant CML at 25°C in D_2O . The spectra shown are: (A); Holo-lysozyme with 1 mM CaCl_2 at pH 4.5 (B); Apo-lysozyme at pH 4.5 (C); Apo-lysozyme at pH 2.0 (D); Apo-lysozyme in 6 M guanidine hydrochloride at pH 4.5.



Fig. 1. Denaturation curves of canine milk lysozyme. The curves show the change in optical activity (or circular dichroism) as a function of temperature. The x-axis represents temperature in degrees Celsius, and the y-axis represents the measured property. The curves are labeled 1, 2, 3, and 4, corresponding to different experimental conditions or samples.

CHAPTER II

Structure and Thermodynamics of the Extraordinarily Stable Molten Globule State of Canine Milk Lysozyme

Abstract

Here I show that an unfolded intermediate of canine milk lysozyme is extraordinarily stable compared with that of the other members of the lysozyme- α -lactalbumin superfamily, which has been studied previously. The stability of the intermediate of this protein was investigated using calorimetry, CD spectroscopy, and NMR spectroscopy, and the results were interpreted in terms of the structure revealed by X-ray crystallography at a resolution of 1.85 Å to an *R*-factor of 17.8 %. Based on the results of the thermal unfolding, this protein unfolds in two clear cooperative stages, and the melting temperature from the intermediate to the unfolded states is about 20° C higher than that of equine milk lysozyme. Furthermore, the ¹H NMR spectra of canine milk lysozyme at 60° C, essentially 100 % of which exists in the intermediate, showed that small resonance peaks that arise from ring-current shifts of aliphatic protons are still present in the upfield region from 0 to -1 ppm. The protein at this temperature (60° C) and pH 4.5 has been found to bind 1-anilino-naphthalene-8-sulphonate (ANS) with enhancement of the fluorescence intensity compared with that of native and thermally unfolded states. I interpret that the extraordinarily stable intermediate is a molten globule state, and the extraordinary stabilization of the molten globule state comes from stronger protection around the C- and D-helix of the aromatic cluster region due to the residue of His-21. The conclusion helps to explain how the molten globule state acquires its structure and stability.

Introduction

To understand the protein folding mechanism, the concept of a folding intermediate, such as a molten globule state, is extremely important. The molten globule state has been thought to be an equilibrium and kinetic intermediate between the native and unfolded states of a number of globular proteins (1-6). For some time, the molten globule state of a protein has been assumed to be a compact intermediate that possesses a stable secondary structure without native tertiary contacts, and α -lactalbumin, cytochrome *c*, apomyoglobin, and several other proteins have often been utilized as a useful model protein for studying the folding mechanisms (4-11). However, with the extensive progress in studying the molten globule state, new evidence has been accumulated that shows specific tertiary interactions or native-like topology within the molten globule state of such proteins as α -lactalbumin (12-19), apomyoglobin (20-23) and equine milk lysozyme (EML) (24-28).

Canine milk lysozyme (CML) belongs to the family of calcium-binding lysozymes (29), and its physicochemical properties are known to be rather different from those of conventional, non-calcium-binding lysozymes. In particular, notable difference has been observed in their unfolding profiles (30). A stable equilibrium intermediate has been shown to accumulate during the guanidine hydrochloride-induced unfolding of CML, although conventional lysozymes usually unfold in a cooperative two-state manner without the accumulation of an intermediate. Thus, the unfolding behaviors of this protein and EML are rather similar to those of α -lactalbumin, which has been shown to exhibit a well characterized molten globule state in its equilibrium unfolding (4, 7, 9).

Recently, the expression system for CML has been established in *Escherichia coli* (*E. coli*), and I have found that the intermediate of this protein shows extraordinary stabilization (31). In this CHAPTER, I show that CML is an excellent model for elucidating the molecular mechanisms by which the intermediate is stabilized. In order to investigate the stability of the intermediate of CML, its unfolding behavior was studied by differential scanning calorimetry (DSC), UV-CD spectroscopy,

and NMR spectroscopy, and the results were interpreted in terms of the structure revealed by X-ray crystallography. Moreover, the 1-anilino-naphthalene-8-sulphonate (ANS) binding experiment was also performed for the purpose of detecting the formation of the loose hydrophobic core and investigating the hydrophobic surface accessible to solvent of the intermediate of this protein. The X-ray crystal structure of this protein was determined by a molecular replacement method and refined at a resolution of 1.85 Å to an *R*-factor of 17.8 % (free *R*-factor = 21.5 %).

From the result of ANS-binding experiment, it appears that the intermediate has been found to bind ANS with enhancement of the fluorescent intensity compared with that of the native and the thermal unfolded states. I interpret that the extraordinarily stable intermediate of this protein is a molten globule state, and the extraordinary stabilization come from its specific packing interactions in the α -domain. In support of this interpretation, the ¹H NMR spectra show that the ring-current effect is certainly present in the intermediate of CML. The relevance of the present results to the physicochemical properties of CML and the structural origins of the stability of the molten globule state will be discussed.

Materials and Methods

Materials

According to the procedures of Koshiba *et al* (1999), expression and purification of CML was carried out using the expression system in *E. coli* (31). EML has been prepared previously (32). All other reagents were of biochemical research grade.

Preparation of apo- and holo-proteins

Apo- and holo-proteins were prepared as described previously (31, 33).

Differential scanning calorimetry (DSC)

Differential scanning calorimetry (DSC) measurements of CML were carried out with an MC-2 microcalorimeter (MicroCal, Inc., Northampton, MA, USA) at a scan rate of 1.0 K/min. Sample solutions for DSC measurements between pH 4.0 and 4.5 were prepared by dissolving the lysozymes in 50 mM sodium acetate buffer, and solutions between pH 2.5 and 4.0 were prepared using 50 mM glycine hydrochloride. The lysozyme concentrations used were 1.2–1.8 mg/mL. The pH of the sample solution was confirmed before and after each measurement. Calorimetric enthalpies (ΔH_{cal}) were obtained using Origin computer software (Microcal, Inc.).

Measurements of CD spectra

CD spectra of CML and EML were measured with a Jasco J-500 spectropolarimeter (Tokyo, Japan). The protein concentrations were 0.5 mg/mL for all measurements. Far- and near-UV CD spectra were measured using a cell with 0.1- and 1 cm optical path lengths, respectively, at pH 4.5 and three different temperatures (9°C, 63°C and 98°C).

Measurements of ANS-binding

Fluorescence spectra of 50 μM ANS in the absence and presence of 25 μM apo-CML were collected using a Hitachi spectrofluorimeter 650-40 (Tokyo, Japan), equipped with a thermostatically controlled cell-holder to keep the temperature. ANS fluorescence was measured using an excitation wavelength of 380 nm. The emission spectrum was recorded over the range of 400 to 640 nm. An ANS stock solution was prepared in methanol, and the concentrations of the ANS stock solutions were determined using an extinction coefficient of $8.0 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ at 372 nm in methanol (49). Fluorescence measurements were carried out using a cell with a 1-cm optical path length at pH 4.5 and three different temperatures (9°C, 63°C and 98°C).

One-dimensional ^1H NMR spectra

One-dimensional ^1H NMR spectra were obtained on a JEOL α -400 spectrometer operated at 400 MHz and several temperatures (30, 40, 50, 60, 65, and 70°C). The sample solution for the ^1H NMR measurement was prepared by dissolving lyophilized lysozyme (of apo-type) in D_2O , and the pH was adjusted with small quantities of NaOD or DCl solution to pH 4.5. In all experiments, the sweep width was 15.0 ppm, and the digital resolution was 0.36 Hz/point. The water signal was suppressed by a DANTE pulse-sequence (34).

Crystallization, data collection, and structure refinement

A diffraction-quality crystal of apo-CML was obtained by vapor diffusion of hanging drops using 0.1 M tri-sodium citrate buffer (pH 5.6) containing 10 mg/mL protein, 0.2 M ammonium acetate, 30 % (w/v) PEG-4000, and 1 mM EDTA. After 7–10 days, a crystal with the size of $0.8 \times 0.2 \times 0.2 \text{ mm}^3$ was obtained. The space group of the crystal was $P6_122$ with cell dimensions $a = b = 68.6 \text{ \AA}$, $c = 107.1 \text{ \AA}$. There is one monomeric molecule in an asymmetric unit. The X-ray diffraction data of the apo-CML crystals were collected from a single crystal at room temperature using a Fuji imaging plate detector with a Rigaku IPR-4080 imaging scanner and a Weissenberg camera for macromolecular crystallography (35) on BL-18B at the Photon Factory, KEK, Japan. Integrated intensities were obtained with DENZO (36) and were reduced using a CCP4 package (37). The initial model of the apo-type CML was obtained by molecular replacement using the program AMoRe (38). A search model was constructed from the atomic coordinates of EML (39). The structural refinements were performed using the program X-PLOR (40).

Estimation of protein concentration

The protein concentrations were estimated spectrophotometrically using the following extinction

coefficients at 280 nm: $E_{1\text{ cm}}^{1\%} = 23.2$ for CML (30) and $E_{1\text{ cm}}^{1\%} = 23.5$ for EML (41).

Results

Differential scanning calorimetry (DSC) of CML

In order to determine the thermodynamic parameters of thermal unfolding, DSC measurements of the CML were carried out at acidic pHs between 2.7 and 4.5 in the absence of calcium ions. Figure 1 shows that the temperature dependence of the partial heat capacity of CML in acidic solutions without calcium ions. Under these conditions, CML is in apo-type. In this pH region, the unfolding of CML is reversible, and the protein exhibits a cooperative thermal unfolding with two remarkably positive heat absorption peaks, as previously reported in EML (25, 26). It is indicating that the first peak represents the unfolding transition from the native state to an unfolded intermediate, and the second peak represents the unfolding transition from the intermediate to the thermally unfolded state. The melting temperatures (T_m), at which both heat absorption peaks reach their maximum, decreases at lower pH values in almost the same manner.

The temperature dependence of the partial heat capacity of CML in 50 mM sodium acetate buffer in the absence (apo-type) and presence (holo-type) of 10 mM CaCl_2 at pH 4.5 is shown in Figure 2. In this pH, the thermal stability and enthalpy of unfolding of CML are insensitive to variations in ionic strength (sodium acetate buffer; from 10 to 100 mM) (data not shown). The temperature dependence of the fractional population of apo-CML at pH 4.5, which was estimated from Figure 2, is shown in Figure 3. From this figure, we can easily see that the intermediate is extraordinarily stable, and essentially 100 % of the protein is in this state at 60° C, compared with 70 % of EML (26). Under the presence of 10 mM CaCl_2 conditions, α -lactalbumin and EML are known to show a single heat absorption peak caused by a cooperative transition from the native to the thermally unfolded states (25, 42–45). However, unlike α -lactalbumin and EML, CML shows

the second heat absorption peak centered at 85° C, indicating that there is a stable intermediate even in the thermal unfolding in the presence of 10 mM CaCl_2 . Thus, the comparison of the DSC curves in the apo- and holo-types of CML also indicates that the extraordinary stabilization of the intermediate of this protein has led to the appearance of the second heat absorption peak in the holo-protein observed above (Figure 2).

Measurements of CD spectra

The far- (215–250 nm) and near-UV (250–310 nm) CD spectra of apo-CML and apo-EML at pH 4.5 and the temperature indicated by arrows (i–iii) in Figure 2 are depicted in Figure 4A and B. The far-UV CD spectrum of apo-CML at 63° C still shows the presence of the secondary structure, much as in the native state (i); moreover, the near-UV CD spectrum at this temperature shows the presence of elements of tertiary-like structure (Figure 4A). In Figure 4B, however for EML, we can see that CD intensity around 220 nm at 63° C is half as much as that of native state (i). Thus, the intermediate of EML has already started to disorder at this temperature. Therefore, it appears that the intermediate of CML still possesses a partially native-like structure and is highly populated at elevated temperatures (70–100° C) both in the absence and the presence of Ca^{2+} (Figure 2, 3, and 4A).

ANS binding

The fluorescent hydrophobic molecule, 1-anilino-naphthalene-8-sulfonate (ANS), is commonly used as a probe of the hydrophobic surface exposed to solvent (47, 48). The fluorescence spectra of apo-CML at pH 4.5 and the temperatures indicated by arrows (i–iii) in Figure 2 are depicted in Figure 5.

The fluorescence intensity of 50 μM ANS in the presence of 25 μM apo-CML at 63° C is

greater than that of the temperatures at 9 and 98° C. This is the characteristics of the molten globule state, which has the hydrophobic core exposed to solvent (19, 28, 48, 49). The flexible nature of the molten globule state permits some internal non-polar groups to become exposed to solvent, thus making the surface of this state more hydrophobic than that of the native or unfolded states (48, 50). Therefore, the result of ANS-binding measurement indicates that the intermediate of this protein has a characteristic of the molten globule state.

One-dimensional ¹H NMR spectra

Figure 6 shows the thermal unfolding transitions of apo-CML, which were monitored with ¹H NMR. Heating of the apo-CML induced a major change in the NMR spectra between 40 to 50° C; this temperature was comparable to the temperature (30-55° C) at which DSC and CD spectra show their first thermal transition (Figure 2, 3 and 4A). The highly dispersed resonances in the chemical shifts from 0 to -1 ppm corresponding to the native conformation show a decrease in intensity concurrent with the downfield chemical shift (31). Similar decreasing in intensity and chemical shift changes as a function of temperature have been observed in α -lactalbumin and EML. However, unlike the ¹H NMR spectra of α -lactalbumin and EML, which show no resonances from 0 to -1 ppm above 60° C (25, 28, 51-53), the spectra of apo-CML show small resonance peaks still present in the upfield region from 0 to -1 ppm above 60° C (Figure 6). Thus, the structure of the intermediate of apo-CML is more order than that of the other members of the lysozyme- α -lactalbumin superfamily.

X-ray crystal structure of apo-CML

Apo-CML was crystallized in space group of *P*6₁22, and the three-dimensional structure was determined by a molecular replacement method. The 1.85 Å structure was refined to an *R*-factor of 17.8 % (*R*_{free} = 21.5 %). The crystallographic parameters and refinement statistics are summarized

in Table 2.

Figure 7 shows the ribbon diagram of apo-CML. A deep cleft containing the active site divides the CML molecule into a large domain (α -domain) and a small domain (β -domain), both of which are similar to those previously reported for the lysozyme- α -lactalbumin superfamily (39, 54-59). The α -domain comprises residues 1-42 and 86-129, and it is composed of four major α helices (A-helix, 5-16; B-helix, 25-36; C-helix, 88-100; D-helix, 109-113) and two short helices (104-107; 123-126). The β -domain is composed of a series of loops and a small three-stranded antiparallel β -pleated sheet (S1: 43-45, S2: 51-55, and S3: 58-61).

The overall structure of the determined apo-CML is shown in Figure 8, together with those of apo-EML (39) and guinea-pig α -lactalbumin (GPLA) (58). The three proteins compared here are very similar to each other, reflecting their high similarity of their amino acid sequences (Figure 9). The root mean square deviations between CML and EML and between CML and GPLA for the main-chain C α atoms were 1.294 and 1.950 Å, respectively. Differences in main-chain conformation are observed in several distinct regions of these structures, the largest of which being in the regions between residues 43 to 55. There are also considerable differences in the regions between residues 64 to 75. One of the features of the crystal structure of apo-CML is that the loop region of the β -domain (residues 43 to 55) is much closer to the α -domain than in EML and GPLA. Therefore, according to the structure of native CML, it is likely that the β -domain structure of CML may be more stable than those of EML and GPLA.

Discussion

The purpose of this CHAPTER is to elucidate the relationship between the structure and the thermodynamics of the intermediate of CML, and compare the intermediate with those of the other members of the lysozyme- α -lactalbumin superfamily, which have been studied previously. In order

to fully understand the effect on the conformational stability of a protein, it is necessary to elucidate the thermodynamic parameters of the unfolding. The thermodynamic parameters for the unfolding [$\Delta H(T)$, $T\Delta S(T)$, and $\Delta G(T)$] of the CML and EML at 41.5° C (transition from the native state to the intermediate; N-I transition) and 66.4° C (transition from the intermediate to the thermally unfolded state; I-U transition) were obtained according to the following equations using the data from the DSC measurements in this study and Griko *et al.* (26):

$$\Delta H(T) = \Delta H(T_m) + \int_{T_m}^T \Delta C_p(T) dT, \quad (1)$$

$$\Delta S(T) = \Delta S(T_m) + \int_{T_m}^T \Delta C_p(T) d \ln T, \quad (2)$$

$$\Delta G(T) = \Delta H(T) - T\Delta S(T), \quad (3)$$

where T_m is the melting temperature and $\Delta S(T_m)$ equals to $\Delta H(T_m)/T_m$. The $\Delta C_p(T)$ values of CML and EML at these temperatures were obtained using the data of Koshiba *et al.* (46) and Griko *et al.* (26), respectively. The parameters obtained are listed in Table 1. The ΔG value of the I-U transition for CML is 2.1 kcal/mol more stable than that of EML at 66.4° C, although the ΔG values of the N-I transition between two proteins are quite similar (−0.3 kcal/mol) to each other at 41.5° C. It is indicating that the thermal stability of the intermediates between CML and EML is very different in spite of the similarity of the N-I transition. In the I-U transition at 66.4° C, the ΔH value for the unfolding of the intermediate of CML (30.9 kcal/mol) is larger than that of EML (29.5 kcal/mol). In addition, the entropy term ($T\Delta S$) in the unfolding of the intermediate of CML (28.8 kcal/mol) is smaller than that of EML (29.5 kcal/mol). These evaluations indicate that both the enthalpy and entropy terms seem to be the major terms for stabilization of the intermediate of CML.

From the results of this study, the intermediate of CML has been shown its extraordinary

stabilization, and I propose that the intermediate is a molten globule state. There are at least three lines of strong evidence for the identification of the intermediate of CML as the molten globule state. First, the intermediate of CML is stably populated at pH 2.0, and the ^1H NMR spectrum is more similar to the spectrum in the unfolded state and the characteristic chemical shift dispersion has mostly disappeared (31). The spectral width is broader than those in the native and the fully unfolded states, and these are characteristics of the molten globule state (25, 53, 80). Second, CML shows a three-state unfolding transition in guanidine hydrochloride (30) with the unfolded intermediate state very similar to the intermediate of the thermal unfolding (this study), and these intermediates are structurally very similar to the acid state (31). Such unfolding behavior and the properties of the intermediate or the acid states are essentially identical to those observed previously in α -lactalbumin and EML (4, 7, 9, 26, 27, 60, 61), for which the unfolding intermediate state has been well characterized and identified as the molten globule state. Finally, the results of the ANS-binding experiment presented in this study show that the hydrophobic fluorescence probe (ANS) has a different affinity to various conformations of CML (Figure 5). This probe is not practically bound to fully unfolded proteins as well as to a number of native proteins (48). Nevertheless, the results of this study show a very good correlation of the highest protein affinity to ANS with the conformation of the intermediate. This is a characteristic of the molten globule state, which has the hydrophobic core exposed to solvent (19, 28, 48, 49).

Recently, with the extensive progress in the study of the molten globule state, new evidence has been accumulated that shows the presence of native-like tertiary fold within the molten globule state of α -lactalbumin (12–19). Peng and Kim (1994) have shown that the helical domain (α -domain) of human α -lactalbumin, in isolation, forms a molten globule state with the same overall tertiary fold as that found in intact α -lactalbumin (12). The same group has also reported that the α -helical domain forms a helical structure with a native-like tertiary fold, while the β -sheet domain

is largely unstructured (14). The studies indicate that the molten globule state of α -lactalbumin has a native-like overall fold, even though it lacks the extensive fixed tertiary interactions that are characteristic of native proteins (12, 62).

Furthermore, Dobson and co-workers have revealed that EML, unlike other lysozymes, can be transformed into the equilibrium molten globule state (24, 25, 27, 28) and that this state is more similar to the native state than the equilibrium state of apo-type α -lactalbumin, which is usually regarded as a "classical" molten globule state (62, 63). Moreover, it has been reported that the molten globule state of EML is stabilized, at least in part, by native-like specific interactions and is more stable than that of α -lactalbumin (60). The α -helices A, B, and D of EML are especially strongly protected in the molten globule state (27, 28).

Thus, the intermediate of CML can be identified as the molten globule state, and it is thought to be similar to the molten globule state of EML, which has been distinguished from a "classical" molten globule state. However, the ^1H NMR spectrum of CML at 60° C (Figure 6), the temperature at which 100 % of the protein exists in the stable molten globule state (Figure 3), has shown interesting characteristics as compared with that of the other lysozyme- α -lactalbumin superfamily. Small resonance peaks are still present in the upfield region from 0 to -1 ppm. These peaks arise from ring-current shifts of methyl protons and are due to specific packing interactions of the aliphatic side chains around the aromatic rings; moreover, this characteristic has previously been observed in the acidic folding intermediate (pH 2.0) (31). Recently, we have sequentially assigned the chemical shifts of CML using ^{15}N -labeled protein (unpublished results), and these peaks appear to come from the methyl protons of Val-98 and βH of Met-105, as shown in Figure 6. Therefore, the NMR studies suggest that the extraordinary stabilization of the molten globule state of CML is due to the specific packing interactions of C-helix (Val-98) and Met-105 around the aromatic rings.

However, why can CML fold to the extraordinarily stable molten globule state as compared with that of EML? In the following, I address this question and show possible reasons that interpret such difference in the molten globule stability between the proteins.

The role of His-21 for the stabilization

The hydrophobic cluster of CML (Figure 10A), which originates from the aromatic residues (Tyr-23, Trp-28, -63, -64, -108, and -111) around Val-98 and Met-105, is compared with that of EML (Figure 10B). In the crystal structure of CML, His-21 is located at the turn region between the A- and B-helices, and is close to the C-helix; however, this position in EML is replaced by glycine (Figure 9). The distance between the $\text{N}^{\epsilon 2}$ atom of His-21 and O atom ($\text{C}=\text{O}$) of Val-99 is 3.13 Å, and a hydrogen bond is formed between these two atoms, as shown in Figure 11. As a result, the A- and B-helix become much closer to the C-helix, and hydrophobicity around this region is increased. Thus, the substitution of Gly-21 with His-21 can be suggested to have two effects on the stabilization of the molten globule state. First, by forming a hydrogen bond between the side chain of His-21 and the main chain of Val-99, the α -helices A and B are connected to the C-helix, and these chains become less flexible. Second, the involvement of the His-21 residue on the hydrophobic cluster makes the structure more stable, and the molten globule state of CML is further stabilized. Indeed, substitutions of some residues in the α -domain of α -lactalbumin with more hydrophobic amino acids increase the stability of the molten globule state of α -lactalbumins (16, 18, 61).

Stability of D-helix

Residue 109 is located in the N terminal of the D-helix; these residues are lysine and valine in EML and CML, respectively (Figure 9, 10A and B). It is well known that a protein is stabilized by an

electrostatic interaction between a negatively charged side chain in a helical N terminal and a positively charged side chain in a helical C terminal due to the α -helix dipole (64-67). In the case of EML, these two residues, which are located in the helical N and C terminals of the D-helix, are both lysine (Lys-109 and 113). From the results of the X-ray structure of EML, the distance between the N^ϵ atom of Lys-109 and the N^ϵ atom of Lys-113 appears to be 3.08 Å. The pK_a value of lysine residue at 25° C is 10.5 (68), so both lysine residues (Lys-109 and 113) are protonated under this experimental condition (acidic pH; 4.5) of DSC (26), UV-CD spectra (Figure 4B) and 1H NMR spectra (25). Thus, these two positively charged side chains are thought to destabilize the D-helix due to electrostatic repulsion. However, this helical N-terminal residue of CML is valine, so that this α -helix dipole is thought to give more stability to the D-helix compared than that in EML. In addition, the distance between the C^β of Val-109 and the N^ϵ atom of Lys-113 is 5.58 Å, which might relieve the steric compression (Figure 10A).

Difference of amino acid sequences

Moreover, I have previously noted the differences between CML and EML in their amino acid sequences (Figure 9), and these differences should be responsible for the differences in the stability of the molten globule state (16, 18, 61, 69-72). CML has 23 amino acid substitutions in comparison with EML. Among these substitutions, that at residue 93 seems to be important in relation to difference in the stability of the molten globule state for the following reason. Residue 93 is located in the middle position of the C-helix (Figure 10A and B). Residue 93 is serine in EML, and alanine in CML. Alanine is thought to be more favorable for forming an α -helix than serine when the alanine residue is located in the central position of an α -helix (73).

Thus, it seems likely that such substitutions (from Gly-21 to His-21; from Lys-109 to Val-109; from Ser-93 to Ala-93) in the aromatic cluster around the C-helix and D-helix increase the

stability of the molten globule state of CML. Therefore, the molten globule state of CML is probably strongly protected by the aromatic cluster around the C-helix and D-helix as compared with EML. This is one of the main reasons why CML has an extraordinarily stable molten globule state as compared with other proteins in the lysozyme- α -lactalbumin superfamily.

In conclusion, the thermal unfolding equilibrium of CML is well represented by the three-state mechanism in which only the native, the intermediate, and unfolding states are populated. Both CD and NMR data indicate that the intermediate of this protein has secondary structure and elements of tertiary structure. Moreover, the results of ANS-binding experiment show the highest affinity of ANS to the intermediate. From these results, I interpret that the extraordinarily stable intermediate is a molten globule state. The molten globule state of this protein is extraordinarily stable compared with those of the other members of the lysozyme- α -lactalbumin superfamily, and this stabilization comes from the stronger protection around the C- and D-helix of the aromatic cluster region. This molten globule state may be more appropriate to define as native-like tertiary folded state. The structural and thermodynamically information will facilitate the study of protein folding.

References

1. Aune, K.C., Salahuddin, A., Zarlengo, M.H. & Tanford, C. (1967) *J. Biol. Chem.*, **242**, 4486-4489.
2. Tanford, C., Aune, K.C. & Ikai, I. (1973) *J. Mol. Biol.*, **73**, 185-197.
3. Ikai, A., Fish, W.W. & Tanford, C. (1973) *J. Mol. Biol.*, **73**, 165-184.
4. Kuwajima, K. (1989) *Proteins: Struct. Funct. Genet.*, **6**, 87-103.
5. Ptitsyn, O.B. (1992) *Protein Folding*, (ed. Creighton, T.E.) pp. 243-300 (Freeman, New York,

- 1992).
6. Ptitsyn, O.B. (1995) *Advan. Protein Chem.*, **47**, 83-229.
7. Kuwajima, K., Nitta, K., Yonayama, M. & Sugai, S. (1976) *J. Mol. Biol.*, **106**, 359-373.
8. Griko, Y.V., Privalov, P.L., Venyaminov, S.Y. & Kutysenko, V.P. (1988) *J. Mol. Biol.*, **202**, 127-138.
9. Kuwajima, K. (1996) *FASEB J.*, **10**, 102-109.
10. Dabora, J.M., Pelton, J.G. & Marqusee, S. (1996) *Biochemistry*, **35**, 11951-11958.
11. Tsui, V., Gracia, C., Cavagnero, S., Siuzdak, G., Dyson, H.J. & Wright, P.E. (1999) *Protein Sci.*, **8**, 45-49.
12. Peng, Z-yu & Kim, P.S. (1994) *Biochemistry*, **33**, 2136-2141.
13. Peng, Z-yu, Wu, L.C. & Kim, P.S. (1995) *Biochemistry*, **34**, 3248-3252.
14. Wu, L.C., Peng, Z-yu & Kim, P.S. (1995) *Nature. Struct. Biol.*, **2**, 281-286.
15. Schulman, B.A. & Kim, P.S. (1996) *Nature. Struct. Biol.*, **3**, 682-687.
16. Wu, L.C. & Kim, P.S. (1997) *Proc. Natl. Acad. Sci.*, **94**, 14314-14319.
17. Song, J., Bai, P., Luo, L. & Peng, Z-yu (1998) *J. Mol. Biol.*, **280**, 167-174.
18. Wu, L.C. & Kim, P.S. (1998) *J. Mol. Biol.*, **280**, 175-182.
19. Mizuguchi, M., Masaki, K. & Nitta, K. (1999) *J. Mol. Biol.*, **280**, 175-182.
20. Kay, M.S. & Baldwin, R.L. (1996) *Nature Struct. Biol.*, **3**, 439-445.
21. Eliezer, D., Yao, J., Dyson, H.J. & Wright, P.E. (1998) *Nature Struct. Biol.*, **5**, 148-155.
22. Fink, A.L., Oberg, K.A. & Seshadri, S. (1998) *Folding & Design*, **3**, 19-25.
23. Kay, M.S. & Ramos, C.H. & Baldwin, R.L. (1999) *Proc. Natl. Acad. Sci.*, **96**, 2007-2012.
24. Morozova, L.A., Haezebrouck, P. & Van Cauwelaert, F. (1991) *Biophys. Chem.*, **41**, 185-191.
25. Van Dael, H., Haezebrouck, P., Morozova, L.A., Arico-Muendel, C. & Dobson, C.M. (1993) *Biochemistry*, **32**, 11886-11894.
26. Griko, Y.V., Freire, E., Privalov, G., Van Dael, H. & Privalov, P.L. (1995) *J. Mol. Biol.*, **252**, 447-459.
27. Morozova, L.A., Haynie, D.T., Arico-Muendel, C., Van Dael, H. & Dobson, C.M. (1995) *Nature Struct. Biol.*, **2**, 871-875.
28. Morozova, L.A., Arico-Muendel, C., Haynie, D.T., Emelyanenko, V.I., Van Dael, H. & Dobson, C.M. (1997) *J. Mol. Biol.*, **268**, 903-921.
29. Grobler, J.A., Rao, K.R., Pervaiz, S. & Brew, K. (1994) *Arch. Biochem. Biophys.*, **313**, 360-366.
30. Kikuchi, M., Kawano, K. & Nitta, K. (1998) *Protein Sci.*, **7**, 2150-2155.
31. Koshiba, T., Hayashi, T., Ishido, M., Kumagai, I., Ikura, T., Kawano, K., Nitta, K. & Kuwajima, K. (1999) *Protein Eng.*, **12**, 429-435.
32. Nitta, K., Tsuge, H., Sugai, S. & Shimazaki, K. (1987) *FEBS Lett.*, **223**, 405-408.
33. Koshiba, T., Tsumoto, K., Masaki, K., Kawano, K., Nitta, K. & Kumagai, I. (1998) *Protein Eng.*, **11**, 683-690.
34. Morris, G.A. & Freeman, R. (1978) *J. Magn. Reson.*, **29**, 433-462.
35. Sakabe, N., Ikemizu, S., Sakabe, K., Higashi, T., Nakagawa, A., Watanabe, N., Adachi, S. & Sasaki, K. (1995) *Rev. Sci. Instrum.*, **66**, 1276-1281.
36. Otwinowski, Z. (1993) pp 80-86, SERC Daresbury Laboratory, Warrington, UK.
37. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr.*, **D50**, 760-763.
38. Navaza, J. (1994) *Acta Crystallogr.*, **A50**, 157-163.
39. Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S. & Miyano, M. (1992) *J. Biochem.*, **111**, 141-143.
40. Brünger, A.T. (1993) X-PLOR :Version 3.1. A system for X-ray Crystallography and NMR, Yale University Press, New Haven, CT, USA.
41. Desmet, J., Van Dael, H., Van Cauwelaert, F., Nitta, K. & Sugai, S. (1989) *J. Inorg. Biochem.*, **37**,

42. Yutani, K., Ogasahara, K. & Kuwajima, K. (1992) *J. Mol. Biol.*, **228**, 347-350.
43. Griko, Y., Freire, E. & Privalov, P.L. (1994) *Biochemistry*, **33**, 1889-1899.
44. Pfeil, W. (1998) *Proteins: Struct. Funct. Genet.*, **30**, 43-48.
45. Griko, Y. & Remeta, D.P. (1999) *Protein Sci.*, **8**, 554-561.
46. Koshiba, T., Nitta, K., Kobashigawa, Y. & Kuwajima, K. (2000) to be published.
47. Stryer, L. (1965) *J. Mol. Biol.*, **13**, 482-495.
48. Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas, A.F. & Gilmanshin, R.I. (1991) *Biopolymers*, **31**, 119-128.
49. Demarest, S.J., Fairman, R. & Raleigh, D.P. (1998) *J. Mol. Biol.*, **283**, 279-291.
50. Semisotnov, G.V., Rodionova, N.A., Kutysenko, V.P., Ebert, B., Blanck, J. & Ptitsyn, O.B. (1987) *FEBS Letters*, **224**, 9-13.
51. Kuwajima, K., Harushima, Y. & Sugai, S. (1986) *Int. J. Peptide Protein Res.*, **27**, 18-27.
52. Baum, J., Dobson, C.M., Evans, P.A. & Hanley, C. (1989) *Biochemistry*, **28**, 7-13.
53. Nitta, K., Tsuge, H. & Iwamoto, H. (1993) *Int. J. Peptide Protein Res.*, **41**, 118-123.
54. Artymiuk, P.J. & Blake, C.C.F. (1981) *J. Mol. Biol.*, **152**, 737-762.
55. Stuart, D.I., Acharya, K.R., Walker, N.P.C., Smith, S.G., Lewis, M. & Phillips, D.C. (1986) *Nature*, **324**, 84-87.
56. Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M. & Phillips, D.C. (1989) *J. Mol. Biol.*, **208**, 99-127.
57. Acharya, K.R., Ren, J., Stuart, D.I., Phillips, D.C. & Fenna, R.E. (1991) *J. Mol. Biol.*, **221**, 571-581.
58. Pike, A.C.W., Brew, K. & Acharya, K.R. (1996) *Structure*, **4**, 691-703.
59. Guss, J.M., Messer, M., Costello, M., Hardy, K. & Kumar, V. (1997) *Acta Cryst.*, **D53**, 355-363.

60. Mizuguchi, M., Arai, M., Yue, K., Nitta, K. & Kuwajima, K. (1998) *J. Mol. Biol.*, **283**, 262-277.
61. Uchiyama, H., Perez-Prat, E.M., Watanabe, K., Kumagai, I. & Kuwajima, K. (1995) *Protein Eng.*, **8**, 1153-1161.
62. Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J. & Dobson, C.M. (1993) *Biochemistry*, **32**, 1707-1718.
63. Dobson, C.M. (1995) Finding the right fold. *Nature Struct. Biol.*, **2**, 513-517.
64. Blagdon, D.E. & Goodman, M. (1975) *Biopolymers*, **14**, 241-245.
65. Wada, A. (1976) *Adv. Biophys.*, **9**, 1-63.
66. Hol, W.G.J., van Duijnen, P.T. & Berendes, H.J.C. (1978) *Nature*, **273**, 443-446.
67. Nicholson, N., Becktel, W.J. & Matthews, B.W. (1988) *Nature*, **336**, 651-656.
68. Tanford, C. (1962) *Adv. Protein. Chem.*, **17**, 69-165.
69. Hughson, F.M., Barrick, D. & Baldwin, R.L. (1991) *Biochemistry*, **30**, 4113-4118.
70. Kiefhaber, T. & Baldwin, R.L. (1995) *J. Mol. Biol.*, **252**, 122-132.
71. Marmorino, J.L. & Pielak, G.J. (1995) *Biochemistry*, **34**, 3140-3143.
72. Marmorino, J.L., Lehti, M. & Pielak, G.J. (1998) *J. Mol. Biol.*, **275**, 379-388.
73. Merutka, G., Lipton, W., Shalongo, W., Park, S.H. & Stellwagen, E. (1990) *Biochemistry*, **29**, 7511-7515.
74. Kraulis, P.J. (1991) *J. Appl. Cryst.*, **24**, 946-950.
75. McKenzie, H.A. & Shaw, D.C. (1985) *Biochem. Int.*, **10**, 23-31.
76. Brew, K., Vanaman, T.C. & Hill, R.L. (1967) *J. Biol. Chem.*, **242**, 3747-3749.
77. Brew, K. (1972) *Eur. J. Biochem.*, **27**, 341-353.
78. Shewale, J.G., Sinha, S.K. & Brew, K. (1984) *J. Biol. Chem.*, **259**, 4947-4956.
79. Hall, L., Craig, R., Edbrooke, M.R. & Campbell, P.N. (1982) *Nucl. Acids Res.*, **10**, 3503-3515.

80. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986) *Biochemistry*, **25**, 6965-6972.

Table 1. Thermodynamic Parameters for the Unfolding of CML and EML in the Absence of Ca^{2+} at pH 4.5

Lysozyme	Type of transition ^a	T_m	ΔH^b	$T\Delta S^b$	ΔG^b	ΔC_p
		°C	kcal/mol	kcal/mol	kcal/mol	kcal/mol/K
at 41.5° C						
EML ^c	N-I	41.5	36.6	36.6	0.0	1.20
CML	N-I	39.0	44.6	44.9	-0.3	1.03 ^d
at 66.4° C						
EML ^c	I-U	66.4	29.5	29.5	0.0	0.61
CML	I-U	86.9	30.9	28.8	2.1	0.70 ^d

^aN, I, and U mean native, intermediate, and unfolded states, respectively. ^bThese values are extrapolated to 41.5 or 66.4° C. ^cFrom Griko *et al.* (26). ^dFrom Koshiba *et al.*, (46).

Table 2. Crystallographic Data, X-ray Processing Statistics, and Refinement Statistics of the apo-CML

<i>Crystal data collection</i>	
Space group	<i>P</i> 6 ₁ 22
Cell constants (Å)	
<i>a</i>	68.6
<i>b</i>	68.6
<i>c</i>	107.1
Resolution (Å)	100-1.85
No. of measured reflections	199,512
No. of independent reflections	13,034
<i>R</i> merge (%) ^a	6.00
<i>Refinement</i>	
Resolution used (Å)	8.0-1.85
No. of used reflections (<i>F</i> > 2σ)	12,679
Data completeness (%)	95.5
Δ bond length (Å)	0.007
Δ bond angle (°)	1.414
Average <i>B</i> -factor (Å ²)	
Main-chain	23.8
Side-chain	28.3
<i>R</i> -factor (%) ^b	17.8
Free <i>R</i> -factor (%) ^c	21.5

^a $R_{\text{merge}} = \Sigma |I - \{I\}| / \Sigma \{I\} \times 100\%$.

^b $R\text{-factor} = \Sigma |F_o| - |F_c| / \Sigma |F_c| \times 100\%$.

^cFree *R*-factor is the same as *R*-factor, except for a 10% subset of all reflections.

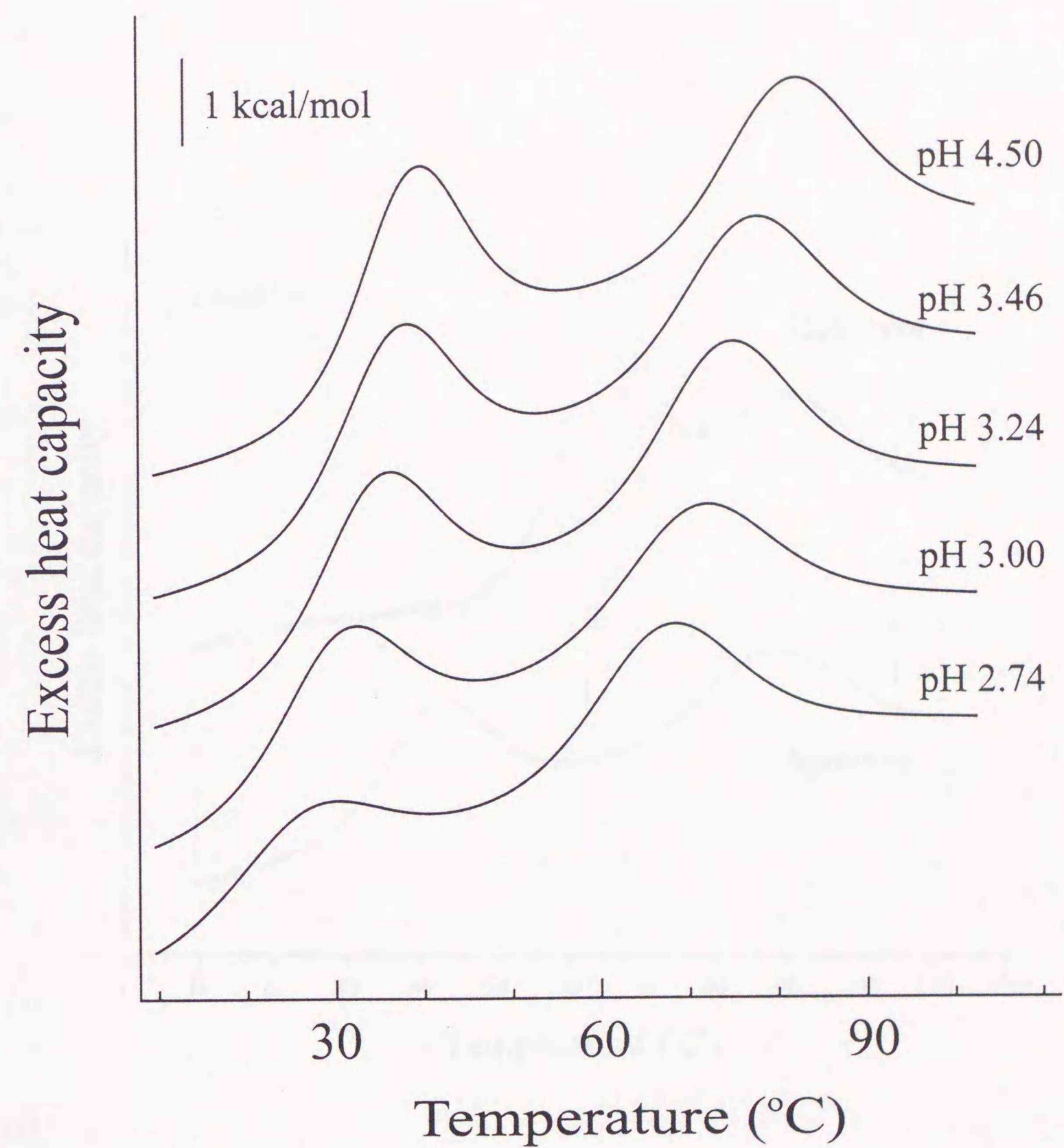


Figure 1. The thermal unfolding transition of CML in the absence of calcium ions at various pHs (2.74, 3.00, 3.24, 3.48, and 4.50). The unfolding transition is expressed by excess heat capacity as a function of temperature.

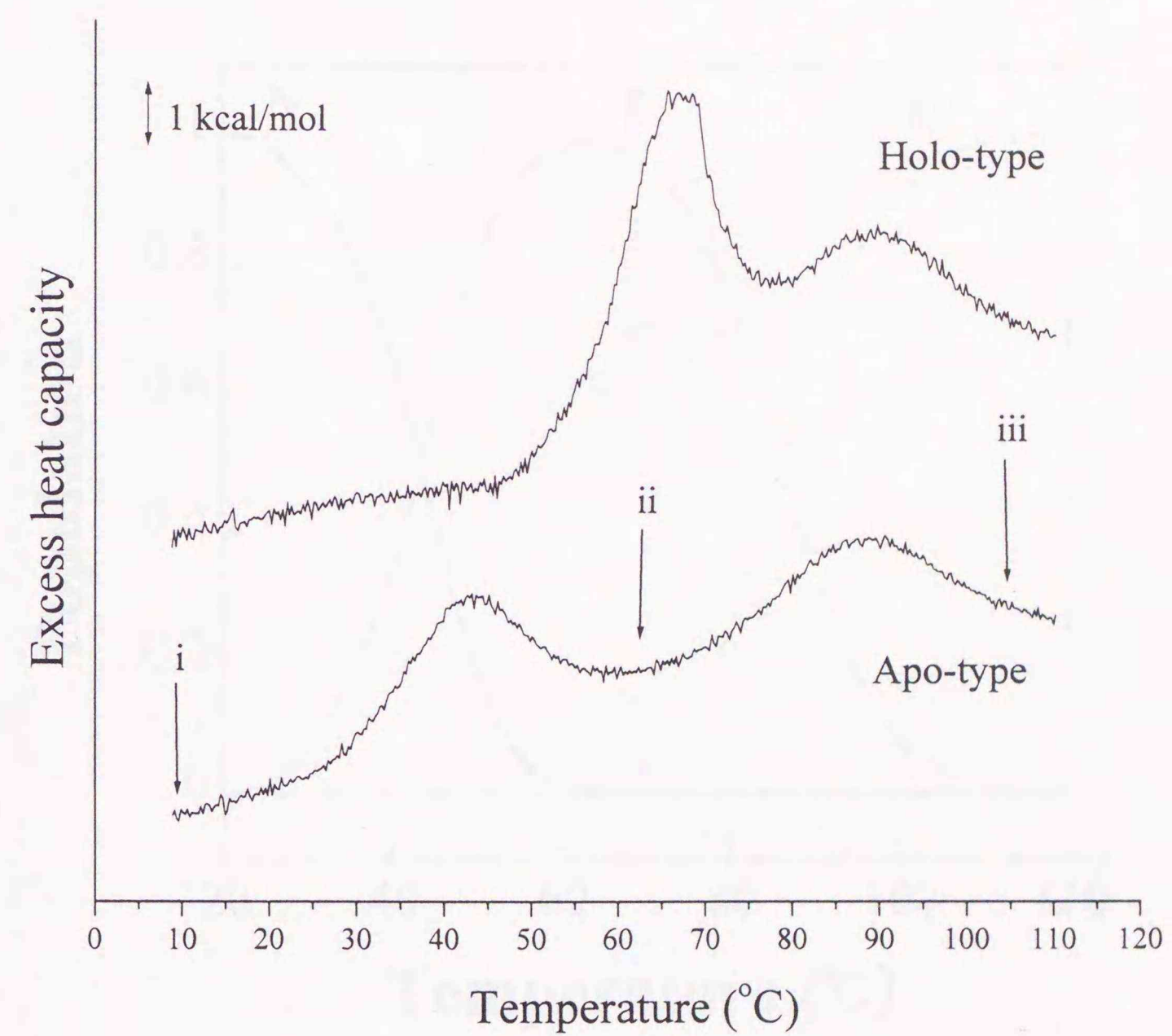


Figure 2. The thermal unfolding transition of CML in the absence (apo-type) and presence of 10 mM CaCl_2 (holo-type) at pH 4.5. The unfolding transition is expressed by excess heat capacity as a function of temperature.

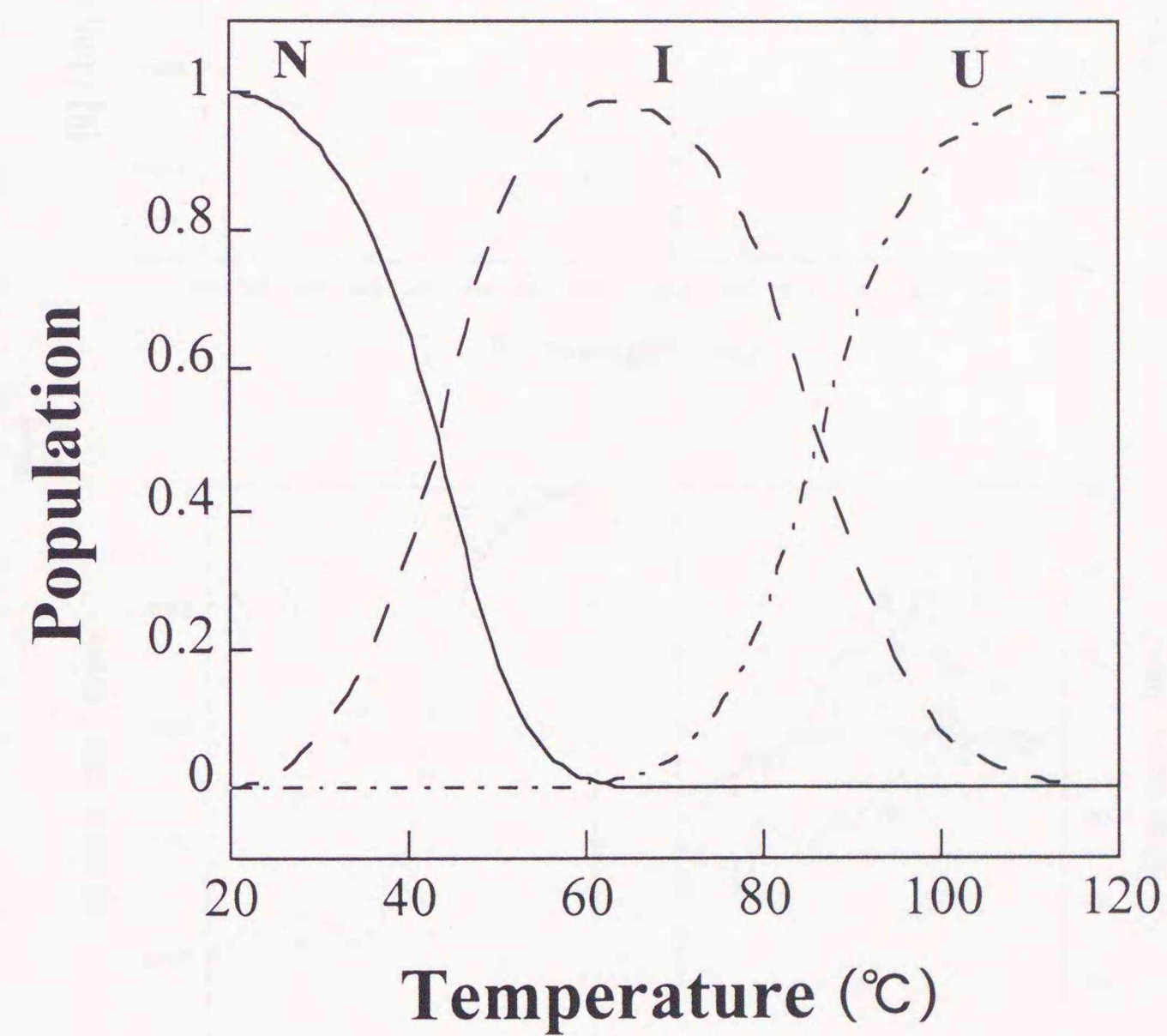


Figure 3. Temperature dependence of the populations of the three states, the native (N), folding intermediate (I) and unfolded (U) states in the absence (apo-type) of CaCl_2 at pH 4.5.

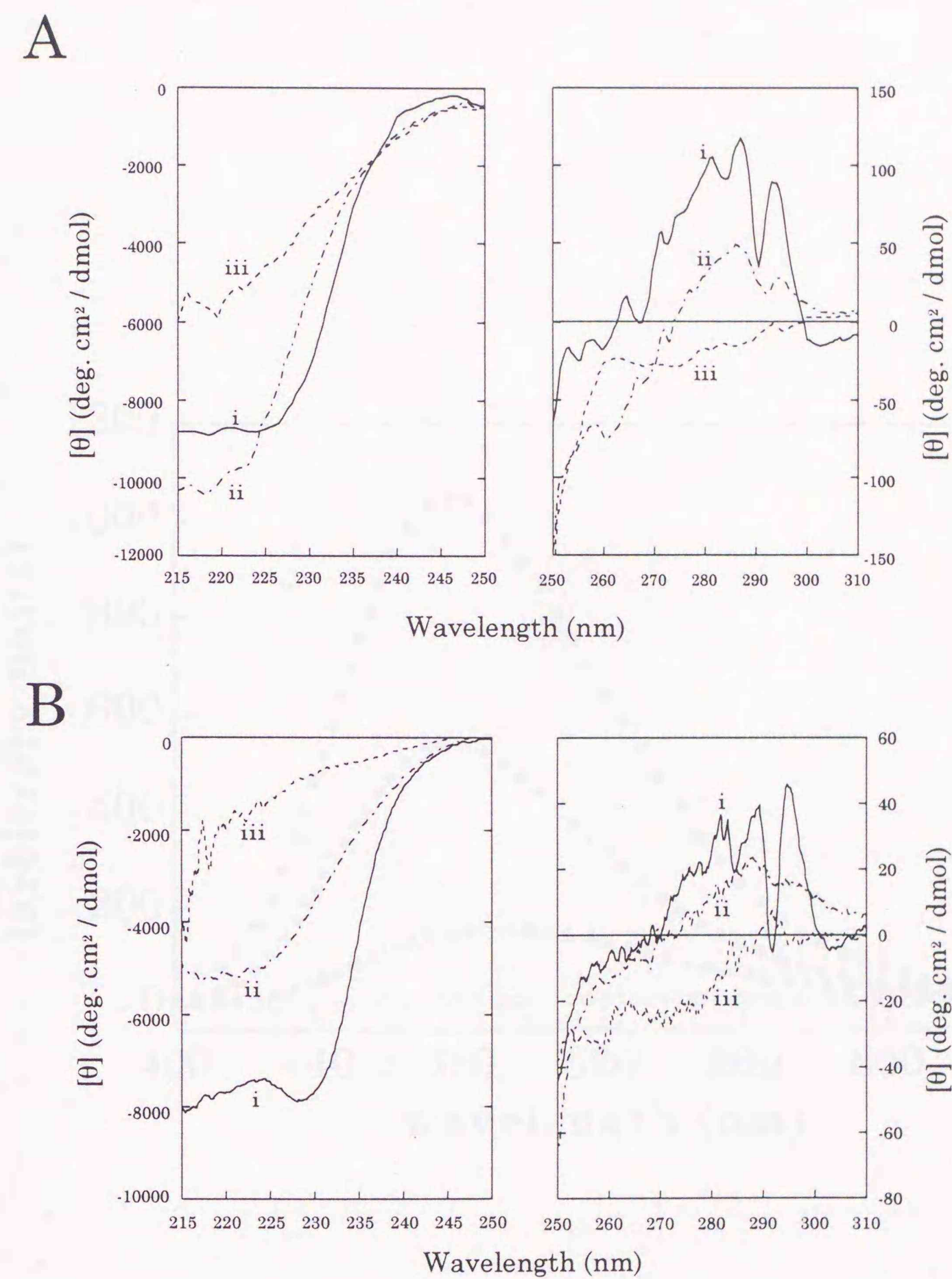


Figure 4. CD spectra of (A) CML and (B) EML in the far-UV (left) and near-UV (right) regions at pH 4.5. Three spectra (i, ii, and iii) were measured at the different temperatures indicated by arrows in Figure 2.

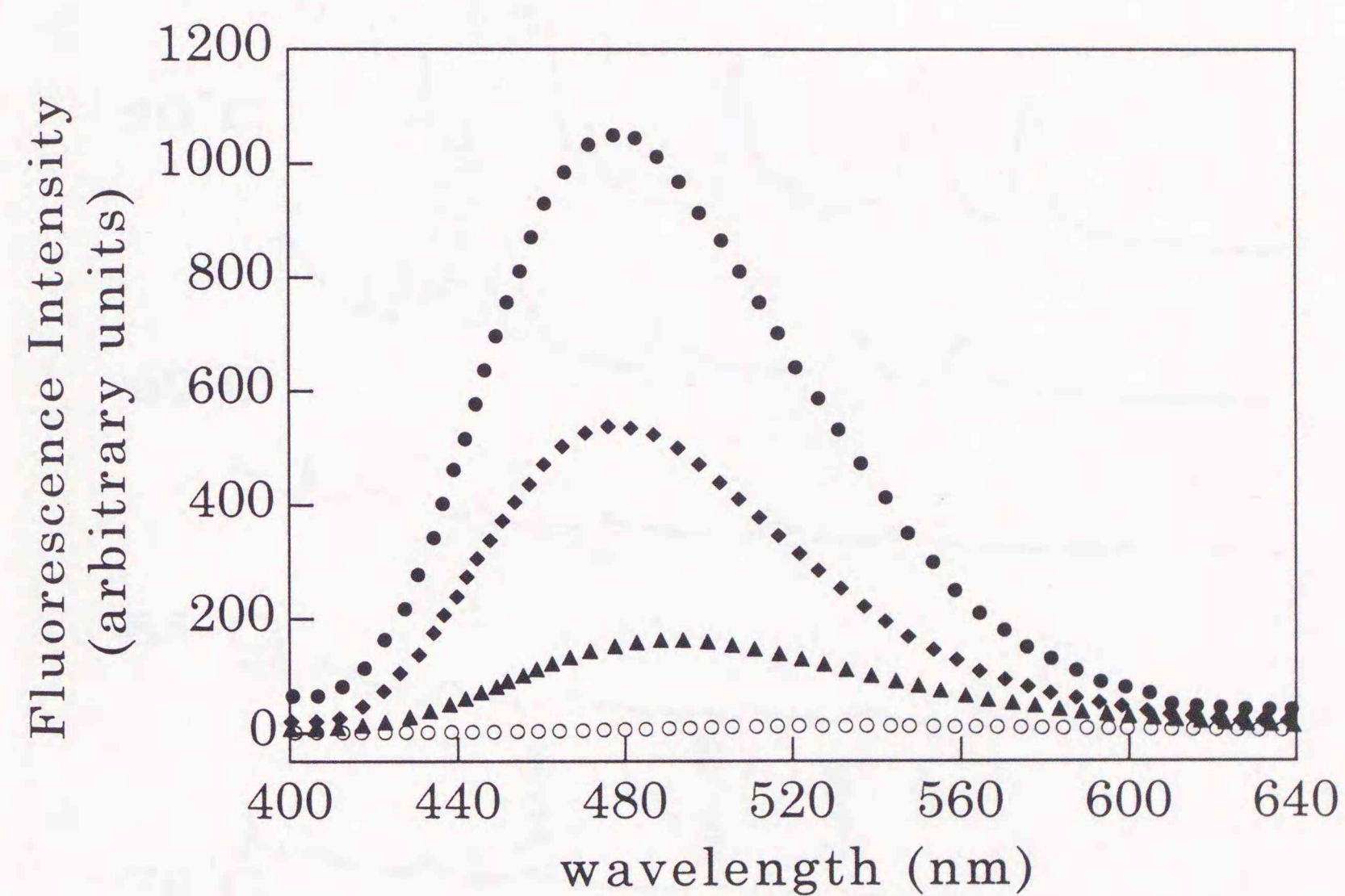


Figure 5. Fluorescence spectra of 50 μM ANS in the absence (open symbols) and presence (filled symbols) of 25 μM apo-CML at pH 4.5. The spectra of the complex of ANS with apo-CML were obtained at 9° C (diamonds), 63° C (circles), and 98° C (triangles). The spectrum of ANS without lysozyme were obtained at 25° C. These three temperatures are indicated in Figure 2 with arrows.

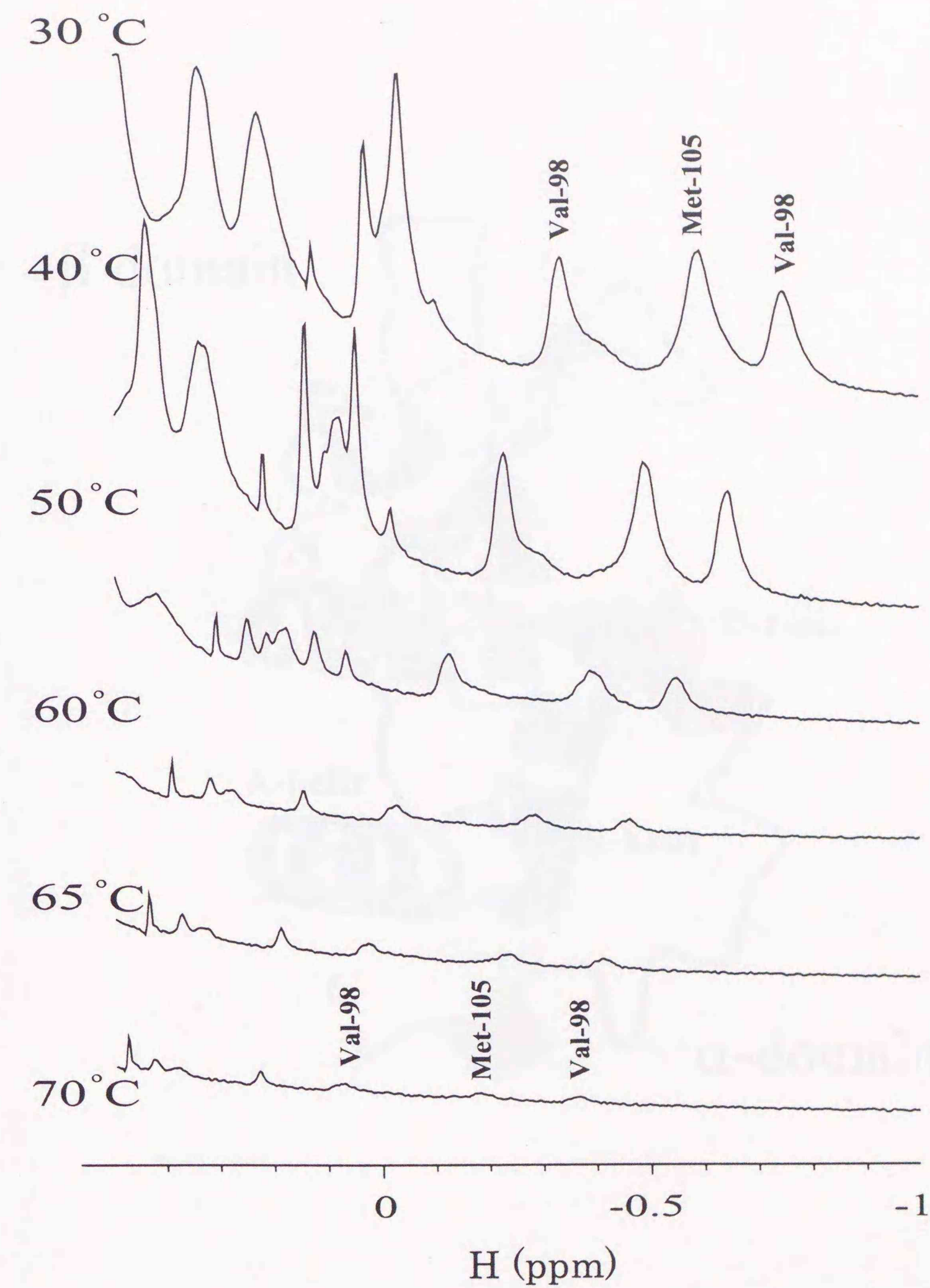


Figure 6. One-dimensional ^1H NMR spectra (400 MHz) of apo-CML as a function of temperature at pH 4.5.

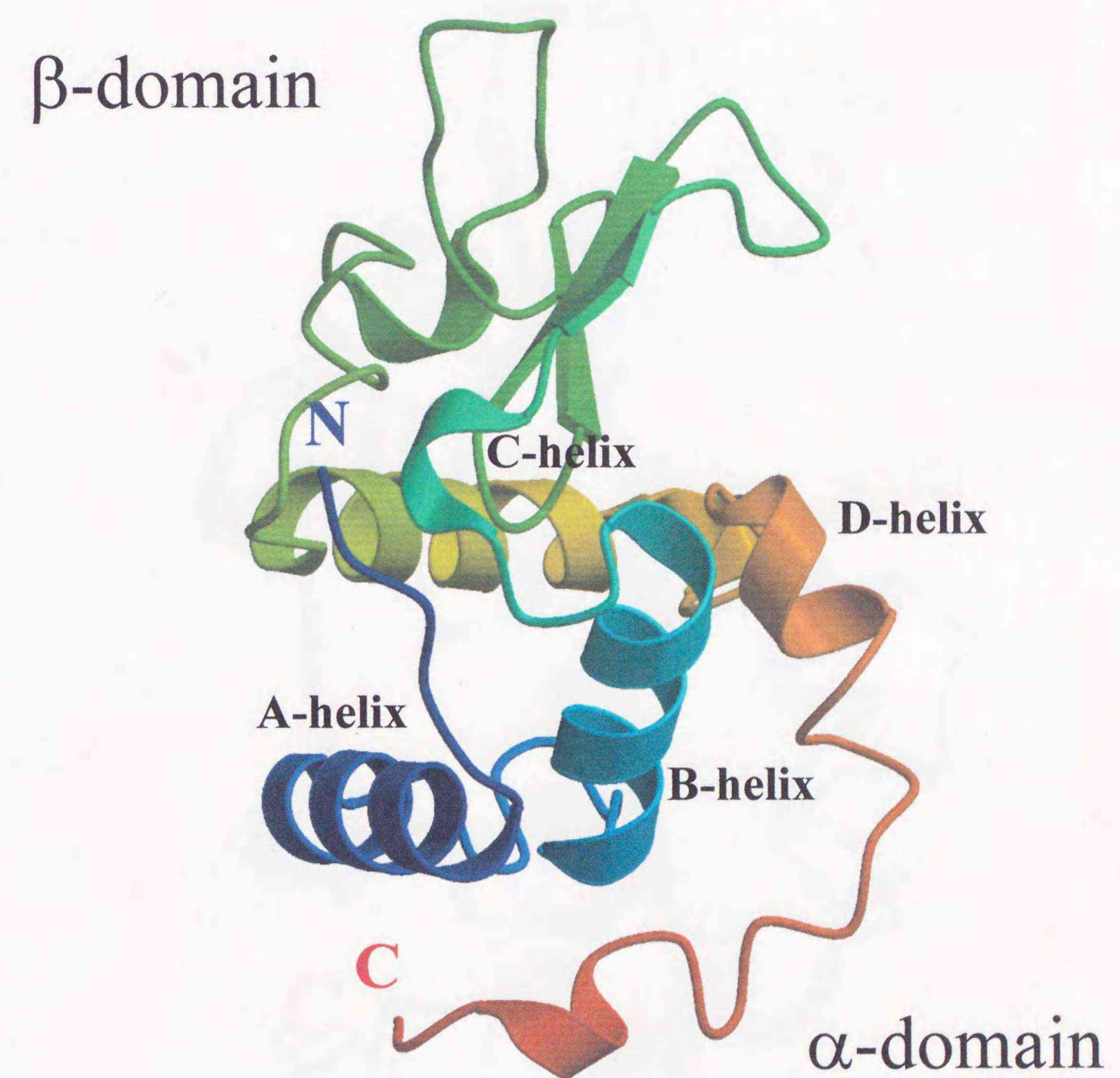


Figure 7. Ribbon model of CML in which α -helices are sequentially labelled from A to D. The structure shows a rainbow color (red to blue) from the N-terminus to the C-terminus (Figure produced using MOLSCRIPT (74)).

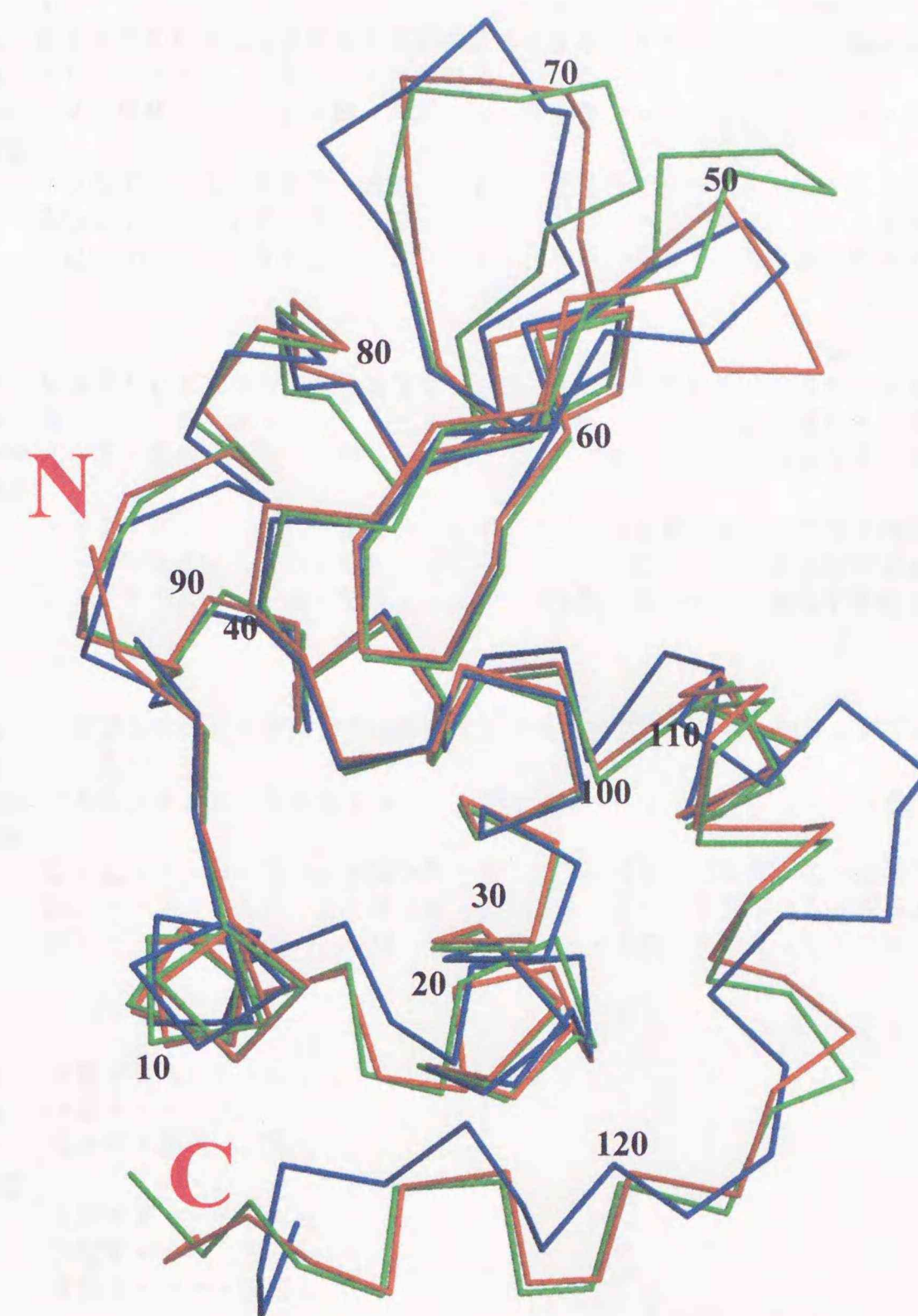


Figure 8. Superimposition of the structures of CML, EML, and GPLA. A structure of the C α traces of CML (red), EML (green) and GPLA (blue) are shown after the least-squares superposition. Every tenth residue of CML is labelled (Figure produced using MOLSCRIPT (74)).

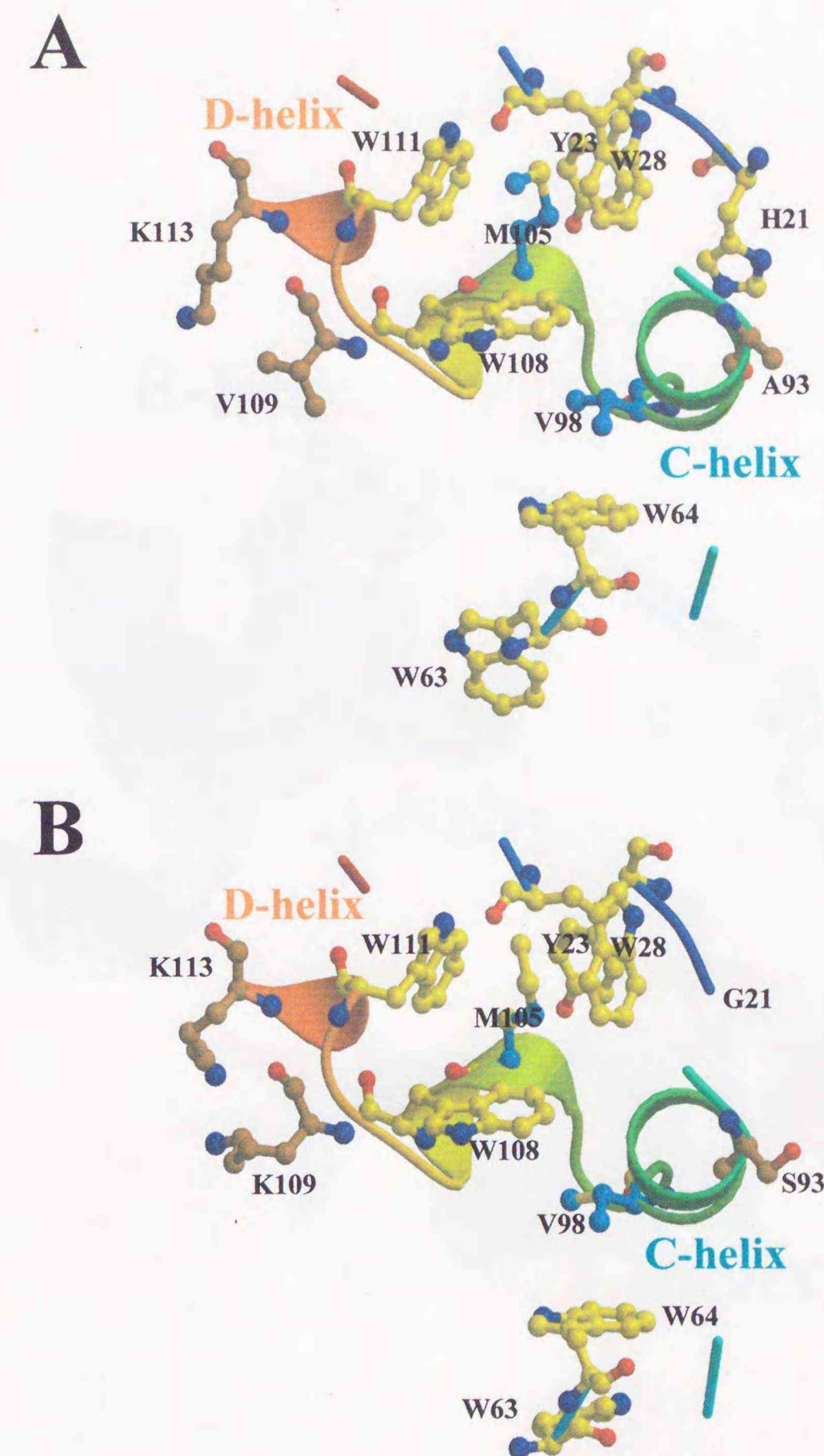


Figure 10. Conformation of the aromatic cluster around the residues Val-98 and Met-105 within 10 Å in distances from their C_{α} atoms. The local structures are shown schematically for (A) CML and (B) EML. Two molecules are viewed in the same orientation (Figure produced using MOLSCRIPT (74)).

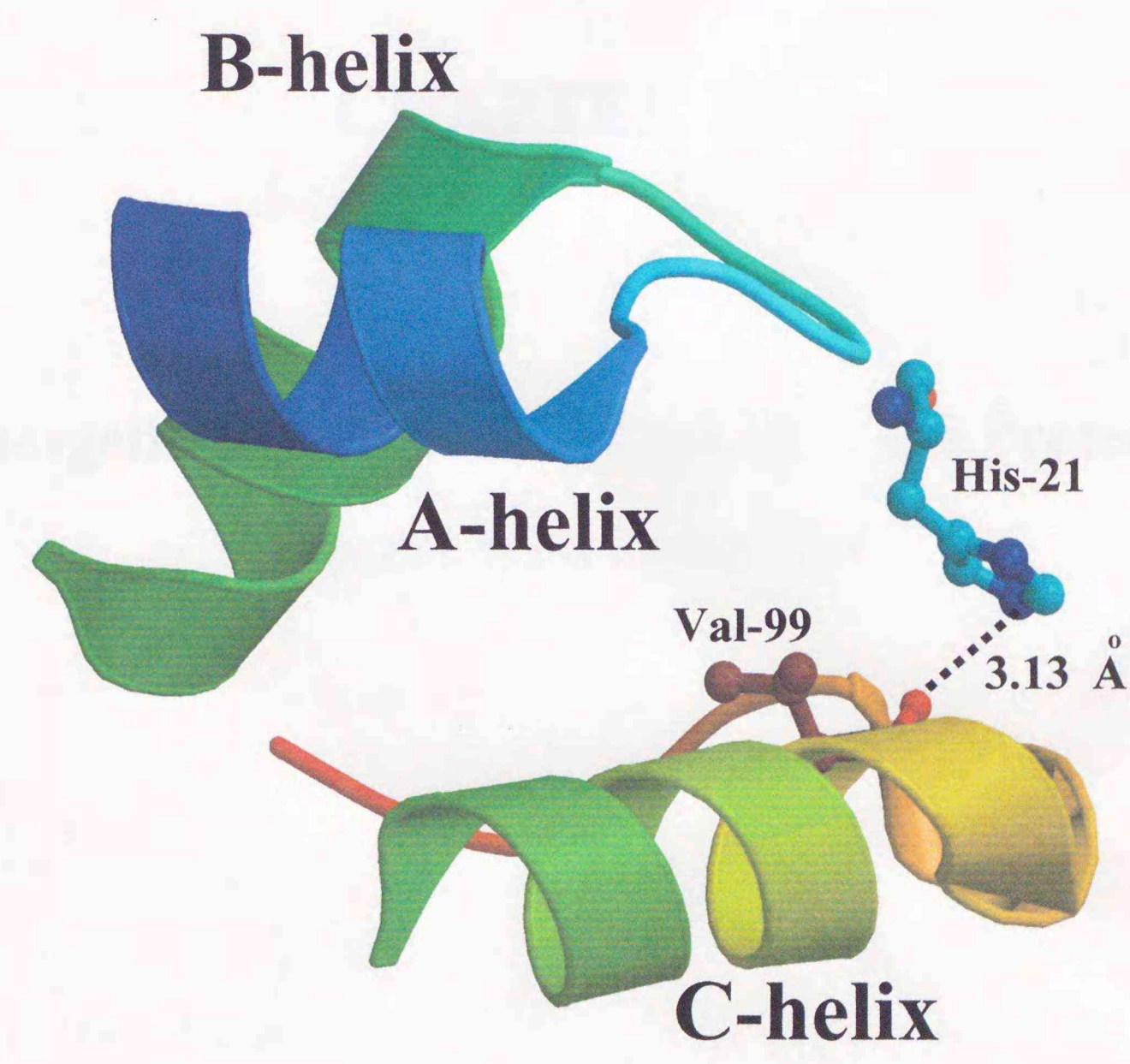
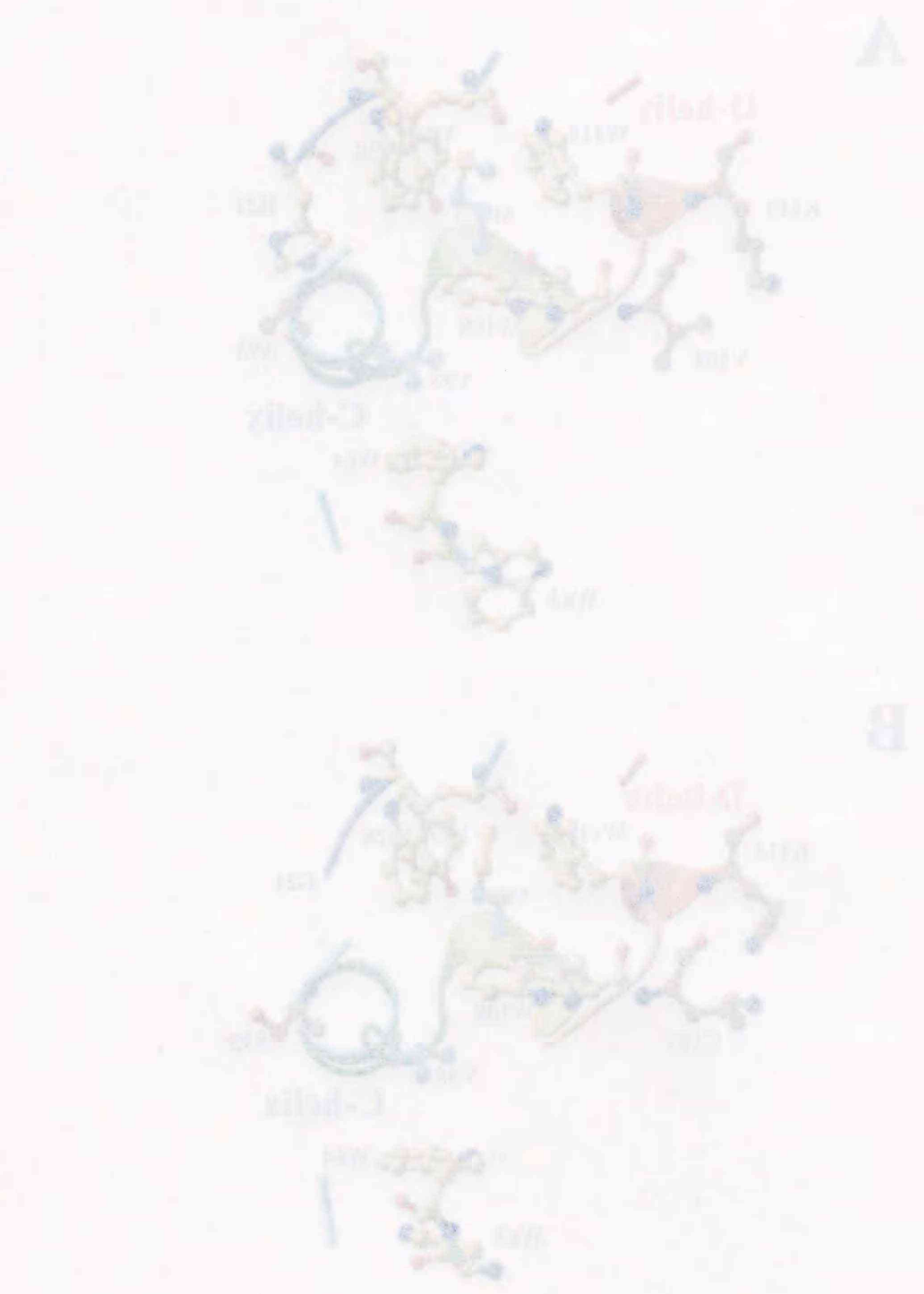


Figure 11. Partial structure of CML around the region of residue His-21. The broken line indicates hydrogen bond (Figure produced using MOLSCRIPT (74)).

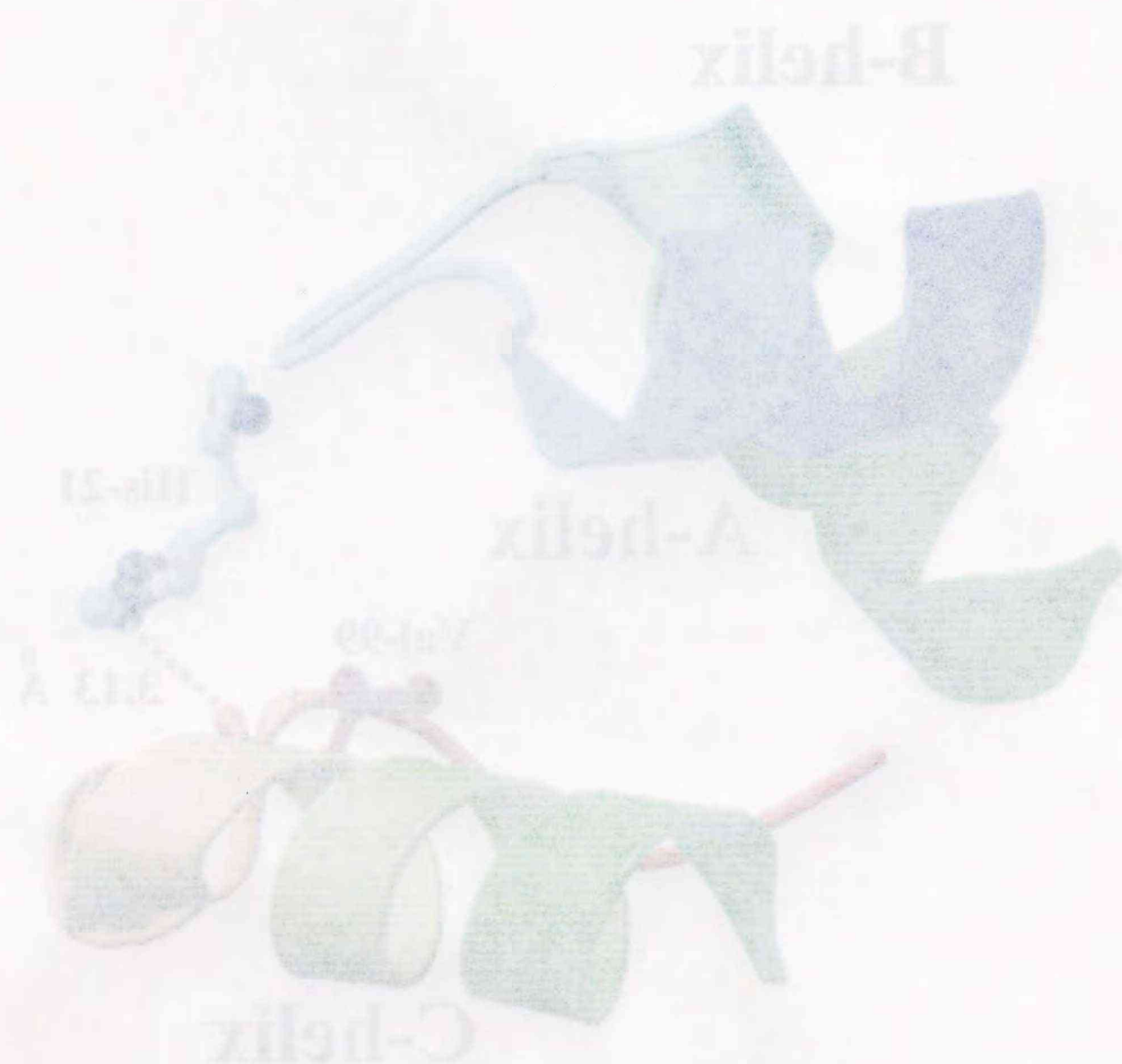


Figure 11-10. Ribbon diagram of the canine milk lysozyme structure. The structure is shown in a ribbon representation, with the three alpha-helices labeled A, B, and C. The protein is colored in shades of blue, green, and orange.

CHAPTER III

Energetics of Three-state Unfolded of a Protein: Canine Milk Lysozyme

Abstract

I studied the thermodynamics of unfolding transitions of a Ca^{2+} -binding lysozyme, canine milk lysozyme (CML), obtained measurement by differential scanning calorimetry, and compared the findings with those for sequentially and structurally homologous proteins, human α -lactalbumin (α -hLA) and equine milk lysozyme (EML). The results showed that CML and EML exhibit two clear heat absorption peaks in the absence of Ca^{2+} ions (apo-form), although the cooperative thermal transition of α -hLA is apparently absent in this form. The first peak represents the unfolding transition from the native to the molten globule state (N-A transition), and the second peak represents that from the molten globule to the thermally unfolded state (A-U transition). It is suggesting that the molten globule state of CML and EML is a thermodynamic state, while that of the α -hLA is thermodynamically indistinguishable from the unfolded state. Moreover, CML also exhibits two clear heat absorption peaks in the presence of Ca^{2+} ions (holo-form), although α -hLA and EML show only a single heat absorption peak in the holo-form. It is shown that extraordinary stability of the molten globule state of CML led to the appearance of the second heat absorption peak in the holo-form. The obtained thermodynamic parameters were discussed in light of the results of CD spectroscopy and hydrogen-exchange experiments. I interpret that the extraordinary stability of the molten globule state of CML is due to the native-like packing interaction in the state, and it comes from the highly persistent structures existing not only within the α -domain but also within the β -domain. The structural and thermodynamics differences between the molten globule state of α -lactalbumins and that of Ca^{2+} -binding lysozymes are further discussed.

Introduction

One of the major goals of structural biology is to understand the protein folding mechanism, and clarification of the concept of a folding intermediate, such as a molten globule state, is an important part of reaching that goal (1). The molten globule state has been thought to be an intermediate state between the native and unfolded states in many globular proteins (2-6). This state is characterized by (i) the presence of a substantial secondary structure; (ii) the absence, in large part, of a specific tertiary structure; (iii) a compact shape; (iv) the formation of a hydrophobic core when the protein is exposed to solvents; and (v) the lack of rigid side-chain packing interactions (7-10). However, with the extensive progress in studying the molten globule state, some of the new evidences have been accumulated about the state. There are a great variety of the folding intermediates, which are including almost fully unfolded state and native-like conformation (11-19).

Canine milk lysozyme (CML), which belongs to the family of Ca^{2+} -binding lysozymes (20-22), and α -lactalbumin have a similar sequence and a very similar three-dimensional structure characterized by the presence of a deep cleft that divides the molecule into two domains (the α - and β -domains) (Figure 1) (23-29). Like α -lactalbumin, CML and equine milk lysozyme (EML) exhibit a stable molten globule state at equilibrium (21, 22, 29-37), although most of the conventional lysozymes usually unfold in a cooperative two-state manner without the accumulation of the molten globule state. Despite the fact that the molten globule states of CML and EML are structurally very similar to that of α -lactalbumin, the molten globule states of CML and EML are much more native-like than that of α -lactalbumin; they are thought to be stabilized by non-specific interactions and by specific side-chain packing interactions (29, 31-33, 36, 37). Thus, although the considerable attentions have been paid to the molten globule states of CML, EML, and α -lactalbumins by using nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopies, there are still lacking the energetics information of the molten globules. For

interpretation of the protein folding mechanism, the thermodynamic characterization of the molten globule state should provide critical information.

In this CHAPTER, I investigated the thermodynamics of unfolding transitions of CML by differential scanning calorimetry (DSC) measurement, and compared these values with those obtained for homologous proteins, human α -lactalbumin (α -hLA) and EML. The results represent that the molten globule state of CML and EML is a thermodynamic state, while that of the α -hLA is thermodynamically indistinguishable from the unfolded state. It may be generally valid that the native-like packing interaction in the molten globule state induces a remarkably positive heat absorption peak with thermal unfolding.

Materials and Methods

Materials

All enzymes used for genetic engineering were purchased from Takara Shuzo (Kyoto, Japan) or Toyobo (Osaka, Japan). Oligonucleotide DNA primers were synthesized by Kurabo (Osaka, Japan). Human α -LA (α -hLA) and equine milk lysozyme (EML) have been prepared previously (36, 38). All other reagents were of biochemical research grade.

Site-directed mutagenesis of CML

Plasmid pHK7CML, which contains T7 transcription and translation signals and DNA encoding the chemically synthesized canine milk lysozyme sequence, was used as a starting material for mutagenesis (22). To obtain the mutant CMLs (H21G, I56L, A93S, and V109K), His-21, Ile-56, Ala-93, or Val-109 in CML were replaced by glycine, leucine, serine, and lysine, respectively. Site-directed mutagenesis was performed as described previously (39). The mutations were confirmed by DNA sequencing.

Expression and refolding of wild type and mutant CMLs

Overexpression of wild type and mutant CMLs in *E. coli*, as inclusion bodies, and refolding procedures using *E. coli* thioredoxin were performed as described previously (22, 39).

Preparation of apo- and holo- proteins

Apo- and holo-proteins were prepared as described previously (22, 29).

Differential scanning calorimetry (DSC)

Differential scanning calorimetry (DSC) measurements of CML, α -hLA, and EML were carried out with an MC-2 microcalorimeter (MicroCal, Inc., Northampton, MA) at the scan rate of 1.0 K/min. Sample solutions for DSC measurements between pH 6.5 and 7.0 were prepared by dissolving the proteins in 50 mM sodium cacodylate buffer; solutions between pH 4.0 and 4.5 were prepared using 50 mM sodium acetate buffer; and between pH 2.0 and 4.0 were prepared using 50 mM glycine hydrochloride. The protein concentrations used were 1.2-1.8 mg/mL. The pH of the sample solution was confirmed before and after each measurement. To expand the temperature range, all DSC measurements were performed under an excess pressure of 1.5 atm.

Circular dichroism (CD) spectra

The CD spectra of CML, α -hLA, and EML were measured on a Jasco J-720 spectropolarimeter (Hitachi, Tokyo, Japan). Far- and near-UV CD spectra were measured using a cell with 0.1 and 1 cm optical path lengths, respectively. The protein concentrations were 0.5 mg/mL for all measurement.

Structure analysis

The atomic coordinates of the proteins analyzed here were obtained from the Protein Data Bank [CML (29), α -hLA (26), and EML (40)]. The hydrogen exchange rate data of CML, α -hLA, and EML at pH 2.0 (5 and 25° C) were obtained from previously reports (31, 32, 37, 41). Molecular graphics were produced using WebLab Viewer (Molecular Simulations Inc.) and MOLSCRIPT (42).

Estimation of protein concentration

The protein concentrations were estimated spectrophotometrically using the following extinction coefficient at 280 nm: $E_{1\text{ cm}}^{1\%} = 23.2$ for wild type and its mutants of CML (21), $E_{1\text{ cm}}^{1\%} = 18.4$ for α -hLA (43), and $E_{1\text{ cm}}^{1\%} = 23.5$ for EML (36).

Results

CD spectra of CML, α -hLA, and EML

Figure 2 shows the CD spectra of CML, α -hLA, and EML in the native (N), the molten globule (A), and the unfolded (U) states, respectively. The N state of all proteins shows characteristic CD bands in the far- and near-UV regions, indicating the presence of the secondary as well as the specific tertiary structures. These characteristics of the native CD spectra are completely lost in 6 M guanidine hydrochloride, where all proteins are in the fully unfolded states.

As shown in Figure 2C, the far-UV CD spectrum of the molten globule state (pH 2.0 and 25° C) of α -hLA shows significant negative ellipticities and the appearance of the negative shoulder at 223 nm, while the near-UV CD spectrum shows that all of the signals are lost. This clearly indicates that the molten globule state of α -hLA has a substantial secondary structure but lacks extensive rigid packing interactions of the aromatic side chains (7, 8, 44). However, CD spectra of CML and EML at pH 2.0 show enhancement of the positive Cotton effects around 288 nm and

disappearance of the troughs observed in the native proteins, indicating a conformational rearrangement around the aromatic residues in the molten globule state (Figure 2A & B). Further, the far-UV CD spectra of the two proteins show enhancement of the negative ellipticity below 227 nm, indicating the presence of a significant secondary structure in the molten globule state. Thus, the molten globule states of CML and EML are much more native-like than that of α -hLA.

Thermal unfolding of CML, α -hLA, and EML in the absence of Ca^{2+} ions

An important question regarding the molten globule state is whether it is a thermodynamic state. To address this question, DSC measurements of the CML, α -hLA, and EML were carried out at acidic pH between 2.7 and 4.5 in the absence of Ca^{2+} ions (apo-form). In this pH region, the thermal unfolding of three proteins were reversible. It was confirmed by reheating the protein solution in the calorimeter cell immediately after cooling from the first scan. Repeated thermal scans of the samples exhibited similar transitions. As shown in Figure 3A (top of the Figure), the cooperative thermal transition is apparently absent in the molten globule state of α -hLA, suggesting that this state is thermodynamically indistinguishable from the unfolded state. This could mean that there are no energy barriers between the molten globule and unfolded states, as previously reported in bovine α -LA (45–47).

However, CML and EML exhibit a cooperative thermal unfolding with two remarkably positive heat absorption peaks (Figure 3A; middle and bottom of the Figure). It is indicating that the first peak represents the unfolding transition from the native to the molten globule state, and the second peak represents the unfolding transition from the molten globule to the thermally unfolded state. Although similar behaviors have been observed in CML and EML, the transition from the molten globule to the thermally unfolded state of CML occurred at a higher temperature and the ΔH_{A-U} was larger. Among the α -LA and lysozyme family of proteins studied thus far (29, 34, 35,

45–48), CML exhibits the most stable molten globule state with the largest ΔH_{A-U} .

Thermal unfolding of CML, α -hLA, and EML in the presence of Ca^{2+} ions

The temperature dependence of the partial heat capacity of CML, α -hLA, and EML in the presence of 10 mM CaCl_2 is shown in Figure 3B. The thermal unfolding of holo-proteins was also reversible. Under this condition, α -hLA and EML show a single heat absorption peak caused by a cooperative transition from the native to the thermally unfolded state (Figure 3B, top & middle). In the case of CML, however, two clear heat absorption peaks have been observed in the holo-form (Figure 3B, bottom). Therefore, the comparison of the heat absorption peaks in the apo- and holo-forms of CML also indicates that the extraordinary stability of the molten globule state of this protein leads to the appearance of the second heat absorption peak even in the holo-protein.

Analysis of the excess heat capacity function of CML

To fully understand the stability of a protein, it is necessary to elucidate the thermodynamic parameters of the thermal unfolding. In order to estimate the thermodynamic parameters such as melting temperature (T_m), enthalpy change (ΔH), entropy change (ΔS), or Gibbs free energy change (ΔG) of a three-state unfolding protein, the heat capacity changes of the protein in a thermodynamic state must first be assumed.

The partial molar heat capacity of a macromolecule under conformational equilibrium, $C_{p,obs}$, is given by the following equation (35):

$$C_{p,obs} = C_{pN} + \{ \langle \Delta H^2 \rangle - \langle \Delta H \rangle^2 \} / RT^2 + \sum f_i \Delta C_{pi} \quad (1)$$

where C_{pN} is the partial molar heat capacity of the native state. The f_i represents the population of

the molecules in the i th state excluding the native state, and ΔC_{pi} is the relative heat capacity of the i th state compared to that of the native state. The second term on the right-hand side of equation (1) represents the fluctuation of conformation of the macromolecule (35), and the third term accounts for the heat capacities of conformers other than the native state.

In this study, the thermal unfolding profile is that of a three-state reversible unfolding (N, A, and U). Therefore, equation (1) can be rewritten as follows:

$$C_{p,obs} = C_{pN} + \{f_A \Delta H_{N-A}^2 + f_U \Delta H_{N-U}^2 - (f_A \Delta H_{N-A} + f_U \Delta H_{N-U})^2\} / RT^2 + f_A \Delta C_{pN-A} + f_U \Delta C_{pN-U} \quad (2)$$

The C_{pN} and the ΔC_{pN-U} of HEL are temperature-dependent, and are represented as a linear and a quadratic function of temperature, respectively (35, 46, 49, 50), as follows:

$$C_{pN}(T) = b_1 + b_2(T - T_0) \quad (3)$$

$$\Delta C_{pN-U}(T) = b_3 + b_4(T - T_0) + b_5(T - T_0)^2 \quad (4)$$

where $T_0 = 298.15$ (K), $b_1 = 4780.2$ (cal/mol/K), $b_2 = 21.334$ (cal/mol/K²), $b_3 = 6568$ (cal/mol/K), $b_4 = 23.84$ (cal/mol/K²), and $b_5 = -0.143$ (cal/mol/K³) (49). Given that the heat capacities of structurally similar proteins are similar (35, 49), I assume for the purpose of our study that the C_{pN} and ΔC_{pN-U} of CML can be approximated as those of HEL. The heat capacity change of each conformational transition of CML is assumed as:

$$\Delta C_{pN-A}(T) = \alpha \Delta C_{pN-U}(T) \quad (6)$$

$$\Delta C_{pA-U}(T) = (1 - \alpha) \Delta C_{pN-U}(T) \quad (7)$$

where the constant α represents the fraction of heat capacity change of the N-A transition relative to that of the N-U transition, and is assumed to be independent of temperature.

The enthalpy and entropy changes, from the native to the molten globule state $\{\Delta H_{N-A}(T), \Delta S_{N-A}(T)\}$ and from the molten globule to the thermal unfolded state $\{\Delta H_{A-U}(T), \Delta S_{A-U}(T)\}$, can be written as:

$$\Delta H_{N-A}(T) = \Delta H_{N-A}(T_{m1}) + \alpha \int_{T_{m1}}^T \Delta C_{pN-U}(T) dT \quad (8)$$

$$\Delta H_{A-U}(T) = \Delta H_{A-U}(T_{m2}) + (1 - \alpha) \int_{T_{m2}}^T \Delta C_{pN-U}(T) dT \quad (9)$$

$$\Delta H_{N-U}(T) = \Delta H_{N-A}(T) + \Delta H_{A-U}(T) \quad (10)$$

$$\Delta S_{N-A}(T) = \Delta H_{N-A}(T_{m1}) / T_{m1} + \alpha \int_{T_{m1}}^T \Delta C_{pN-U}(T) d \ln T \quad (11)$$

$$\Delta S_{A-U}(T) = \Delta H_{A-U}(T_{m2}) / T_{m2} + (1 - \alpha) \int_{T_{m2}}^T \Delta C_{pN-U}(T) d \ln T \quad (12)$$

$$\Delta S_{N-U}(T) = \Delta S_{N-A}(T) + \Delta S_{A-U}(T) \quad (13)$$

where T_{m1} and T_{m2} are the transition temperatures of the N-A and A-U transitions, respectively. In actual analyses of the experimental data, an instrumentally introduced constant A , due to the drift of base lines, etc., is introduced as follows:

$$C_{p,obs} = A + C_{pN} + f_A \Delta C_{pN-A} + f_U \Delta C_{pN-U} + [f_A \Delta H_{N-A}^2(T) + f_U \Delta H_{N-U}^2(T) - \{f_A \Delta H_{N-A}(T) + f_U \Delta H_{N-U}(T)\}^2] / RT^2 \quad (14)$$

The quantities f_A and f_U are expressed as functions of $\Delta H_{N-A}(T_{m1})$, $\Delta H_{A-U}(T_{m2})$, T_{m1} , T_{m2} , α and $\Delta C_{pN-U}(T)$. The constants $\Delta H_{N-A}(T_{m1})$, $\Delta H_{A-U}(T_{m2})$, T_{m1} , T_{m2} , α and A were determined for each set

of experimental data using the non-linear least-squares method.

Determination of the thermodynamic parameters of wild type and mutant CMLs

First, the parameter α was treated as an unknown constant to be determined by the non-linear least-squares method for six data sets of wild type apo-CML under different pHs (Figure 4). The mean of the resulting value for α was 0.57 with a standard deviation of 0.05. Then, the thermodynamic parameters were recalculated with a fixed α of 0.57, because the heat capacity change is usually independent of pH. This value of α was assumed for holo- and mutant CMLs.

Table 1 summarizes the thermodynamic parameters of wild type CML obtained by the above methods with the fixed α at several pHs. The temperature dependence of $\Delta H_{N-A}(T_{m1})$ and $\Delta H_{A-U}(T_{m2})$ is shown in Figure 5. The data of $\Delta H_{N-A}(T_{m1})$ and $\Delta H_{A-U}(T_{m2})$ are taken from Table 1. These plots were well fitted to each enthalpy function as follows:

$$\Delta H_{N-A}(T_{m1}) = \Delta H_{N-A}(T_0) + \alpha \int_{T_0}^{T_{m1}} \Delta C_{pN-U}(T) dT \quad (15)$$

$$\Delta H_{A-U}(T_{m2}) = \Delta H_{A-U}(T_0) + (1 - \alpha) \int_{T_0}^{T_{m2}} \Delta C_{pN-U}(T) dT \quad (16)$$

Table 2 shows the thermodynamic parameters of the wild type, H21G, I56L, A93S, and V109K mutants of CML and of α -hLA and EML in the absence of calcium ions at pH 4.5. The thermal stability of the molten globule state of all mutant lysozymes decreased relative to that of wild type CML. In fact, the thermal stability of the molten globule state of the H21G and V109K mutants was about 3 or 4° C less stable than that of wild type CML. Thus, the data indicated that the four mutated residues in CML played an important role in creating the extraordinary stability of the molten globule state.

Discussion

Despite the potential importance of this area of study and the considerable attention already received, few structural details of the molten globule state have been obtained, even under equilibrium conditions, because the flexibility of the state itself hampers the application of X-ray crystallographic studies (molten globules have not been crystallized successfully). Thanks to the progress of the NMR measurement, however, we could have obtained a lot of evidences for protected amide hydrogens and hence identified regions of persistent structure in helical regions of several molten globules over recent years (11, 16, 31, 32, 37, 41, 51–56). However, there are still lacking the energetics information of the molten globules (17, 35, 46–48) compared with those of the structural details of CD and NMR methodologies. The purpose of this CHAPTER is to understand the energetics of the molten globule state of three homologous proteins (CML, α -hLA, and EML) by using calorimetry, and to discuss the results of structural details obtained from CD spectroscopy and hydrogen-exchange experiments.

Non-native and native-like packing of the molten globule states

Given that the overall crystal structures of CML (29), α -hLA (26), and EML (40) are very similar, I consider that the difference in the stability of the molten globule states among these proteins is not due to a difference in overall native structure. The near-UV CD spectrum of the molten globule state (A-state) of α -hLA shows an apparent loss of the signal relative to the native α -hLA (Figure 2C), suggesting disrupted side-chain packing (7, 44). However, the difference of the amplitude between N and A spectra for CML and EML are very small relative to that for α -hLA. The near-UV CD spectra of the molten globule states of CML and EML indicate an intermediate level of signal between the N and U spectra, and the spectra are much more close to those of the N proteins (Figure 2A & B). In spite of a disappearance of the positive Cotton effects around 294 nm, the

positive bands around 288 nm show significant signals which is about 65 % and 73 % intensities of native CML and EML, respectively. These results demonstrate the presence of packing of aromatic side-chains in the molten globule states of CML and EML, while the absence of such native-like packing interactions in the molten globule state of α -hLA. A hydrophobic cluster involving three tryptophanyl residues (Trp-28, 108, and 111) seems to be more densely packed in the molten globule states in CML and EML than in α -hLA (these three residues are Leu-26, Trp-104, and His-107 in α -hLA), and this may be responsible for the enhanced near-UV CD band (29, 31, 32).

The thermodynamic analysis by using calorimetry (DSC) is well reflecting the results of CD spectra. The non-cooperative thermal transition of the molten globule state, such as in the case of α -hLA (Figure 3A; top), may simply be interpreted in terms of a very small enthalpy change, because the enthalpy change of the protein unfolding primarily arises from the disorder of the specific side-chain packing interaction (57). The cooperative thermal transition of the molten globule state in the case of CML and EML (Figure 3A; middle and bottom) suggests the presence of specific packing interactions; in contrast, such interactions are thought to be minimal in the molten globule state of α -hLA. In general, typical values for the heat capacity change (ΔC_p) for the reactions of the proteins show that thermal unfolding results in a large positive change, which is ascribed to the exposure of hydrophobic groups to the aqueous environment (58). The ΔC_p values of the A-U transition of CML and EML were 2.9 kJ/mol/K and 2.5 kJ/mol/K at 66.4° C and pH 4.5 (Figure 5), respectively, although that of the α -hLA was close to zero. These results indicate that the interactions stabilizing the molten globule state are more extensive and cooperative in CML and EML than in α -hLA.

Structural basis of the cooperative thermal transitions

Cooperative folding, which is specific to globular proteins, results from tight and unique packing of

amino acid residues in the native structure (59–61). Therefore, the thermal transition of the “classical” molten globule states is expected to be a weakly or non-cooperative process, since they do not possess a unique tertiary structure (6–9, 44). However, my results show that the molten globule states of CML and EML can be considered as the highly cooperative intermediate, while that of the α -hLA can be considered as the non-cooperative intermediate (Figure 3). In order to explain the cooperative thermal transitions of the molten globule states of CML, α -hLA, and EML, the hydrogen exchange properties (protection factors) of the A-states were compared among these proteins (Figure 6). There are several important features of the amide hydrogen exchange protection exhibited in the α and β -domains of CML, α -hLA, and EML. In the case of α -hLA, the protected protons were detected only within the α -domain, and none of the amide protons within the β -domain were significantly protected from solvent exchange in the A-state (41) (Figure 6; top). It is indicating that molten globule state of α -hLA has a bipartite structure with a disordered β -domain and an α -domain containing substantial secondary structure and a native-like tertiary topology (13–15, 41). Therefore, the cooperative thermal transition of α -hLA is apparently absent in this state because of the compact but dynamic structure (low order structure).

In contrast, the number and magnitude of protected amide protons in the A-state of CML (64 amide protons) and EML (54 amide protons) were much larger than that of α -hLA (34 amide protons) (31, 32, 37). The protected protons of CML and EML in this state were observed not only within the α -domain but also within the β -domain (Figure 6; middle & bottom). It is suggesting that highly persistent structures exist in the α - and β -domain of CML and EML in the A-state, and especially, the β -domain of the two proteins are more ordered than that of α -hLA. Thus, it can be proposed that the cooperative thermal transitions of the molten globule states, which have been observed in CML and EML, may come from a highly ordered structure of the β -domain in the A-states. This is one of the main reasons why the cooperative thermal transitions are observed in

the molten globule state of CML and EML, and not in that of the α -hLA.

Comparison of the molten globule state of CML with EML

As discussed above, it has been shown that CML and EML show the presence of native-like tertiary folding in the molten globule state, which, in these two proteins, is a thermodynamic state. The molten globule state of CML, in particular, has been shown to be extraordinarily stable in comparison with that of EML; comparatively, it was shown to be 9.0 kJ/mol more stable at pH 4.5 and 66.4° C in the absence of Ca^{2+} ions (Table 3). However, what causes these differences in the stability of the molten globule state between CML and EML? One way to investigate the differences is to thoroughly compare the physicochemical properties of the molten globule state of CML and EML.

In the crystal structure of CML, shown in Figure 7A, His-21 is located at the turn region between the A- and B- helices and is close to the C-helix (also see Figure 1). The distance between the $\text{N}^{\text{H}}2$ atom of His-21 and the O atom ($\text{C}=\text{O}$) of Val-99 is 3.13 Å, indicating the formation of a hydrogen bond between these two atoms in the native structure. As a result, the A- and B- helices become much closer to the C- helix, and hydrophobicity around this region is increased. In fact, hydrogen exchange rates in this region of CML have exhibited uniform behavior. Eleven aromatic and aliphatic residues within this region (Arg-10, Lys-11, Leu-12, Lys-13, and Met-15 of the A-helix; Phe-20 and Tyr-23 of the A-B loop; and Cys-30, Met-31, Ala-32, and Glu-33 of the B-helix) exhibit very high protection factors ($P > 100$) in the A-state (pH 2.0 and 25° C). In this condition, the most protected residues of the A- and B-helices are located in the interface of these helices with the C-helix. This suggests that the topology of the A-, B-, and C-helices in the A-state might be similar to that of the thermally induced molten globule state at pH 4.5. However, protection factors of EML in this region were much lower ($10 < P < 1$) than those of CML in the

A-state (Figure 7B).

Moreover, residue 109 is located in the N terminal of the D-helix. This residue is valine and lysine in CML and EML, respectively (Figure 8A and B). The helix contact between the B- and D-helices is revealed clearly in the hydrogen exchange data for the A-state (pH 2.0 and 25° C) of CML and EML. Seven aromatic and aliphatic residues around this region (Cys-30, Met-31, Ala-32, and Glu-33 of the B-helix; Trp-108 of the C-D loop; Val-109 of the D-helix; and His-114 of the loop) exhibit the highest protection factors ($P > 100$) at pH 2.0 and 25° C in CML. The most protected residues of the B-helix are also located in the interface of this helix with the D-helix. The Val-109 is particularly strongly immobilized in the A-state ($P = 330$), and seems to stabilize the D-helix. However, the residue at this site is lysine in EML, and such high protection is not observed in the case of the A-state of EML. This may suggest that the hydrogen exchange kinetics of Lys-109 occurs so rapidly that the protection factor cannot be determined. This exchange of Val-109 and Lys-109 must be a major reason for the destabilization of the D-helix of EML.

The present results show that the amino acid replacements discussed above actually decrease the stability of the molten globule state of CML. In particular, the values for H21G and V109K are considerably lower, 2.0 kJ/mol and 1.8 kJ/mol lower, respectively, than those of the wild-type CML. The thermodynamic parameters suggest that the Gly-21 $\rightarrow$ His-21 replacement has two effects on the stabilization of the molten globule state. First, by forming a hydrogen bond between the side chain of His-21 and the main chain of Val-99, the A- and B-helices are connected to the C-helix, and these chains become less flexible. Therefore, it can be estimated that the replacement of a histidine residue by a glycine residue would destabilize the protein at most by 2.0 kJ/mol in the ΔG of the A-U transition, and by 7.7 kJ/mol in the ΔH of the A-U transition. Second, the involvement of the His-21 residue on the hydrophobic cluster makes the structure more stable, and the molten globule state of CML is further stabilized. In the case of the V109K mutant, these

two positively charged side chains (Lys-109 and -113) are thought to destabilize the native structure due to electrostatic repulsion. Among the thermodynamic parameters for the A-U transition of EML and the V109K mutant, the major terms to destabilize the molten globule state are the entropy terms ($T\Delta\Delta S = 3.0$ kJ/mol for EML and $T\Delta\Delta S = 5.8$ kJ/mol for V109K) (Table 3). The enthalpy terms seem to compensate only partly for the entropic destabilization of the EML and V109K. It appears that the destabilization of the D-helix is caused by the increase of chain entropy in the molten globule state, which increase is caused by the existence of the lysine residue at the N terminal of the D-helix. This may indicate that the helical contact between the B- and D-helices of EML is much weaker than that of CML.

Conclusion

I have shown that CML and EML exhibit two clear heat absorption peaks in the absence of Ca^{2+} ions (apo-form), although the cooperative thermal transition of α -hLA is apparently absent in this form. Moreover, CML also exhibits two clear heat absorption peaks in the presence of Ca^{2+} ions (holo-form), although α -hLA and EML show only a single heat absorption peak in the holo-forms. My data indicate that the molten globule state of CML is extraordinarily stable compared with those of α -hLA and EML, and that it is a thermodynamic state. I interpret that the extraordinary stability of the molten globule state of CML is due to the native-like packing interaction in the molten globule state, and it comes from the highly persistent structures existing not only within the α -domain but also within the β -domain.

References

1. Kim, P. S. & Baldwin, R. L. (1990) *Annu. Rev. Biochem.*, **59**, 631-660.

2. Ohgushi, M. & Wada, A. (1983) *FEBS Lett.*, **164**, 21-24.
3. Dolgikh, D. A., Abaturon, L. V., Brazhnikov, E. V., Lebedev, I. O., Chirgadze, I. N. & Ptitsyn, O. B. (1983) *Dokl. Akad. Nauk. SSSR*, **272**, 1481-1484.
4. Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. (1985) *Biochemistry*, **24**, 874-881.
5. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986) *Biochemistry*, **25**, 6965-6972.
6. Ptitsyn, O. B. (1992) In "Protein folding" (T. E. Creighton, ed.). Freeman & Company, New York, p243-300.
7. Kuwajima, K. (1989) *Proteins: Struct. Funct. Genet.*, **6**, 87-103.
8. Kuwajima, K. (1996) *FASEB J.*, **10**, 102-109.
9. Ptitsyn, O. B. (1995) *Adv. Protein Chem.*, **47**, 83-229.
10. Arai, M. & Kuwajima, K. (2000) *Adv. Protein Chem.*, **53**, 209-283.
11. Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. (1989) *Biochemistry*, **28**, 7-13.
12. Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990) *Science*, **249**, 1544-1548.
13. Peng, Z-yu & Kim, P. S. (1994) *Biochemistry*, **33**, 2136-2141.
14. Peng, Z-yu, Wu, L. C. & Kim, P. S. (1995) *Biochemistry*, **34**, 3248-3252.
15. Wu, L. C., Peng, Z-yu & Kim, P. S. (1995) *Nat. Struct. Biol.*, **2**, 281-286.
16. Eliezer, D., Jennings, P. A., Dyson, H. J. & Wright, P. E. (1997) *FEBS Lett.*, **417**, 92-96.
17. Privalov, P. L. (1996) *J Mol Biol.*, **258**, 707-725.
18. Fink, A. L., Oberg, K. A. & Seshadri, S. (1997) *Folding Design*, **3**, 19-25.
19. Gilmanshin, R., Dyer, R. B. & Callender, R. H. (1997) *Protein Sci.*, **6**, 2134-2142.
20. Grobler, J. A., Rao, K. R., Pervaiz, S. & Brew, K. (1994) *Arch. Biochem. Biophys.*, **313**, 360-366.
21. Kikuchi, M., Kawano, K. & Nitta, K. (1998) *Protein Sci.*, **7**, 2150-2155.
22. Koshiba, T., Hayashi, T., Ishido, M., Kumagai, I., Ikura, T., Kawano, K., Nitta, K. &

- Kuwajima, K. (1999) *Protein Eng.*, **12**, 429-435.
23. Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S. & Murai, N. (1980) *Biochem. Biophys. Res. Commun.*, **95**, 1098-1104.
 24. Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M. & Phillips, D. C. (1986) *Nature*, **324**, 671-675.
 25. Nitta, K., Tsuge, H., Shimazaki, K. & Sugai, S. (1988) *Biol. Chem. Hoppe-Seyler*, **369**, 671-675.
 26. Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C. & Fenna, R. E. (1991) *J. Mol. Biol.* **221**: 571-581.
 27. Sugai, S. & Ikeguchi, M. (1994) *Adv. Biophys.*, **30**, 37-84.
 28. Pike, A. C. W., Brew, K. & Acharya, K. R. (1996) *Structure*, **4**, 691-703.
 29. Koshiba, T., Yao, M., Kobashigawa, Y., Demura, M., Nakagawa, A., Tanaka, I., Kuwajima, K. & Nitta, K. (2000) *Biochemistry*, **39**, 3248-3257.
 30. Morozova, L. A., Haezebrouck, P. & Van Cauwelaert, F. (1991) *Biophys. Chem.*, **41**, 185-191.
 31. Morozova, L. A., Haynie, D. T., Arico-Muendel, C., Van Dael, H. & Dobson, C. M. (1995) *Nature Struct. Biol.*, **2**, 871-875.
 32. Morozova, L. A., Arico-Muendel, C., Haynie, D. T., Emelyanenko, V. I., Van Dael, H. & Dobson, C. M. (1997) *J. Mol. Biol.*, **268**, 903-921.
 33. Morozova, L. A., Jones, J. A., Noppe, W. & Dobson, C. M. (1999) *J. Mol. Biol.*, **289**, 1055-1073.
 34. Van Dael, H., Haezebrouck, P., Morozova, L. A., Arico-Muendel, C. & Dobson, C. M. (1993) *Biochemistry*, **32**, 11886-11894.
 35. Griko, Y. V., Freire, E., Privalov, G., Van Deal, H. & Privalov, P. L. (1995) *J. Mol. Biol.*, **252**, 447-459.

36. Mizuguchi, M., Arai, M., Yue, K., Nitta, K. & Kuwajima, K. (1998) *J. Mol. Biol.*, **283**, 262-277.
37. Kobashigawa, Y., Demura, M., Koshiba, T., Kumaki, Y., Kuwajima, K. & Nitta, K. (2000) *Proteins: Struct. Funct. Genet.*, **40**, 579-589.
38. Mizuguchi, M., Masaki, K. & Nitta, K. (1999) *J. Mol. Biol.*, **292**, 1137-1148.
39. Koshiba, T., Tsumoto, K., Masaki, K., Kawano, K., Nitta, K. & Kumagai, I. (1998) *Protein Eng.*, **11**, 683-690.
40. Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S. & Miyano, M. (1992) *J. Biochem.*, **111**, 141-143.
41. Schulman, B. A., Redfield, C., Peng, Z-yu, Dobson, C. M. & Kim, P. S. (1995) *J. Mol. Biol.*, **253**, 651-657.
42. Kraulis, P. J. (1991) *J. Appl. Crystallog.*, **24**, 946-950.
43. Nozaka, M., Kuwajima, K., Nitta, K. & Sugai, S. (1978) *Biochemistry*, **17**, 3753-3758.
44. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. (1976) *J. Mol. Biol.*, **106**, 359-373.
45. Yutani, K., Ogasahara, K. & Kuwajima, K. (1992) *J. Mol. Biol.*, **228**, 347-350.
46. Griko, Y., Freire, E. & Privalov, P. L. (1994) *Biochemistry*, **33**, 1889-1899.
47. Pfeil, W. (1998) *Proteins: Struct. Funct. Genet.*, **30**, 43-48.
48. Griko, Y. (2000) *J. Mol. Biol.*, **297**, 1259-1268.
49. Gomez, J., Hilser, V., Xie, D. & Freire, E. (1995) *Proteins: Struct. Funct. Genet.*, **22**, 404-412.
50. Xie, D., Fox, R. & Freire, E. (1995) *Protein Sci.*, **3**, 2175-2184.
51. Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol.*, **3**, 84-91.
52. Chyan, C. L., Wormald, C., Dobson, C. M., Evans, P. A. & Baum, J. (1993) *Biochemistry* **32**, 5681-5691.
53. Chamberlain, A. K. & Marqusee, S. (1998) *Biochemistry*, **37**, 1736-1742.

54. Forge, V., Wijesinha, J., Balbach, K., Brew, K., Robinson, C. V., Redfield, C. & Dobson, C. M. (1999) *J. Mol. Biol.*, **288**, 673-688.
55. Wrabl, J. & Shortle, D. (1999) *Nat. Struct. Biol.*, **6**, 876-883.
56. Englander, S. W. (2000) *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 213-238.
57. Makhatadze, G. I. & Privalov, P. L. (1995) *Adv Protein Chem.*, **47**, 307-425.
58. Edsall, J. T. & Gutfreund, H. (1983) *Biothermodynamics*. John Wiley & Sons, Chichester 215-217.
59. Privalov, P. L. (1979) *Adv. Protein Chem.*, **33**, 167-241.
60. Privalov, P. L. & Gill, S. L. (1988) *Adv. Protein Chem.*, **39**, 191-234.
61. Shakhnovich, E. & Finkelstein, A. V. (1989) *Biopolymers*, **28**, 1667-1680

Table I. Thermodynamic parameters of the thermal transitions (N-A and A-U) of the wild type CML in the absence (apo-form) and presence (holo-form) of Ca^{2+} ions at several pHs.

pH	$T_m^{\text{N-A}}(^{\circ}\text{C})$	$\Delta H^{\text{N-A}}(\text{kJ/mol})$	$T_m^{\text{A-U}}(^{\circ}\text{C})$	$\Delta H^{\text{A-U}}(\text{kJ/mol})$
2.74	20.54	89.3	70.13	158.0
3.00	28.40	133.0	73.98	154.4
3.24	33.74	148.9	78.58	179.7
3.46	36.39	157.8	81.54	175.8
3.99	39.74	164.7	84.97	168.1
4.50	39.01	176.0	86.92	185.8
4.50 (holo)	65.13	282.0	86.25	185.8

Table II. Thermodynamic parameters of the thermal transitions (N-A and A-U) of the wild type and mutant CMLs and those of α -hLA and EML in the absence (apo-form) of Ca^{2+} ions at pH 4.5.

Protein	$T_m^{\text{N-A}}(^{\circ}\text{C})$	$\Delta H^{\text{N-A}}(\text{kJ/mol})$	$T_m^{\text{A-U}}(^{\circ}\text{C})$	$\Delta H^{\text{A-U}}(\text{kJ/mol})$
CML				
wild type	39.01	176.0	86.92	185.8
V109K	41.78	183.1	82.76	179.1
H21G	34.13	124.9	83.64	169.7
I56L	26.19	99.6	84.48	183.1
A93S	39.37	165.6	84.33	183.4
EML [#]	41.50	153.7	66.44	123.8
α -hLA	N.D.	N.D.	N.D.	N.D.

[#]Previous study (Griko et al., 1995).

N.D., not detected by the present method.

Table III. Thermodynamic parameters of the thermal transition (A-U) of the wild type and mutant CMLs and that of EML in the absence (apo-form) of Ca^{2+} ions at pH 4.5 and 66.4 °C.

Lysozyme	T_m (°C)	ΔH^* (kJ/mol)	$T\Delta S^*$ (kJ/mol)	ΔG^* (kJ/mol)	$\Delta\Delta G^*$ (kJ/mol)
CML					
wild type	86.93	129.5	120.5	9.0	0.0
V109K	82.76	133.5	126.3	7.2	-1.8
H21G	83.64	121.8	114.8	7.0	-2.0
I56L	84.48	133.1	125.1	8.0	-1.0
A93S	84.33	133.7	125.8	7.9	-1.1
EML [#]	66.40	123.5	123.5	0.0	-9.0

*These values are extrapolated to 66.4 °C by equations (9) and (12).

[#]Estimated from previous study (Griko et al., 1995).

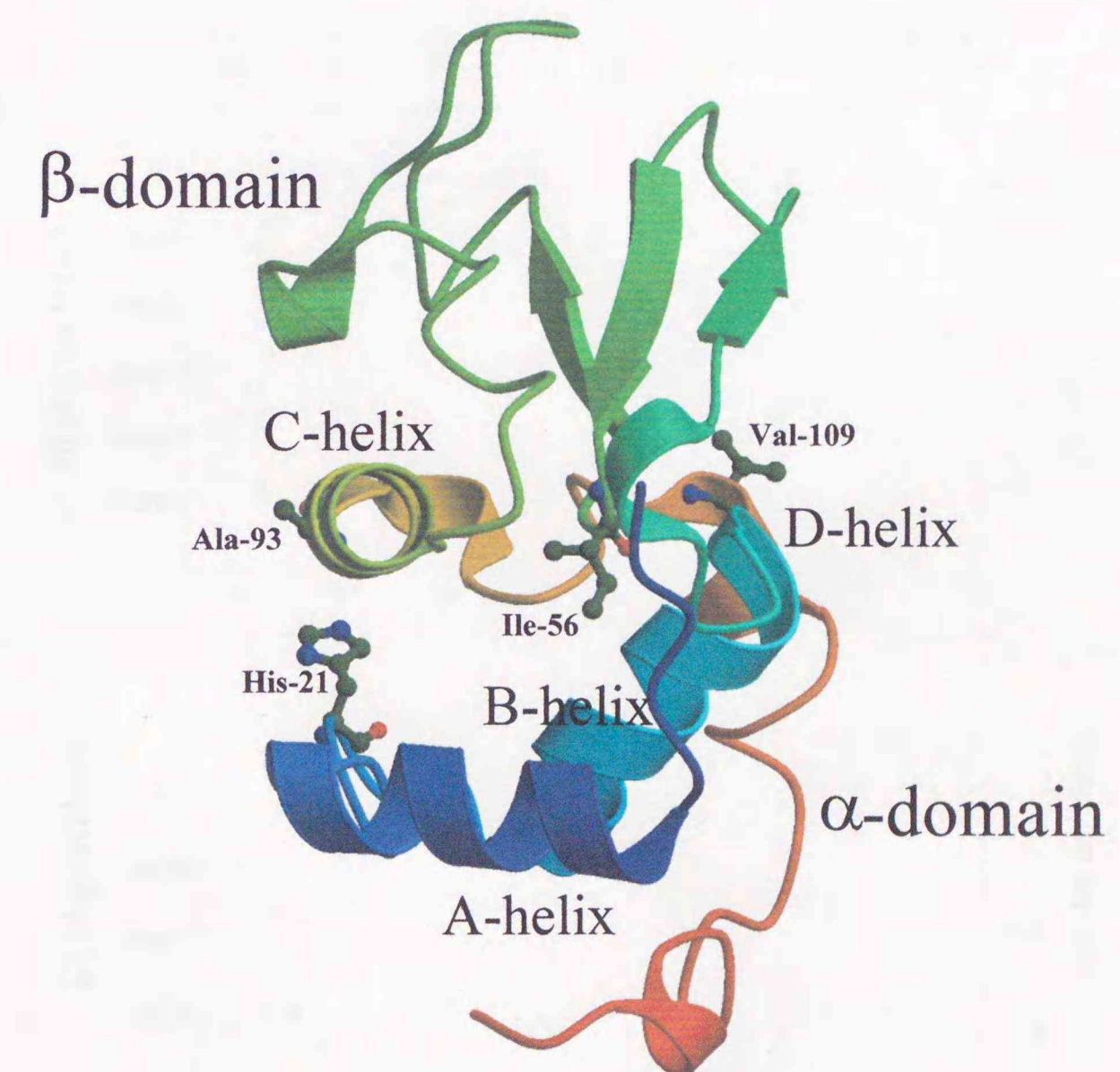


Figure 1. Schematic representation of the crystal structure of apo-CML (29) showing the backbone and four residues, His21, Ile56, Ala93, and Val109, substituted in this study. α -Helices are sequentially labeled from A to D. The structure shows a rainbow color (red to blue) from the N-terminus to the C-terminus. This figure was produced using MOLSCRIPT (42).

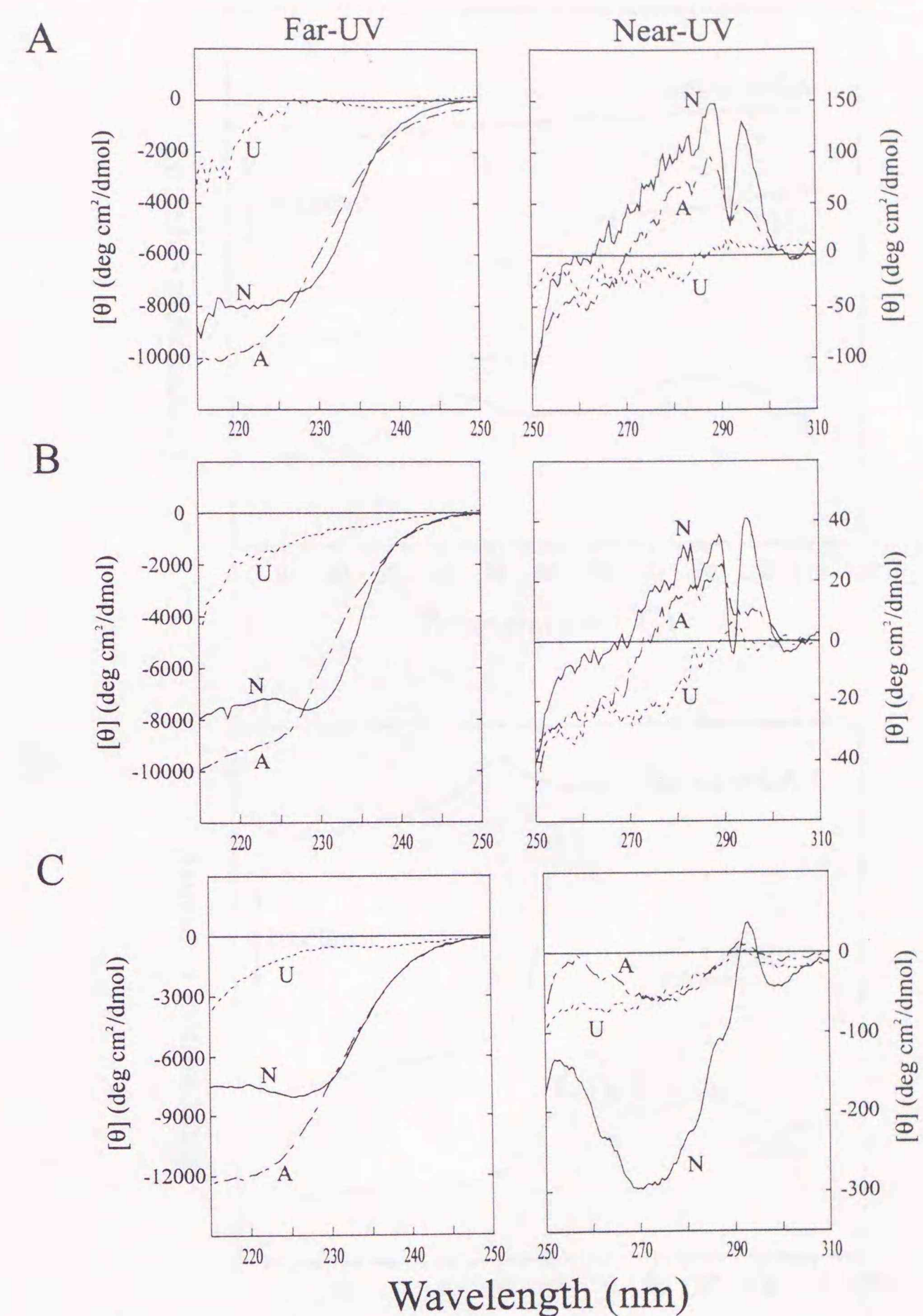
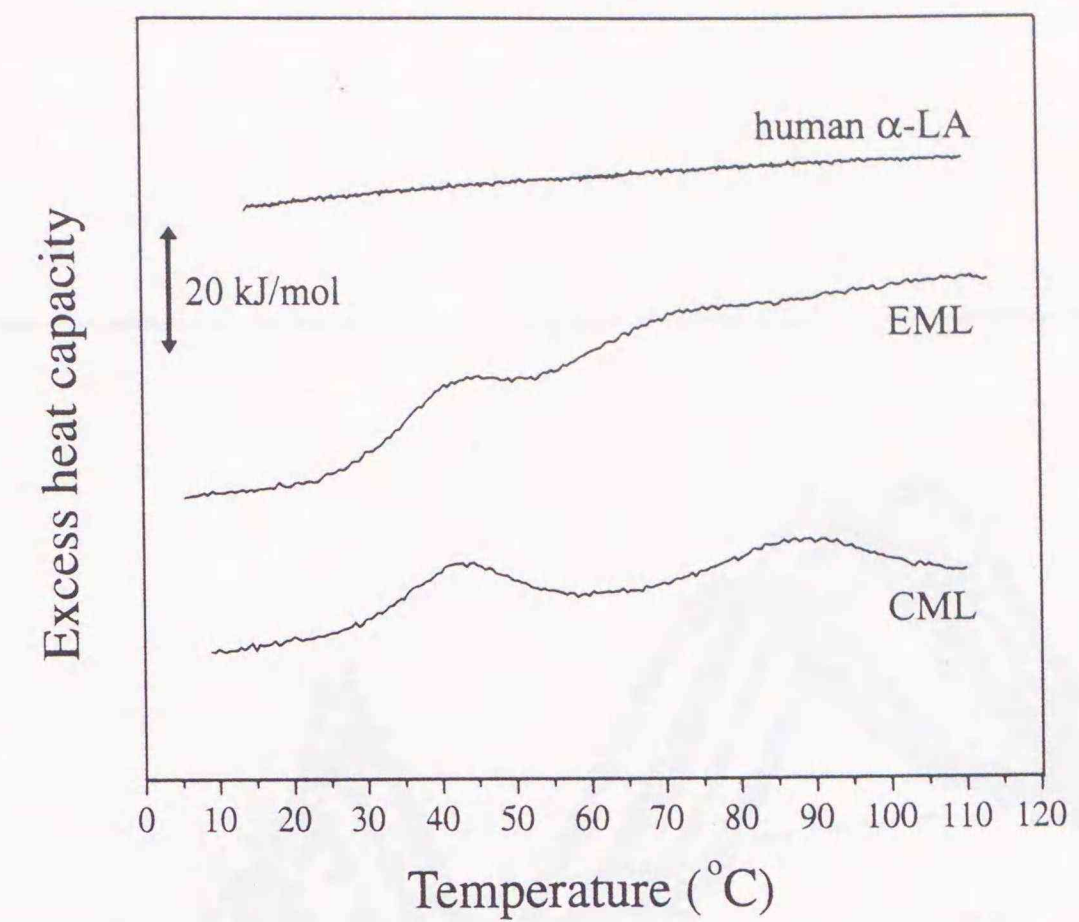


Figure 2. CD spectra of native (N, solid lines), molten globule (A, long broken lines), and unfolded states (U, short broken lines) of (A) CML, (B) EML, and (C) α -hLA in the far-UV (left) and near-UV (right) regions. The native proteins were dissolved in 50 mM Tris-HCl containing 10 mM CaCl₂ (pH 8.0), the molten globule proteins in 50 mM glycine-HCl (pH 2.0), and the unfolded proteins in 6 M guanidine hydrochloride (pH 8.0). All spectra were measured at 25° C.

A



B

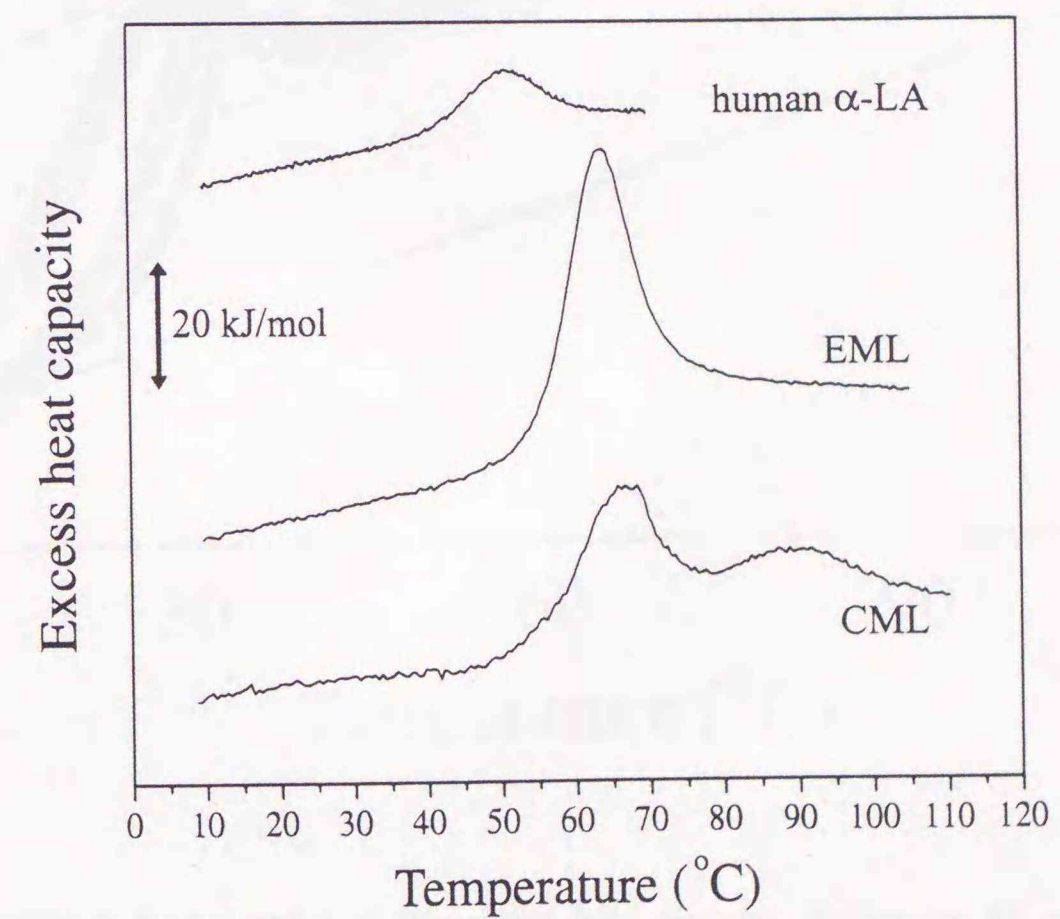


Figure 3. (A) The thermal unfolding transition of CML, α -hLA, and EML in the absence (apo-form) of Ca^{2+} ions at pH 4.5. (B) The thermal unfolding transition of CML, α -hLA, and EML in the presence of 10 mM CaCl_2 at pH 4.5 (CML and EML) and pH 8.0 (α -hLA), respectively. The unfolding transitions are expressed by excess heat capacity as a function of temperature.

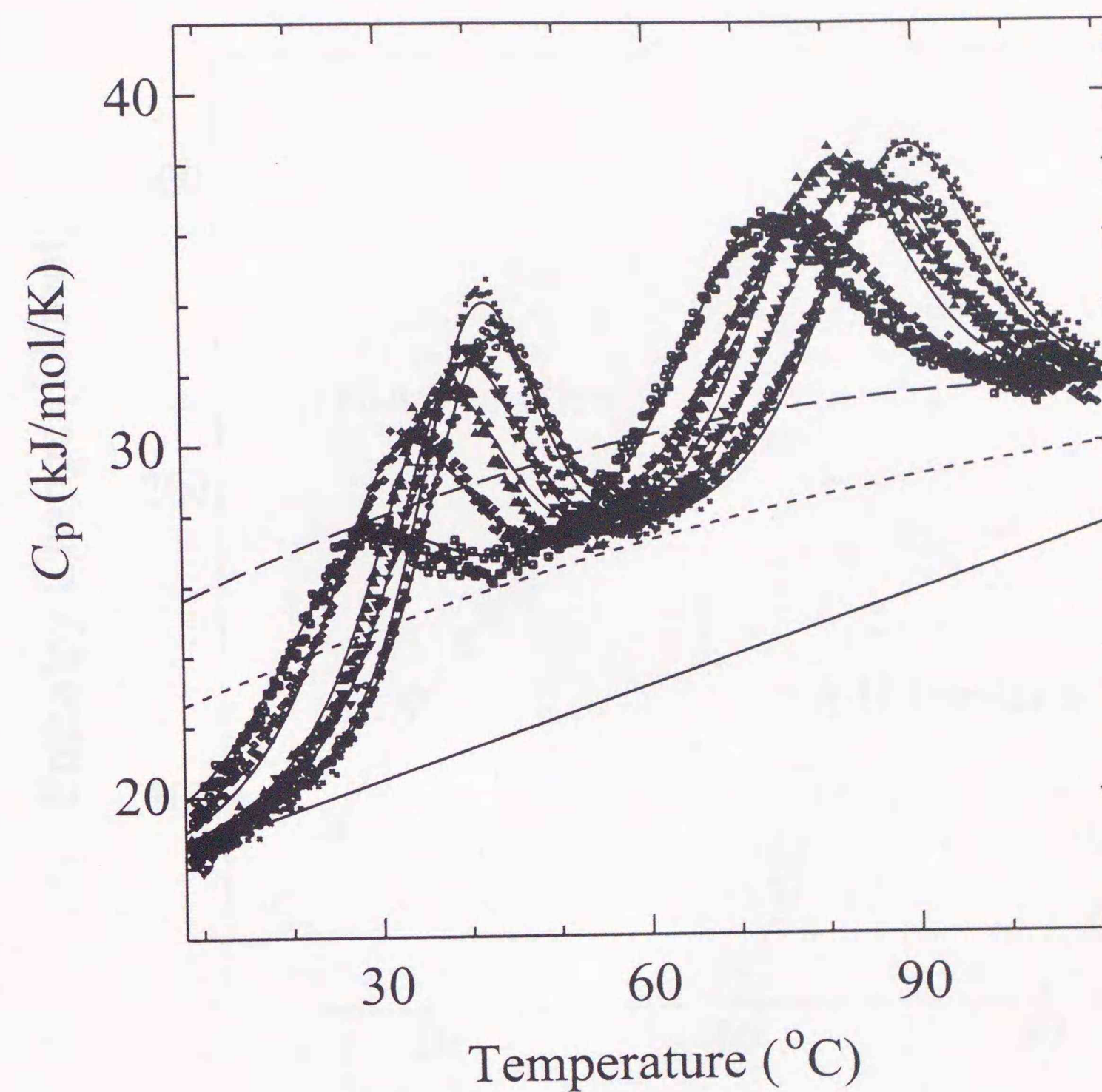


Figure 4. Temperature dependencies of the partial heat capacity of CML in 50 mM sodium acetate or glycine hydrochloride solutions without Ca^{2+} ions at various pHs. Open squares, pH 2.74; diamonds, pH 3.00; triangles, pH 3.24; reversed triangles, pH 3.46; crosses, pH 3.99; circles, pH 4.50. The thick solid and broken lines represent the heat capacity functions of the native and unfolded states of HEL (49), which are assumed to be applicable to CML in this study. The dotted line represents the heat capacity function of the molten globule state of CML with an α parameter of 0.57 (see text).

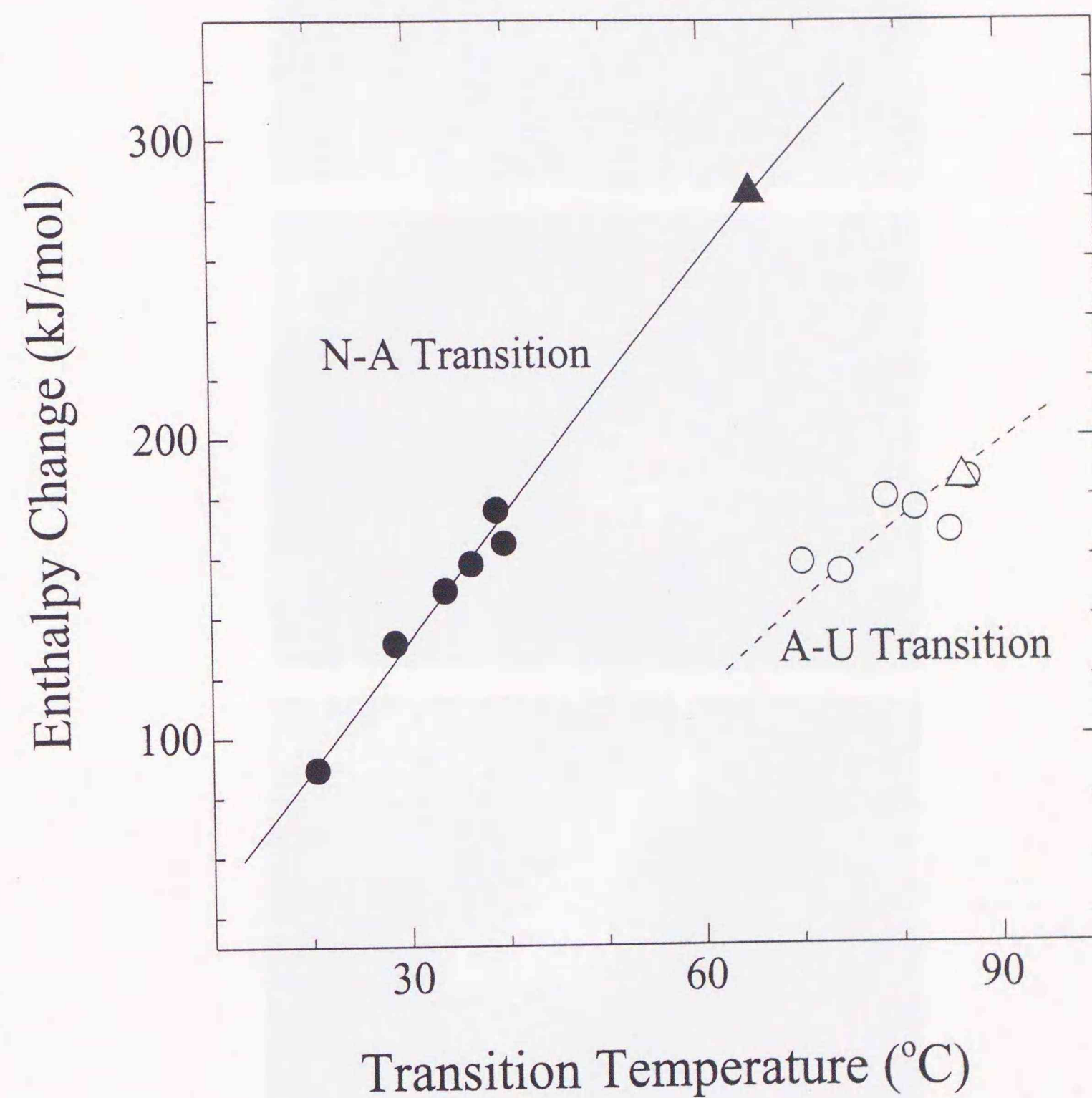


Figure 5. Temperature dependencies of the enthalpy change for the N-A transition (filled symbols) and the A-U transition (open symbols) of CML in the absence (circle) and presence (triangle) of Ca^{2+} ions. The solid and broken lines represent the enthalpy change functions calculated from equations (15) and (16) with $\alpha = 0.57$, $\Delta H_{\text{N-A}}(T_0) = 110.5$ kJ/mol, and $\Delta H_{\text{A-U}}(T_0) = 4.2$ kJ/mol, respectively.

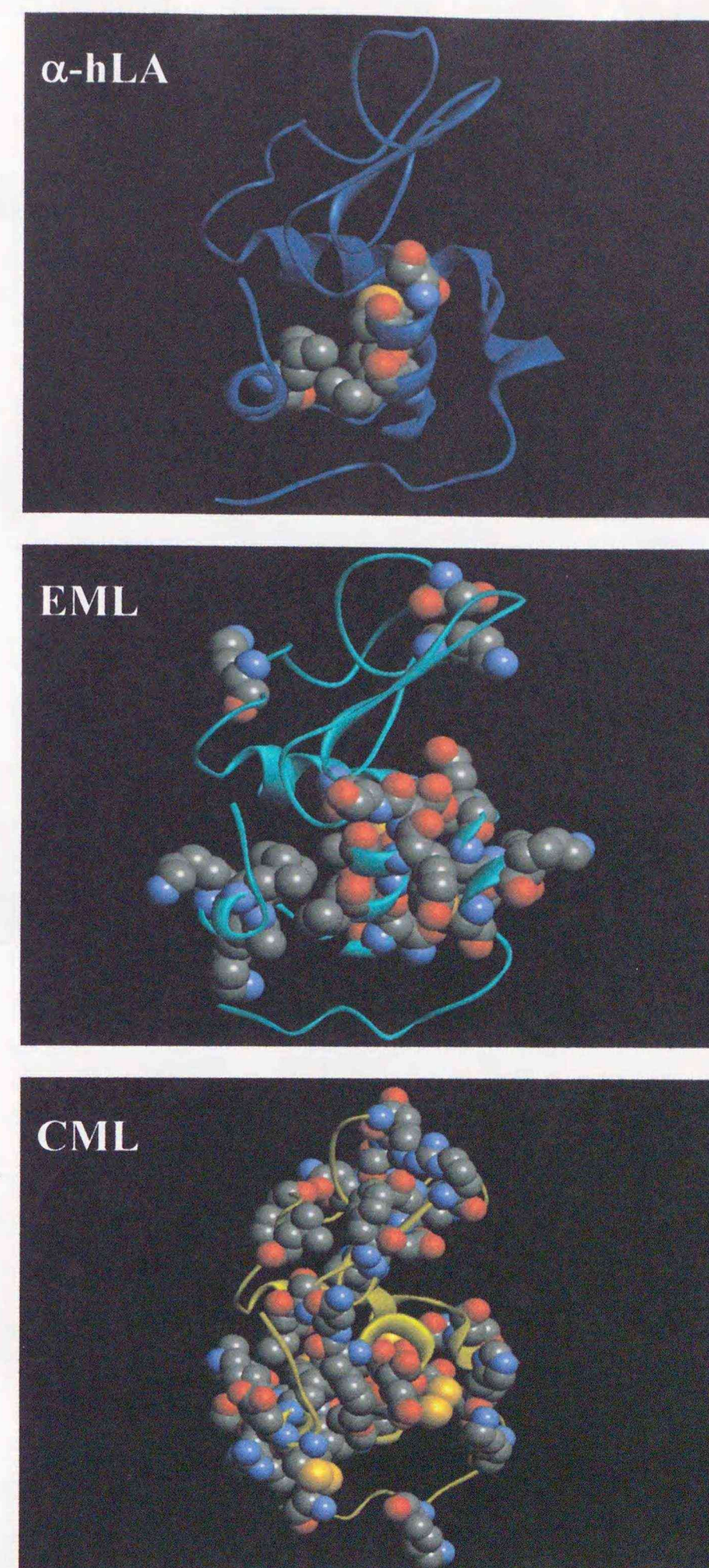


Figure 6. The most protected amide protons ($P > 20$) in the A-states are highlighted on the crystal structures of α -hLA (26), EML (40), and CML (29) with the CPK models. The hydrogen exchange rate data of CML (37), α -hLA (41), and EML (31, 32) at pH 2.0 were obtained from previous reports. Molecular graphics were produced using WebLab Viewer (Molecular Simulations Inc.).

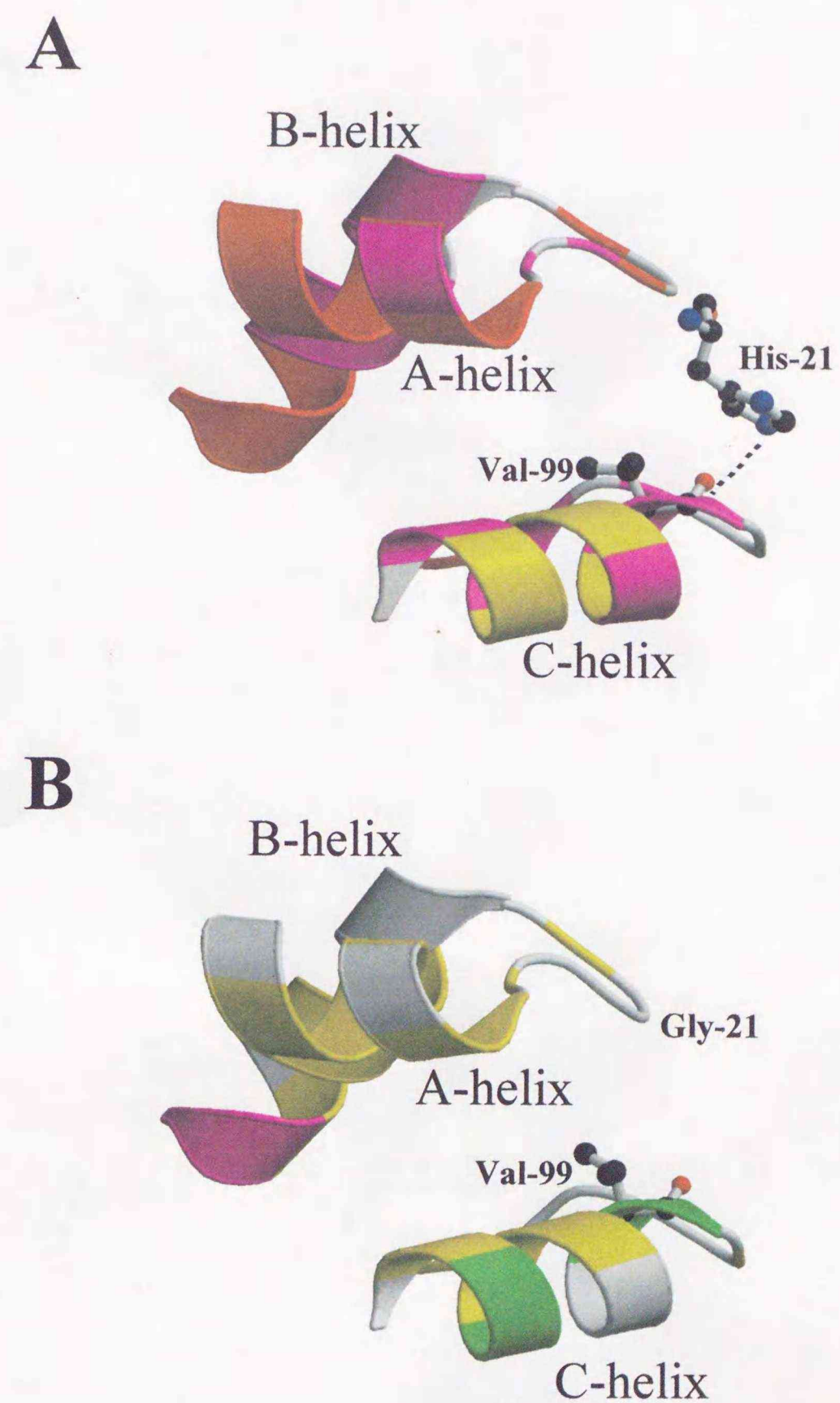


Figure 7. Partial structures of (A) CML and (B) EML around the region of the A-, B-, and C-helices. The broken line indicates a hydrogen bond. The structures are indicated by four colors that represent the protection factors as $P > 100$ (red), $100 > P > 10$ (magenta), $10 > P > 1$ (yellow), and $1 > P$ (green). The figures were produced using MOLSCRIPT (42).

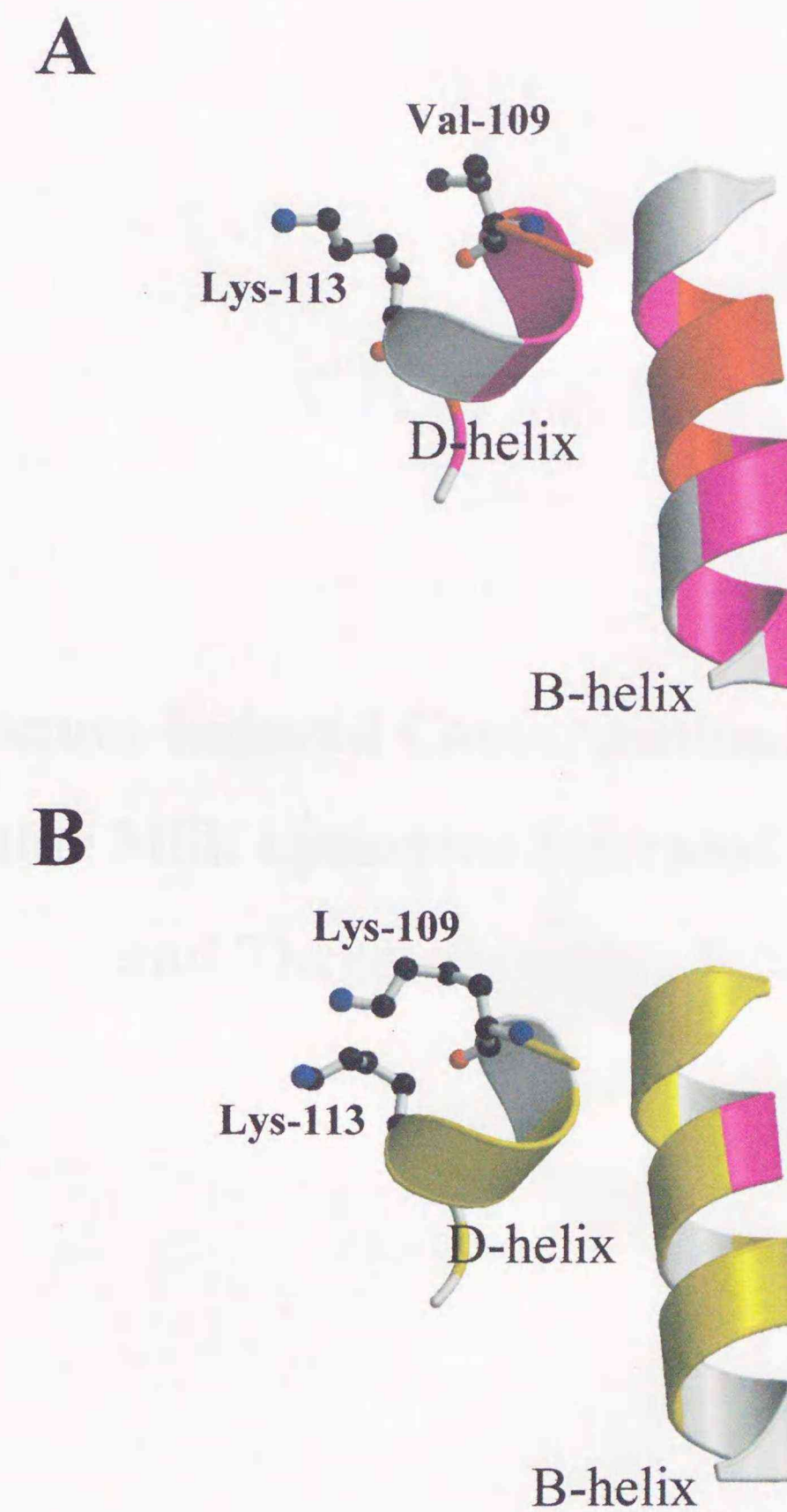


Figure 8. Partial structures of (A) CML and (B) EML around the region of the B- and D-helices. The structures are indicated by three colors that represent the protection factors as $P > 100$ (red), $100 > P > 10$ (magenta), $10 > P > 1$ (yellow). The figures were produced using MOLSCRIPT (42).



Figure 4.1. Ribbon diagrams of the protein structure in the open (A) and closed (B) states. The structure was determined by X-ray crystallography. The open state (A) is shown in purple and the closed state (B) is shown in yellow.

CHAPTER IV

Calcium-Induced Conformational Changes of Canine Milk Lysozyme Revealed by Structural and Thermodynamic Evidence

Abstract

In order to demonstrate the conformational changes between the Ca^{2+} free (apo-) and binding (holo-) forms of a Ca^{2+} -binding protein, the thermodynamic changes of canine milk lysozyme (CML) were investigated using isothermal titration calorimetry and interpreted in terms of structural information obtained from nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. On the basis of the results of calorimetry, I determined that the Ca^{2+} -binding reaction to CML was an exothermic process and the enthalpic contribution for the reaction was larger than that of the entropy. The negative heat capacity changes ($\Delta C_p = -0.21$ kcal/mol/K) of CML observed upon Ca^{2+} -binding suggesting a characteristic of a prominent hydrophobic effect. The structure analysis using ^1H - ^{15}N HSQC spectra and X-ray diffraction showed that two large conformational changes occurred during the Ca^{2+} -binding. One was a local structural shift in the vicinity of the Ca^{2+} -binding site, and the other was observed in a hydrophobic core region that included Lys-62, Trp-63, and Trp-64, as if the hydrophobic interaction increased around the cluster. These conformational changes are consistent with the thermodynamic data obtained by calorimetry. This is the first example of investigating the Ca^{2+} -induced conformational change of a Ca^{2+} -binding protein by thermodynamic study, NMR spectroscopy, and X-ray crystallography.

Introduction

The calcium ion is essential for living organisms; it is used in cells as a signal transduction element. An increase in the cytoplasmic Ca^{2+} concentration leads to the binding of Ca^{2+} by intracellular regulatory proteins, an event that initiates a wide variety of cellular processes (1–5). The Ca^{2+} -binding reaction sometimes accompanies conformational changes of the protein between the Ca^{2+} -free form (apo-form) and the Ca^{2+} -bound form (holo-form), which are considered to be responsible for biological functions such as signal transduction and the formation of macromolecular complexes (6–8). Therefore, to elucidate the Calmodulin has been investigated as a potential model protein in which conformational changes can be studied using nuclear magnetic resonance (NMR) spectroscopy (9–15). However, there is a lack of information about the conformational changes of proteins between the apo- and holo-forms by X-ray crystallography.

My approach in this CHAPTER is to investigate the conformational changes of a Ca^{2+} -binding protein between the apo- and holo-forms using structural (X-ray crystallography and NMR spectroscopy) and thermodynamic analysis. For this purpose, I chose canine milk lysozyme (CML) because we have found that this protein is one of a Ca^{2+} -binding protein (16), and have recently determined the structure of apo-CML by X-ray crystallography at a resolution of 1.85 Å to an *R*-factor of 17.8 % (17). In addition, the results of the ^1H resonance assignments of holo-CML have been reported recently (18).

I noted that two large conformational changes occur during the Ca^{2+} -binding of CML. In order to demonstrate these conformational changes of CML, I investigated the thermodynamic changes of the protein using isothermal titration calorimetry (ITC) and interpreted the findings in terms of structural information obtained from ^1H NMR and X-ray crystallography. Calorimetric study revealed that the heat capacity changes (ΔC_p) of CML, which are most closely related to the structural changes of the molecule on ligand-binding, were large negative changes, suggesting that

major conformational changes have occurred between the apo- and holo-forms of this protein. In support of this interpretation, the crystal structures of apo- and holo-CML show that the side chains of Lys-62 and Trp-63 were rotated around the χ_1 angle and χ_2 angle, respectively, as if the hydrophobic interaction increases around the cluster, and the results were substantially consistent with the ^1H NMR spectra information. The results provide valuable insights about the structural and thermodynamic evidence of Ca^{2+} -binding protein.

Materials and Methods

Materials

According to the procedures of Koshiba et al (1999), expression and purification of CML were carried out using the expression system in *E. coli* (19). Mutant human lysozyme in which Glu-86 and Ala-92 were replaced with aspartic acid (HLD86/92) was prepared previously (20). All other reagents were of biochemical research grade.

Preparation of Apo- and Holo-proteins

Apo- and holo-proteins were prepared as described previously (20).

Isothermal Titration Calorimetry (ITC)

I performed isothermal titration calorimetry (ITC) to determine the enthalpy change upon the binding of Ca^{2+} to the CML, HLD86/92, and organic chelator (EDTA) using a MicroCal OMEGA titration calorimeter (21). Sample solutions for ITC measurements were prepared by dissolving lyophilized apo-lysozymes or EDTA (Ca^{2+} -free) in 50 mM sodium acetate buffer (pH 5.5) to make 0.1 mM solutions, and each 1.6 mL of this solution was injected into the cell and titrated with 5 μL portions of a 5 mM CaCl_2 solution at several temperatures. The calorimetric enthalpy (ΔH), the

binding constant (K_b), and the number of bound Ca^{2+} ions (n) were obtained by using the computer program ORIGIN (MicroCal, Inc., Northampton, MA).

Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectra were obtained on a JEOL α -500 spectrometer (JEOL Co., Tokyo, Japan) operated at 500 MHz and 20° C. Spectral assignments were supplemented by temperature titration and assignment results carried out at pH 4.5 and 30° C for holo-CML (18). Sample solutions for multi-dimensional NMR were dissolved in 10% $\text{D}_2\text{O}/\text{H}_2\text{O}$ to 1-1.5 mM, pH was adjusted to 4.5 by NaOH or HCl. For the sample of Ca^{2+} -saturated state (holo-CML), CaCl_2 was added to 10 mM. Conventional ^{15}N -edited spectra, i.e., HSQC (22), TOCSY-HSQC and NOESY-HSQC (23); were acquired. All ^{15}N -edited spectra were gradient-enhanced. The NOESY and TOCSY mixing times for all measurements were set at 200 ms and 80 ms, respectively. Sweep widths of 8000 Hz were used for the ^1H and of 1800 Hz for the ^{15}N dimensions, with the carrier frequency set at the water resonance. Data sets for ^{15}N -edited TOCSY and NOESY experiments typically comprised 512 (^1H) $\times$ 128 (^1H) $\times$ 64 (^{15}N) of 8 scan. Data sets for HSQC comprised 1024 (^1H) $\times$ 128 (^{15}N) of 8 scan, with a sweep width of 8000 Hz for the ^1H and of 2200 Hz for the ^{15}N dimensions. All proton chemical shifts were referenced to the methyl resonance of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) used as an internal standard. The ^{15}N chemical shift was indirectly referenced using the following ratios of the zero-point frequencies: 0.101 to ^1H (24). Data processing was performed on a Silicon Graphics Indy Work station using the program packages NMRpipe (25) and XEASY (26).

Crystallization, Data Collection, and Structure Refinement of Holo-CML

The crystallization of holo-CML was carried out by vapor diffusion of hanging drops using 30 mM potassium phosphate buffer (pH 5.8) containing 20 mg/mL protein, 2.1 M ammonium sulphate, and

1.5 mM CaCl_2 . After 3 weeks, crystals with the size of $0.6 \times 0.3 \times 0.05 \text{ mm}^3$ were obtained at 18° C, and the crystals were hemihedrally twinned. The space group was $P3_2$ with cell dimensions $a = b = 31.4 \text{ \AA}$, $c = 202.5 \text{ \AA}$ and $\gamma = 120^\circ$. There are two monomeric molecules in an asymmetric unit. The X-ray diffraction data of the holo-CML crystal was collected from a single crystal at room temperature using a Fuji imaging plate detector with a Rigaku IPR-4080 imaging scanner and a Weissenberg camera for macromolecular crystallography (27) on BL-18B at the Photon Factory, KEK, Japan. Integrated intensities were obtained using Denzo, and reduced using a Scalepack (28). Twin was detected using program CNS (29). The twin operator was two-fold along a^* (b^*)-axis and the reflection related to (h, k, l) was ($h+k, -k, l$), which gives rise to a diffraction pattern of pseudo Laue symmetry of $-31m$. The initial twin factor of 0.45 indicated that the crystals are nearly perfect hemihedral twinning. Since the two molecules in an asymmetric unit were related with a local two-fold axis parallel to the twin axis (a^* or b^*), twinned data was directly used for solving the structure by a molecular replacement method (30). The initial model of the holo-CML was obtained using the program AMoRe (31), and a search model was constructed from the atomic coordinates of apo-CML (17). The structural refinements were performed using the program CNS by perfect twinning protocol (30).

Determination of Protein Concentration

The protein concentrations were estimated spectrophotometrically using the following extinction coefficients at 280 nm: $E_{1\text{cm}}^{1\%} = 23.2$ for CML (16) and $E_{1\text{cm}}^{1\%} = 25.6$ for HLD86/92 (32).

Results

Isothermal Titration Calorimetry (ITC) of CML

Thermodynamic parameters for the binding of Ca^{2+} to CML, EDTA, and HLD86/92 were

investigated by isothermal titration calorimetry (ITC) using a MicroCal OMEGA calorimeter in 50 mM acetate buffer at pH 5.5 and several temperatures. This condition was chosen for the following reasons. At first, the pH of the solution is near to the pH conditions at which the tertiary structures for these proteins were determined (17, 33). Second, the intermolecular associations of lysozyme are negligible below this pH (34). Finally, the enthalpy change for ionization of acetic acid is so small that the proton release for the binding reaction of Ca^{2+} can be omitted (35).

In Figure 1, the typical titration process of the CML at 20° C is shown with the similar titration of EDTA and HLD86/92. The Ca^{2+} -binding constants (K_b), the numbers of bound Ca^{2+} ions (n), and enthalpy changes (ΔH) were obtained from the analysis of these curves by using the calorimetric analysis program ORIGIN (MicroCal Inc.), and the heat capacity changes (ΔC_p) of the Ca^{2+} -binding were estimated from the temperature dependence of the enthalpy changes. The Gibbs free energy changes ($\Delta G = -RT \ln K_b$) of Ca^{2+} -binding were calculated from the binding constants, and the entropy changes ($T\Delta S = -\Delta G + \Delta H$) of the binding could also be estimated. The results are summarized in Table 1.

For the binding of Ca^{2+} to EDTA and HLD86/92, the binding constants were determined to be $0.9\text{--}2 \times 10^5 \text{ M}^{-1}$ and $2\text{--}8 \times 10^5 \text{ M}^{-1}$, respectively, at the temperature region from 20 to 40° C. The contributions of ΔH for the binding of Ca^{2+} were unfavorable (2–3 kcal/mol). A large release of entropy term (9–11 kcal/mol as $T\Delta S$) was found to drive the Ca^{2+} -binding to EDTA and HLD86/92 in these conditions, as reported previously (36–38). From the temperature dependence of ΔH shown in Figure 2, the ΔC_p of EDTA and HLD86/92 were estimated to be 0.05 and 0 kcal/mol/K, respectively, by the least-squares fitting method.

In contrast, the thermodynamic parameters of CML showed different and interesting behaviors compared with those of EDTA and HLD86/92. The binding constants of Ca^{2+} for CML were $2\text{--}9 \times 10^6 \text{ M}^{-1}$ at the temperature region from 11 to 24° C, which were slightly higher than those

of EDTA and HLD86/92 (Table 1). Moreover, the contribution of ΔH for the binding of Ca^{2+} was favorable ($\Delta H = -5.1 \pm 1.3 \text{ kcal/mol}$) compared with those of the entropy terms ($3.9 \pm 1.1 \text{ kcal/mol}$ as $T\Delta S$). Therefore, the binding Ca^{2+} for CML is primarily driven by enthalpic contributions. The ΔC_p values of CML were estimated to be -0.21 kcal/mol/K , suggesting a prominent hydrophobic effect (39). Thus, the results indicate that there is significant conformational change accompanying Ca^{2+} binding in the case of CML, although no significant conformational change was seen in the HLD86/92 (33).

¹H-¹⁵N HSQC Spectra of CML

Figure 3A and B show the ¹H-¹⁵N HSQC spectra of apo- and holo-CML at pH 4.5 and 20° C. Due to the solubility problem of the sample, NMR measurements were carried out at pH 4.5 (solubility > 1mM) instead of pH 5.5 (solubility < 0.6mM). In the apo and holo-CML, most peaks are overlapped between two species. This indicates that overall structure of apo- and holo-CML are quite similar, although some spectral differences can be observed (see Figure 3C and D). One of the major differences between the ¹H-¹⁵N HSQC spectra of apo- and holo-CML is that the chemical shift of indole amide protons (Trp-63 and Trp-64) are slightly shifted to the upfield and the downfield region respectively, and indole amide proton of Trp-108 is largely shifted to the downfield region from 10.1 to 10.4 ppm (0.3 ppm) as a result of the addition of Ca^{2+} ions. Moreover, the chemical shifts of main-chain amide protons of Asp-91 and Ile-56 were also shifted to the downfield and upfield regions, respectively. Most of these residues are located near the Ca^{2+} -binding site and the hydrophobic cluster (in the β -domain) consisted from Lys-62, Trp-63, and Trp-64, which has been observed in the X-ray crystal structure of apo-CML (17). These shifts also indicate that a conformational change has mainly occurred in the Ca^{2+} -binding site and β -domain of CML. Thus, the addition of Ca^{2+} ions for apo-CML is not only restricted a significant conformational change in

the Ca^{2+} binding site but is also affected for the hydrophobic cluster consisted from Lys-62, Trp-63, and Trp-64 in the β -domain.

X-ray Crystal Structure of holo-CML

The 1.9 Å structure of holo-CML was determined by a molecular replacement method and was refined to an *R*-factor of 15.6 % ($R_{\text{free}} = 25.7$ %). The crystallographic parameters and refinement statistics are summarized in Table 2.

Figure 4 shows the ribbon diagram of holo-CML in complex with a Ca^{2+} ion. A deep cleft containing the active site divides the CML molecule into a large α -helix rich domain (α -domain) and a small β -sheet rich domain (β -domain), both of which are quite similar to the domains previously reported for the apo-CML (17). The root mean square deviation (r.m.s.d.) between apo- and holo-CML for the main chain C α atoms was 0.879 Å.

The local structures in the vicinity of the Ca^{2+} -binding site of apo- and holo-CML are shown in Figure 5. The Ca^{2+} -binding site in the holo-CML structure is essentially identical in terms of both conformation and ligand coordination to the binding site previously reported in the various α -lactalbumins and Ca^{2+} -binding lysozymes (33, 40–43). The Ca^{2+} ion is coordinated by a constellation of seven oxygen groups that form a distorted pentagonal bipyramid (holo-form). Five of the ligating groups are contributed by the main chain (Lys-82 and Asn-87) and side chains (Asp-85, -90, and -91). The remaining coordinating groups are supplied by two water molecules. All the Ca^{2+} -protein ligand (Ca-O) distances are between 2.3 and 2.7 Å, corresponding to a tight Ca^{2+} -binding loop in the structure. However, in the structure of the apo-CML, the side chains of Asp-85 and -90 are rotated around the χ_1 angle toward the solvent, as if to reduce the negative charge repulsion in the Ca^{2+} -binding site. The water molecules occupied the same position as the bound Ca^{2+} . The result was similar to the previous reports of several apo-forms of Ca^{2+} -binding

lysozymes (33, 38, 44, 45), which is thought to relieve the destabilization of apo-protein.

Discussion

The Ca^{2+} -binding reaction sometimes accompanies conformational changes of the protein between the Ca^{2+} -free- (apo-) and Ca^{2+} -bound-forms (holo-form), which are considered to be responsible for biological functions (6–8). The purpose of this CHAPTER is to elucidate the relationship between the structure and thermodynamics of Ca^{2+} -induced conformational changes of a Ca^{2+} -binding protein. For this purpose, the thermodynamic change of CML was investigated using ITC and interpreted in terms of structural information obtained from NMR spectroscopy and X-ray crystallography.

The results of ITC measurement showed that the thermodynamic change upon the binding of Ca^{2+} to CML showed different behaviors from those of binding to EDTA and HLD86/92. The Ca^{2+} -binding reaction to CML was an exothermic process, and the enthalpic contribution for the reaction was larger than that of the entropy; the binding reaction to EDTA and HLD86/92 were endothermic processes, and the contributions were driven primarily by entropy terms (Table 1). Previously, Kuroki and co-workers (1992) revealed that the binding of Ca^{2+} to EDTA and HLD86/92 displayed similar thermodynamic behaviors, and this similarity may indicate that the interaction between Ca^{2+} and ligand residues does not depend on the kind of ligand (37). Kuroki et al also found that the conformational change of HLD86/92 contributes little to the thermodynamics parameters observed upon the binding of Ca^{2+} by the results of ITC and X-ray crystallography (33, 37, 38). Therefore, the enthalpy term (ΔH^*) of CML, EDTA, and HLD86/92, shown in Table 3, was obtained by subtracting the thermodynamic contribution of Ca^{2+} -binding observed in EDTA as the intrinsic contribution of Ca^{2+} -binding alone. The resulting value should correspond to the contribution of the conformational change of the proteins upon Ca^{2+} -binding. The ΔH^* of CML was a large negative value (-8.8 kcal/mol), although that of HLD86/92 was close to zero (0.6

kcal/mol). This indicates that there are significant conformational changes not only in the Ca^{2+} -binding site but also in the other region of CML accompanying Ca^{2+} -binding. The thermodynamic changes (ΔH and ΔC_p) that occur when Ca^{2+} binds to other Ca^{2+} -binding proteins (46–53) determined in previous studies are summarized in Table 3. All of these proteins show large negative ΔH and ΔC_p values as seen in those of CML. In most ligand binding processes result in negative values for ΔC_p , which are accounted for by a decrease in exposed surface, possibly due to tightening of the holo-structure. My results following the ITC measurement of CML may indicate that a substantial area of the hydrophobic groups are sequestered from the surface to the interior of the molecule.

The structural analysis using ^1H - ^{15}N HSQC spectra and X-ray diffraction supported the thermodynamic data obtained by calorimetry well. As shown in Figure 3A and B, the chemical shift of indole amide proton (Trp-64) and main-chain amide protons of Asp-91 are largely shifted to the downfield region from 11 to 10 ppm as a result of the addition of Ca^{2+} ions. In the aliphatic region, the chemical shift of γH (Lys-62) is also shifted to the upfield region from 0 to -1 ppm as a result of the addition of Ca^{2+} ions (data not shown). The residues of Lys-62 and Trp-64 are located in the β -domain and exist on the opposite side of the Ca^{2+} -binding site (Figure 4). The NMR studies suggest that the conformational change between apo- and holo-CML has occurred not only in the Ca^{2+} -binding site but also around the region of Lys-62 and Trp-64.

In the crystal structure of apo- and holo-CML, Lys-62, Trp-63, Trp-64, and Val-98 form a hydrophobic cluster, as shown in Figure 6A and B. However, the Ca^{2+} ion influences this hydrophobic cluster, and it induced a conformational change between the apo- and holo-CML at this region. In the structure of holo-CML (Figure 6B), the side chains of Lys-62 and Trp-63 were rotated around the χ_1 angle (100°) and χ_2 angle (120°), respectively, as if the hydrophobic interaction increases around the cluster. The value of the accessible surface area (ASA) of

holo-CML in the cluster was about $218 \text{ \AA}^2$, which is smaller than that of the apo-CML ($252 \text{ \AA}^2$). The hydrophobic contacts caused by the Ca^{2+} -binding reaction of CML contribute about $30 \text{ \AA}^2$ to the hydrophobic interaction. Moreover, the distance from the C^δ atom of Lys-62 to the C^γ atom of Val-98 is $11.5 \text{ \AA}$ at holo-CML, although that of the apo-CML is $13.7 \text{ \AA}$. We can see it like a "burst into blossom" at the structure of holo-CML and as a "fade-away blossom" at the structure of apo-CML. These conformational changes are consistent with the thermodynamic results obtained by ITC measurements. Therefore, I believe that the large negative ΔC_p value of CML, which is accounted for by a decrease in the exposed surface, is caused by the decrease of the exposed surface area ($\Delta\text{ASA} = 30 \text{ \AA}^2$) to the solvent and estimated to be about -0.21 kcal/mol/K . This is one of the main reasons why the Ca^{2+} -binding reaction to CML is an exothermic process and why the ΔC_p is a negative value.

This is the first example of a Ca^{2+} -binding protein in which the Ca^{2+} -induced conformational change was thoroughly investigated by using thermodynamic study, NMR spectroscopy, and X-ray crystallography. I hope our structural and thermodynamic information will facilitate the study of the ligand-binding mechanism.

References

1. Ebashi, S., and Endo, M. (1968) *Prog. Biophys. Mol. Biol.*, **18**, 123-183.
2. Applebury, M. L., and Coleman, J. E. (1969) *J. Biol. Chem.*, **244**, 308-318.
3. Hill, R. L., and Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol.*, **43**, 411-490.
4. Potter, J. D., and Gergely, J. (1975) *J. Biol. Chem.*, **250**, 4628-4633.
5. Levine, B. A., and Williams, R. J. P. (1982) *Calcium and Cell Function* (Cheung, W.-Y., ed) Vol. II, pp. 1-38, Academic Press, New York.
6. Nemoto, Y., Ikeda, J., Katoh, K., Koshimoto, H., Yoshihara, Y., and Mori, K. (1993) *J. Cell Biol.*,

123, 963-976.

7. Dell' Angelica, E. C., Schleicher, C. H., and Santome, J. A. (1994) *J. Biol. Chem.*, **269**, 28929-28936.
8. Hughes, R. E., Brzovic, P. S., Klevit, R. E., and Hurley, J. B. (1995) *Biochemistry*, **34**, 11410-11416.
9. Ikura, M., Hiraoki, T., Hikichi, K., Mikuni, T., Yazawa, M., and Yagi, K. (1983) *Biochemistry*, **22**, 2573-2579.
10. Barbato, G., Ikura, M., Key, L. E., Pastor, R. W., and Bax, A. (1992) *Biochemistry*, **31**, 5269-5278.
11. Ikura, M., Barbato, G., Klee, C. B., and Bax, A. (1992) *Cell Calcium*, **13**, 391-400.
12. Finn, B. E., Drakenberg, T., and Forsén, S. (1993) *FEBS Lett.*, **36**, 368-374.
13. Zhang, M., Tanaka, T., and Ikura, M. (1995) *Nature Struct. Biol.*, **2**, 758-767.
14. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren H., Klee, C. B., and Bax, A. (1995) *Nature Struct. Biol.*, **4**, 768-776.
15. Finn, B. E., Evenäs, J., Drakenberg, T., Waltho, J. P., Thulin, E., and Forsén, S. (1995) *Nature Struct. Biol.*, **4**, 777-783.
16. Kikuchi, M., Kawano, K., and Nitta, K. (1998) *Protein Sci.*, **7**, 2150-2155.
17. Koshiba, T., Yao, M., Kobashigawa, Y., Demura, M., Nakagawa, A., Tanaka, I., Kuwajima, K., and Nitta, K. (2000) *Biochemistry*, **39**, 3248-3257.
18. Kobashigawa, Y., Demura, M., Koshiba, T., Kumaki, Y., Kuwajima, K., and Nitta, K. (2000) *Proteins: Struct. Funct. Genet.*, **40**, 579-589.
19. Koshiba, T., Hayashi, T., Ishido, M., Kumagai, I., Ikura, T., Kawano, K., Nitta, K., and Kuwajima, K. (1999) *Protein Eng.*, **12**, 429-435.
20. Koshiba, T., Tsumoto, K., Masaki, K., Kawano, K., Nitta, K., and Kumagai, I. (1998) *Protein*

Eng., **11**, 683-690.

21. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L.-N. (1989) *Anal. Biochem.*, **179**, 131-137.
22. Bax, A., Griffey, R. H., and Hawkins, B. L. (1983) *J. Magn. Reson.*, **55**, 301-315.
23. Kay, L. E., Marion, D., and Bax, A. (1989) *J. Magn. Reson.*, **84**, 72-84.
24. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1989) *J. Mol. Biol.*, **280**, 933-952.
25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277-293.
26. Bartels, C., Xia, T.-H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1-10.
27. Sakabe, N., Ikemizu, S., Sakabe, K., Higashi, T., Nakagawa, A., Watanabe, N., Adachi, S., and Sasaki, K. (1995) *Rev. Sci. Instrum.*, **66**, 1276-1281.
28. Otwinowski, Z. (1993) pp. 80-86, SERC Daresbury Laboratory, Warrington, UK.
29. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstlere, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr.*, **D54**, 905-921.
30. Yao, M., Koshiba, T., Nitta, K., and Tanaka, I. (2000) manuscript preparing.
31. Navaza, J. (1994) *Acta Crystallogr.*, **A50**, 157-163.
32. Parry, R. M., Chandan, R. C., and Shahani, K. M. (1969) *Arch. Biochem. Biophys.*, **130**, 59-65.
33. Inaka, K., Kuroki, R., Kikuchi, M., and Matsushima, M. (1991) *J. Biol. Chem.*, **266**, 20666-20671.
34. Sophianopoulos, A. J. and van Holde, K. E. (1964) *J. Biol. Chem.*, **239**, 2516-2524.
35. Fukada, H. and Takahashi, K. (1998) *Proteins: Struct. Funct. Genet.*, **33**, 159-166.
36. Kuroki, R., Kawakita, S., Nakamura, H., and Yutani, K. (1992) *Proc. Natl. Acad. Sci. USA*, **89**,

6803-6807.

37. Kuroki, R., Nitta, K., and Yutani, K. (1992) *J. Biol. Chem.*, **267**, 24297-24301.
38. Kuroki, R. and Yutani, K. (1998) *J. Biol. Chem.*, **273**, 34310-34315.
39. Tanford, C. (1973) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, John Wiley & Sons, New York.
40. Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., and Phillips, D. C. (1986) *Nature*, **324**, 84-87.
41. Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., and Fenna, R. E. (1991) *J. Mol. Biol.*, **221**, 571-581.
42. Pike, A. C. W., Brew, K., and Acharya, K. R. (1996) *Structure*, **4**, 691-703.
43. Guss, J. M., Messer, M., Costello, M., Hardy, K., and Kumar, V. (1997) *Acta Crystallogr.*, **D53**, 355-363.
44. Yao, M., Tanaka, I., Hikichi, K., and Nitta, K. (1992) *J. Biochem.*, **111**, 1-3.
45. Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S., and Miyano, M. (1992) *J. Biochem.*, **111**, 141-143.
46. Griko, Y. V. (1999) *Biophys. Chem.*, **79**, 117-127.
47. Tanokura, M., and Yamada, K. (1987) *Biochemistry*, **26**, 7668-7674.
48. Tanokura, M. (1990) *Eur. J. Biochem.*, **188**, 23-28.
49. Kometani, K., and Yamada, K. (1983) *Biochem. Biophys. Res. Commun.*, **114**, 162-167.
50. Yamada, K., and Kometani, K. (1982) *J. Biochem.*, **92**, 1505-1517.
51. Sellers, P., Laynez, J., Thulin, E., and Forsén, S. (1991) *Biophys. Chem.*, **39**, 199-204.
52. Tanokura, M., and Yamada, K. (1986) *J. Biol. Chem.*, **261**, 10749-10752.
53. Griko, Y. V., and Remeta, D. P. (1999) *Protein Sci.*, **8**, 554-561.
54. Kraulis, P. J. (1991) *J. Appl. Cryst.*, **24**, 946-950.

Table 1. Thermodynamic changes in the binding of Ca^{2+} to CML, HLD86/92, and EDTA at pH 5.5 and several temperatures

sample	temp. (°C)	<i>n</i>	K_b (M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
CML	11.6	0.9	1.91×10^6	-8.18	-3.78	4.40
	16.1	0.9	9.29×10^6	-9.22	-4.20	5.02
	19.1	0.9	5.29×10^6	-8.99	-4.84	4.15
	20.2	0.9	2.23×10^6	-8.52	-5.09	3.43
	22.2	0.9	6.08×10^6	-9.16	-5.71	3.45
	24.1	0.9	5.03×10^6	-9.11	-6.38	2.73
HLD86/92	20.0	0.9	2.04×10^5	-7.12	2.95	10.07
	25.2	0.9	4.23×10^5	-7.68	2.97	10.65
	30.2	0.8	5.87×10^5	-8.00	3.15	11.15
	40.2	0.8	8.13×10^5	-8.47	2.85	11.32
EDTA	11.3	1.1	7.03×10^4	-6.31	1.58	7.89
	15.8	1.0	6.27×10^4	-6.34	2.05	8.39
	20.2	1.0	1.10×10^5	-6.76	2.15	8.91
	24.2	1.0	8.70×10^4	-6.72	2.37	9.09
	30.1	0.9	2.21×10^5	-7.41	2.50	9.91
	40.2	0.9	2.24×10^5	-7.67	3.11	10.78

Table 2. Crystallographic data, X-ray processing statistics, and refinement statistics of the holo-form CML

<i>Crystal data collection</i>	
Space group	<i>P</i> 32
Cell constants (Å)	
<i>a</i>	31.3
<i>b</i>	31.3
<i>c</i>	202.5
Resolution (Å)	100-1.90 (1.97-1.90)
No. of measured reflections	43072 (3326)
No. of independent reflections	16169 (1423)
Averaged <i>I</i> / $\sigma(I)$	11.1 (3.0)
^a <i>R</i> _{merge} (%)	7.5 (29.0)
<i>Refinement</i>	
Resolution used (Å)	10.0-1.90
No. of used reflections (<i>F</i> > 2 σ)	15,524
Data completeness (%)	89.2
Δ bond length (Å)	0.006
Δ bond angle (°)	1.338
average <i>B</i> -factor (Å ²)	
main-chain	27.1
side-chain	28.9
^b <i>R</i> -factor (%)	15.9
^c Free <i>R</i> -factor (%)	25.7

$$^a R_{\text{merge}} = \Sigma |I - \{I\}| / \Sigma \{I\} \times 100\%.$$

$$^b R\text{-factor} = \Sigma |F_o| - |F_c| / \Sigma |F_c| \times 100\%.$$

^cFree *R*-factor is the same as *R*-factor, except for a 10% subset of all reflections.

Table 3. Enthalpy and heat capacity changes of several Ca^{2+} binding proteins upon a binding of Ca^{2+}

Protein	Buffer (pH)	temp. °C	ΔH kcal/mol	ΔH^* kcal/mol	ΔC_p kcal/mol/K	Ref.
EDTA	AcOH-Na (5.5)	24	2.37	0	0.05	this study
	Tris-HCl (7.5)	25	12.1	0	0.02	46
	MOPS (7.5)	25	5.70	0	0.02	46
	PIPES (7.5)	25	3.24	0	0	46
HLD86/92	AcOH-Na (5.5)	25	2.97	0.6	0	this study
CML	AcOH-Na (5.5)	24	-6.38	-8.75	-0.21	this study
Bullfrog parvalbumin						
α -type	PIPES (7.5)	25	-3.87	-7.11	-0.16	47
β -type	PIPES (7.5)	25	-7.93	-11.17	-0.24	47
Toad parvalbumin						
Site 1	PIPES (7.5)	25	-13.14	-16.38	-0.57	48
Site 2	PIPES (7.5)	25	-5.26	-8.5	-0.14	48
Bovine cardiac troponine C	PIPES (7.0)	25	-5.64	-8.88	-0.21	49
Rabbit skeletal troponine C						
Site 1	PIPES (7.0)	25	-13.5	-16.73	-0.43	50
Site 3, 4	PIPES (7.0)	25	-1.86	-5.1	-0.05	50
Bovine brain calmodulin						
normal	HEPES (7.0)	25	-9.8	N. D.	-0.38	51
TFP-complex	PIPES (7.0)	25	-6.43	-9.67	-0.12	52
Bovine α -lactalbumin	Tris-HCl (8.0)	5	-1.7	-13.79	N. D.	53

N. D., Not detected

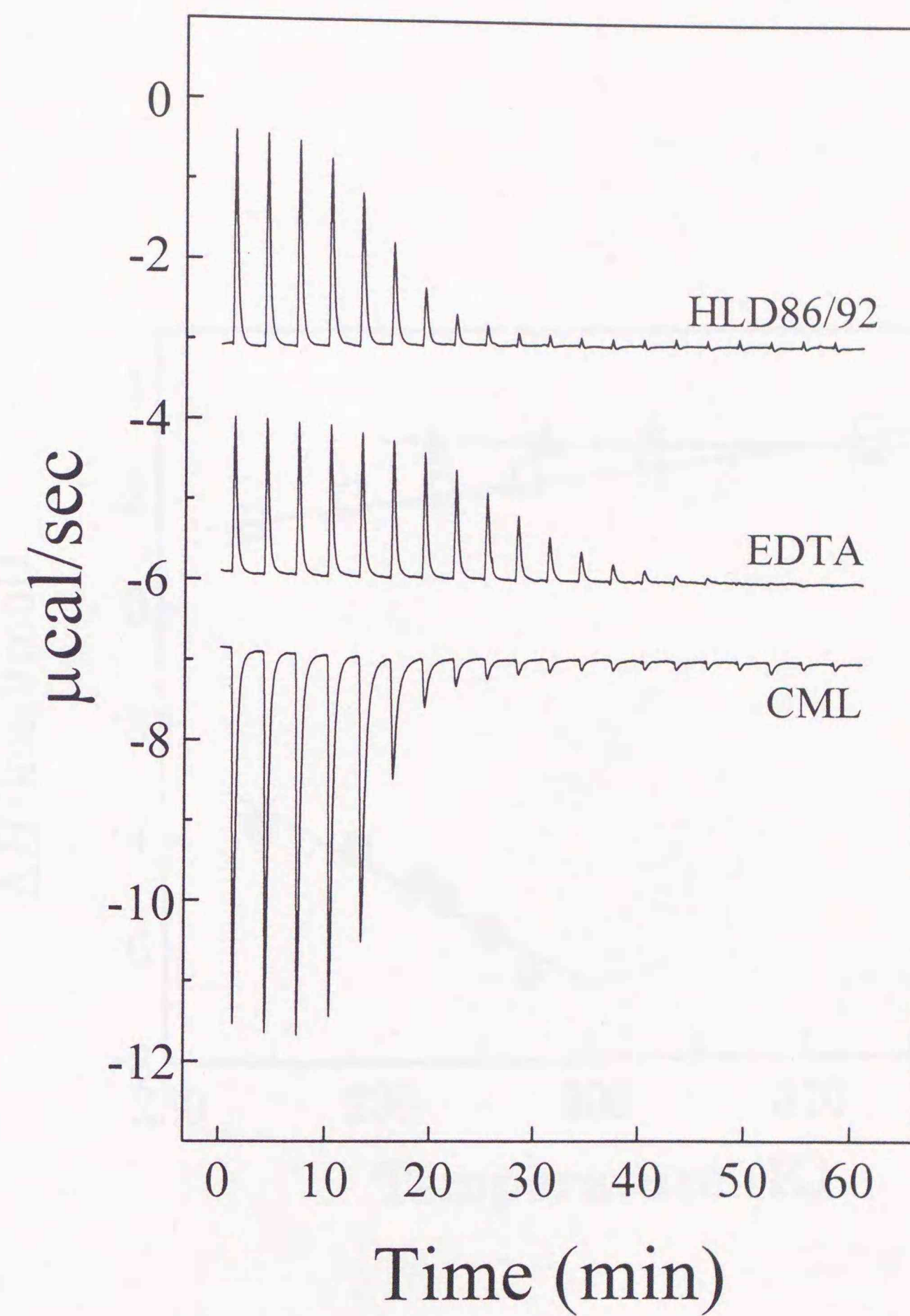


Figure 1. Typical calorimetric titration of CML, EDTA, and HLD86/92 in 50 mM sodium acetate buffer at 20° C and pH 5.5. The sample solutions (about 0.1 mM) were titrated with 50 mM sodium acetate buffer containing 5 mM CaCl_2 .

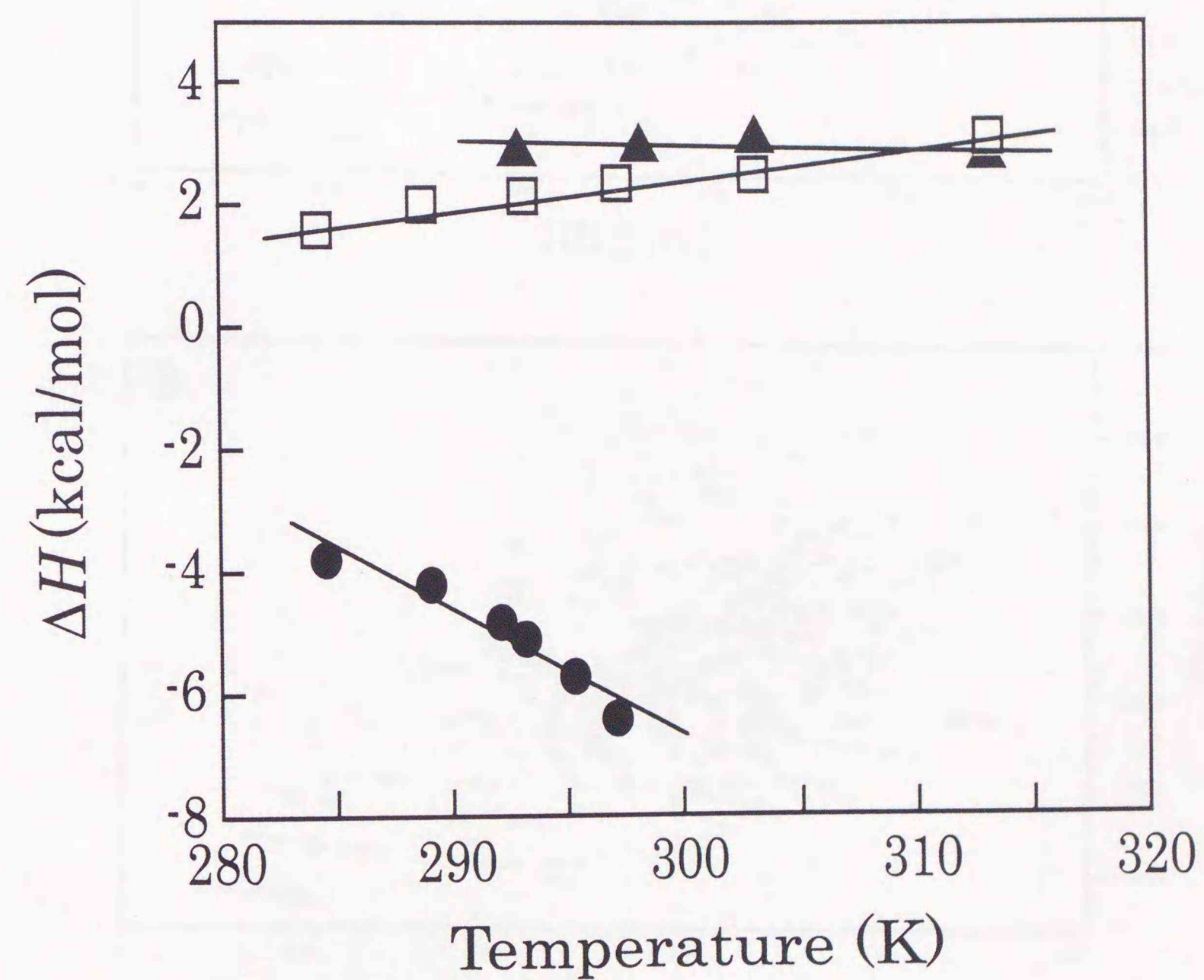
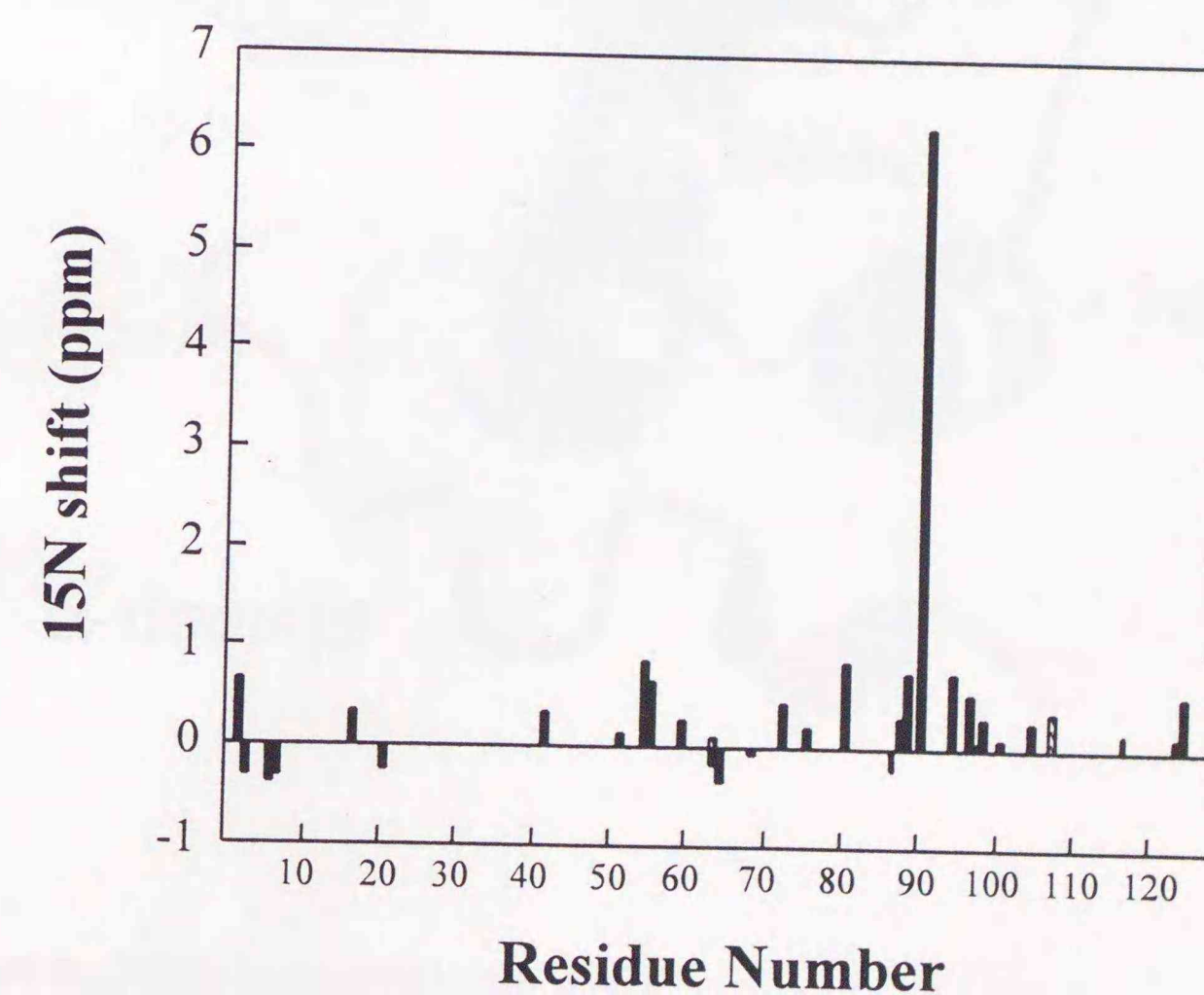
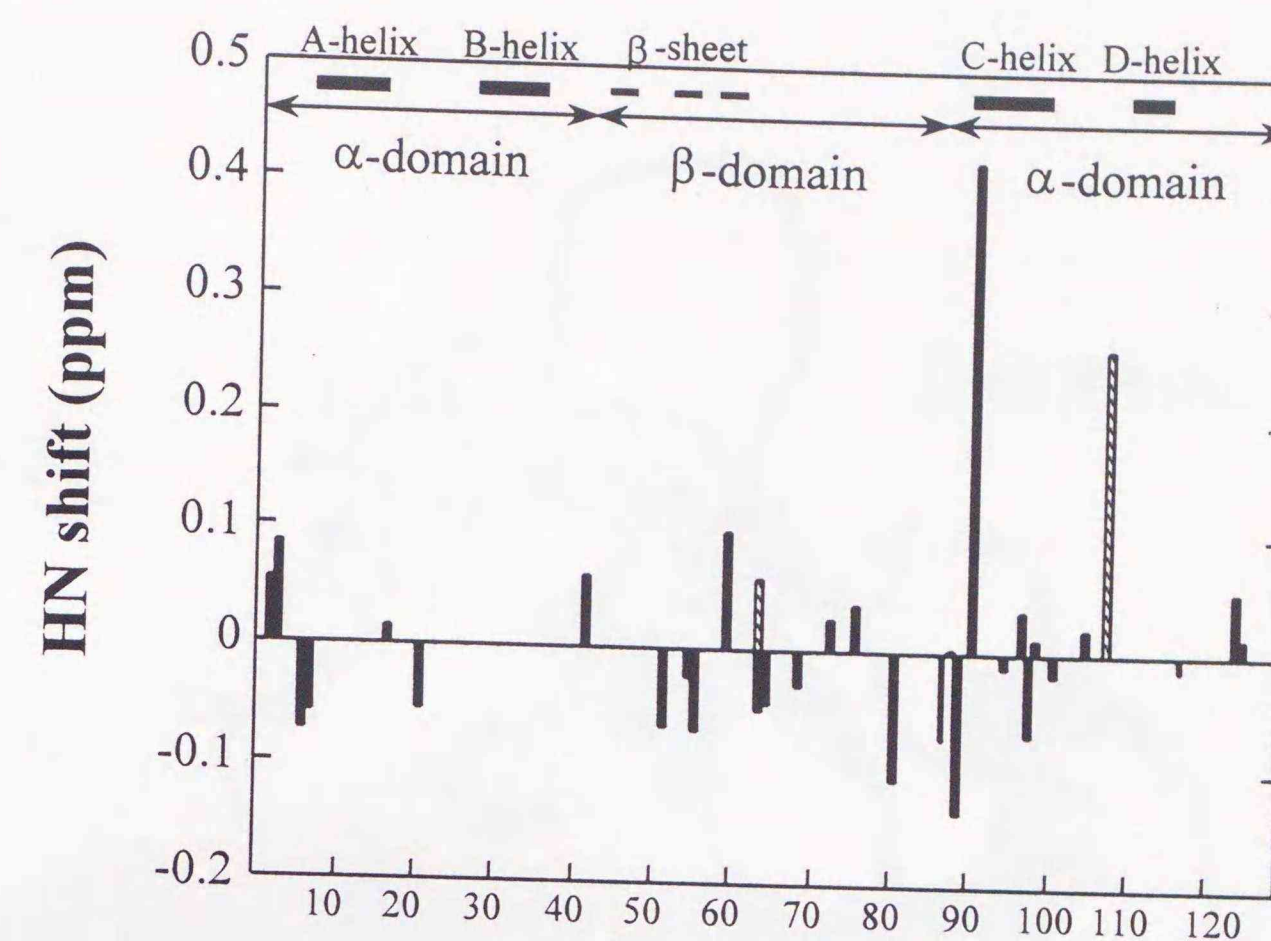


Figure 2. Temperature dependence of the enthalpy change for the binding of Ca^{2+} of CML (filled circles), EDTA (open squares), and HLD86/92 (filled triangles). The straight lines were obtained by a least-squares method for each sample.

C



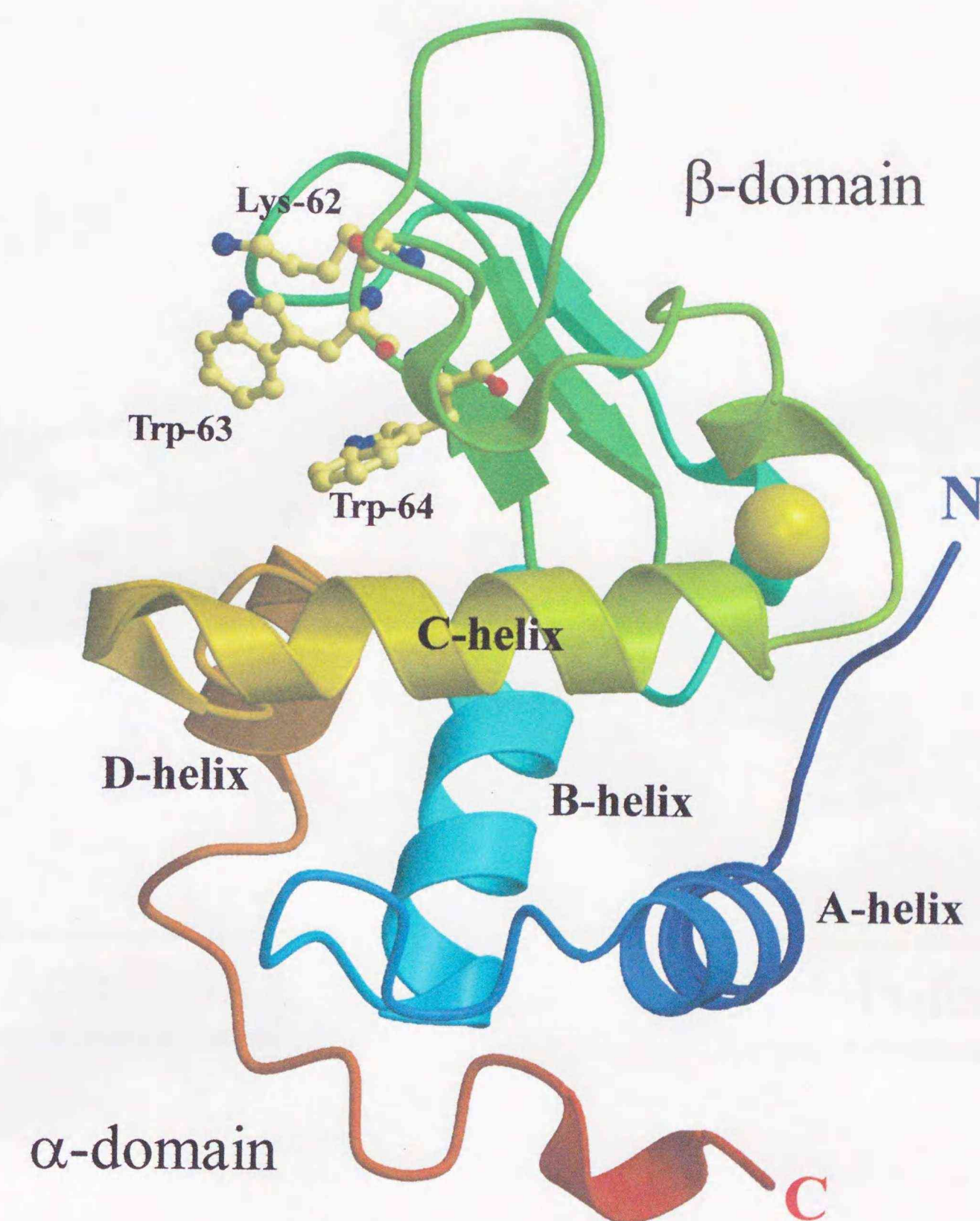


Figure 4. Ribbon model of holo-CML in which α -helices are sequentially labeled from A to D. The yellow sphere represents a Ca^{2+} ion. The structure shows a rainbow color (red to blue) from the N-terminus to the C-terminus (figure produced using MOLSCRIPT (54)).

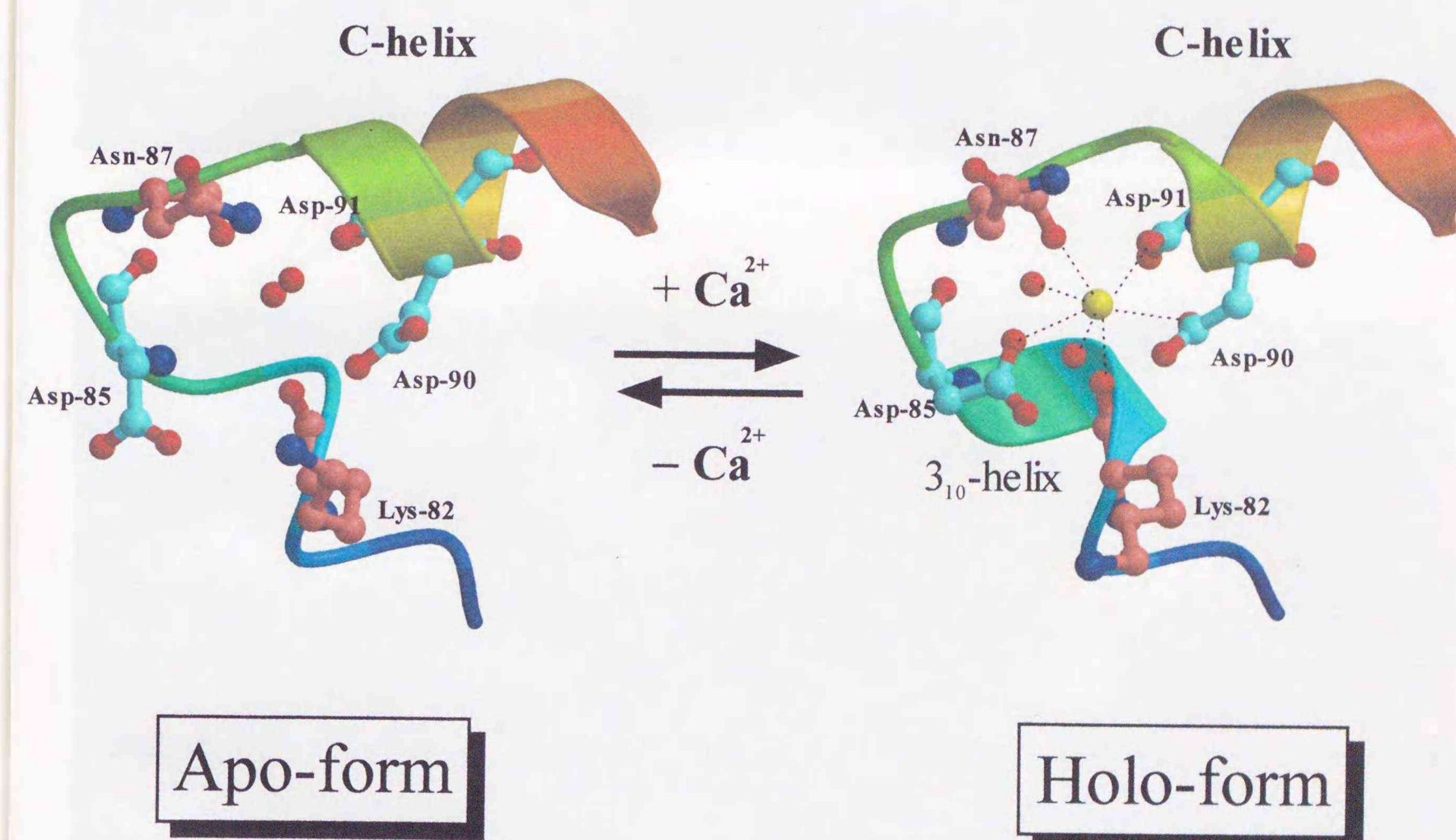
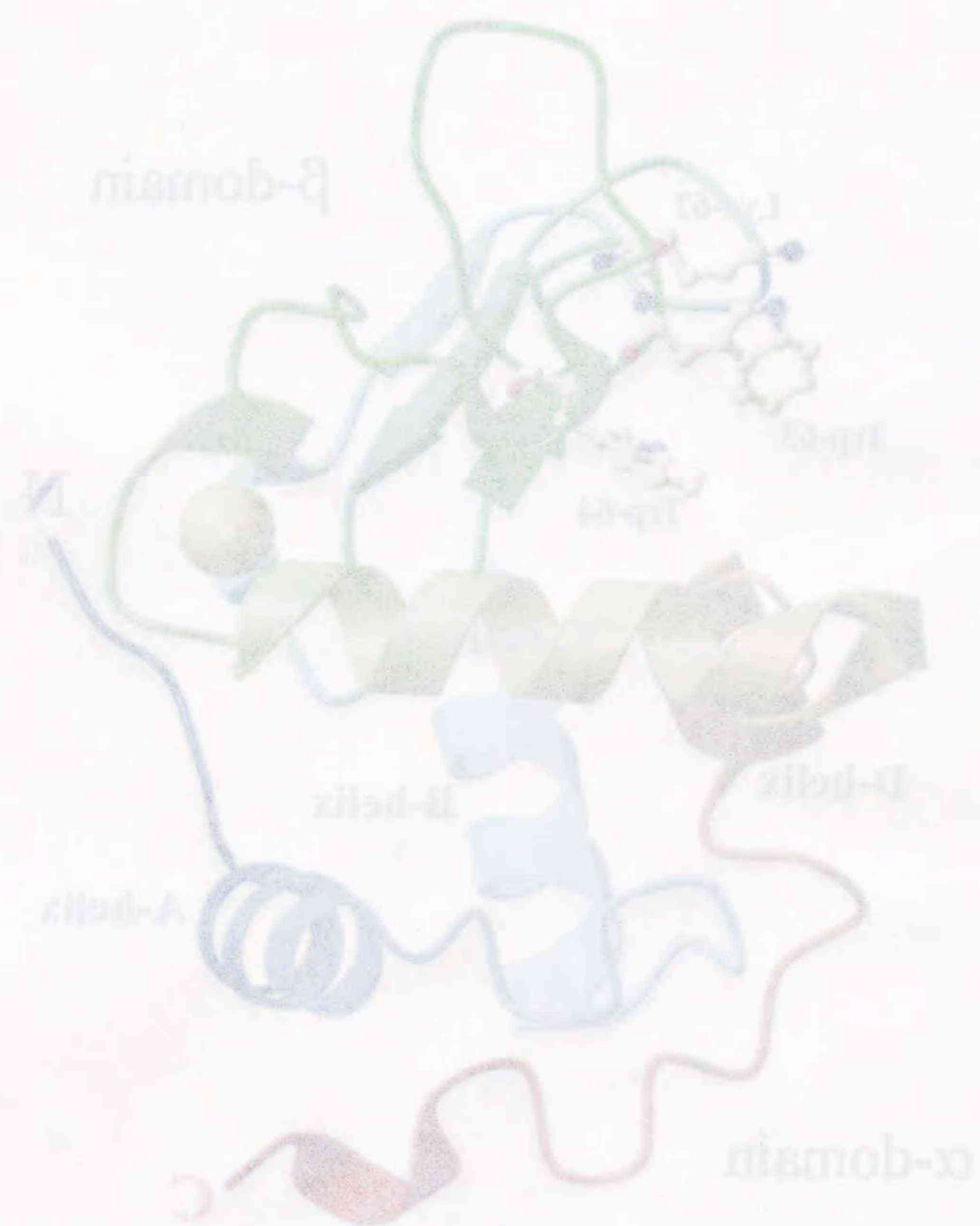


Figure 5. Comparison of the Ca^{2+} binding sites of apo- and holo-CML. The yellow sphere represents a Ca^{2+} ion, and red spheres represent the water molecules. Only side and main chains interacting with Ca^{2+} or those important for maintaining the Ca^{2+} -binding site are shown in fully labeled (figure produced using MOLSCRIPT (54)).

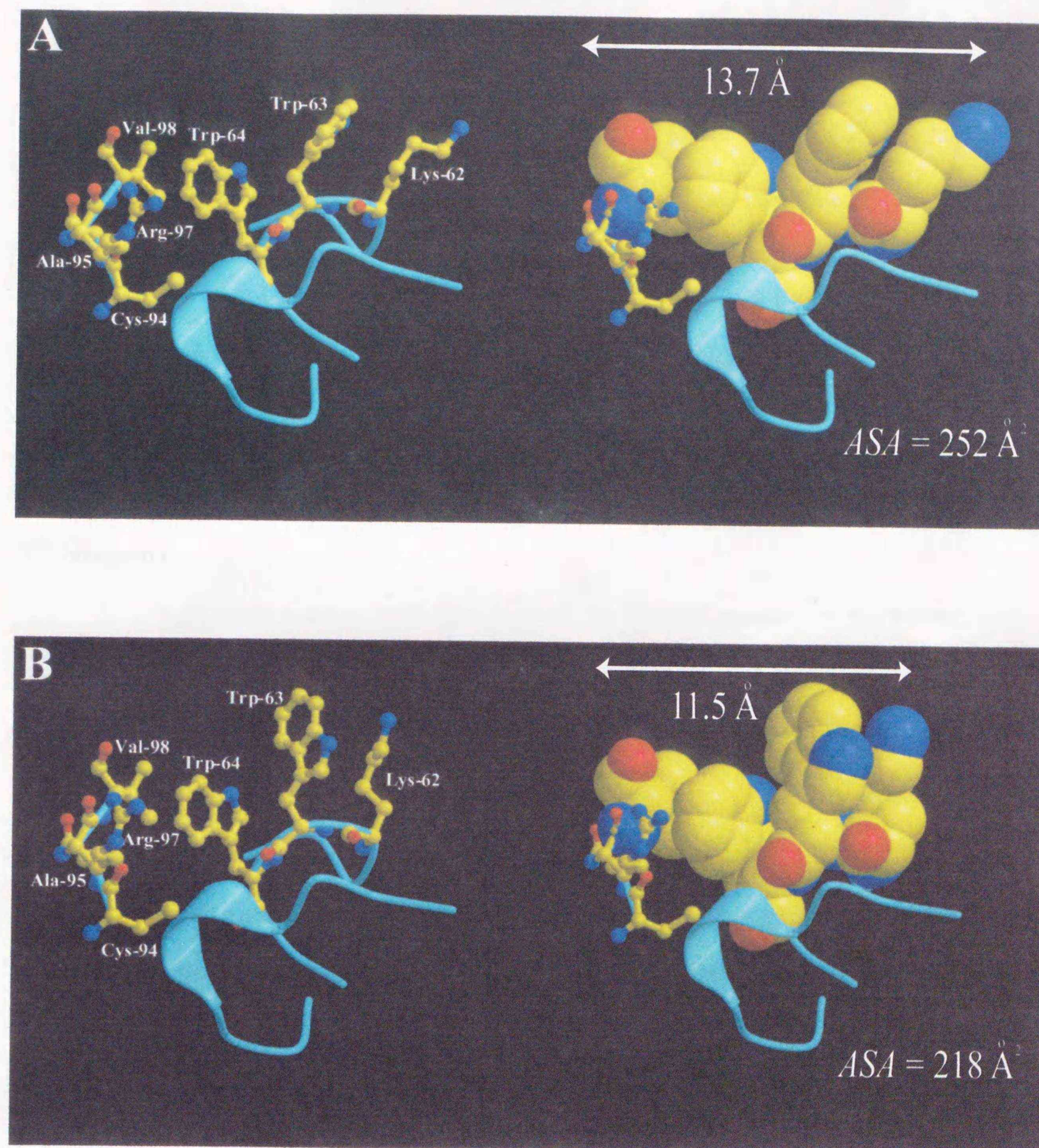


Figure 6. Conformation of a hydrophobic cluster around Trp-63 and -64. The local structures are shown schematically for (A) apo-CML, and (B) holo-CML. Two molecules are viewed in the same orientation and right structure of each figure drawing in CPK models (figure produced using MOLSCRIPT (54)).



ACKNOWLEDGEMENTS

I am deeply indebted to Professor Katsutoshi NITTA in Graduate School of Science, Hokkaido University for a lot of helpful supports, suggestions and continuous guidance.

I also wish to express my sincere gratitude to Professors Makoto DEMURA in Graduate School of Science, Hokkaido University, Kunihiro KUWAJIMA of Department of Physics, School of Science, University of Tokyo, Keiichi KAWANO of Department of Structural Biology, Faculty of Pharmaceutical Sciences, Toyama Medical and Pharmaceutical University, and Isao TANAKA in Graduate School of Science, Hokkaido University for their helpful supports and valuable discussions.

I am deeply indebted to Professor Izumi KUMAGAI of Department of Biochemistry and Engineering, Graduate School of Engineering, Tohoku University, for his helpful advice and discussions about preparation of the protein.

I would deeply like to thank Dr. Ryota KUROKI of Central Laboratories for Key Technology, Kirin Brewery Co., Ltd. for his valuable advices and helpful collaborations.

I am indebted to Dr. Min YAO in Graduate School of Science, Hokkaido University and Dr. Atsushi NAKAGAWA of Reserch Center for Structural Biology, Institute for Protein Research, Osaka University for their valuable comments and technical assistance on protein crystallography.

I am grateful to Mr. Yoshihiro KOBASHIGAWA for his valuable comments and suggestions of the measurement of NMR spectroscopy.

Mr. Tomohiro HAYASHI, Miss Miwako ISHIDO, and Dr. Teikichi IKURA of Department of Physics, School of Science, University of Tokyo is acknowledged for construction of the synthetic canine milk lysozyme gene.

Finally, I heartily thank my parents (Tesshu & Setsuko) and my friends for their kind assistance.

This work is supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan and also by a for JSPS Research Fellowships for Young Scientists.

March 1, 2001

Ⓒ Takumi KOSHIBA

*Division of Biological Sciences
Graduate School of Science
Hokkaido University
Kita-ku, Sapporo
JAPAN*

