Ultraviolet Resonance Raman Studies on Myoglobin Dynamics: Ligand Binding and Protein Folding

Nami Haruta

Department of Functional Molecular Science
School of Mathematical and Physical Science
The Graduate University for Advanced Studies

2001

PREFACE

Since tertiary structures of proteins construct on the base of a large amount of noncovalent interactions such as hydrophobic interaction, electrostatic interaction and hydrogen bonding, protein structures are flexible and fluctuating in a solution. The fluctuation of protein allows a conformational change with a rapid response to an environment change. One of major environment changes for proteins is a ligand binding. Reaction processes of proteins have started from the binding and provoke the conformational changes. Although the fluctuations of proteins are certain to be essential for these functions, it is difficult to combine the fluctuations to functions with apparently evidence. Vibrational spectroscopy is one of the strong methods to challenge this problem because of its specificity and sensitivity. In this thesis, author used the ultraviolet resonance Raman (UVRR) spectroscopy to detect the conformational changes of myoglobin during the ligand binding and protein folding. Since process of protein folding is increasing the noncovalent interactions and reducing the fluctuations, the understanding of folding mechanism will predict the fluctuations and the functions of the native structures.

This thesis consists of three charters. The general backgrounds of the fluctuation of proteins and UVRR spectroscopy are described in chapter 1. Chapter 2 contains the overview of heme proteins respecting the conformational change and the results of UVRR studies that detected the slightly conformational change of myoglobin during the ligand binding and selectively the iron-coordinated Tyr in the H64Y mutant. In chapter 3, UVRR spectroscopy was applied to equilibrium and kinetic folding reactions and provided a new aspect of folding process of apomyoglobin.

ACKNOWLEDGMENT

This thesis is a summary of the author's studies from 1998 to 2001 at the Department of Functional Molecular Science, the Graduate University for Advanced Studies. This work is carried out under the supervision of Prof. Teizo Kitagawa at Institute of Molecular Science (IMS). Author would like to express her cordial gratitude for his constant guidance, criticism, and encouragement. The author is also obliged to Dr. Yasuhisa Mizutani for his fruitful discussion and acute suggestion. She sincerely owes it to Dr. Michihiko Aki that this work could not execute without his careful guidance in spectroscopic technique.

It should be emphasized that the studies in this thesis have required the collaboration with a number of groups. Author is grateful to Prof. Yoshihito Watanabe (IMS) and his group members, especially Dr. Shin-ichi Ozaki, Dr. Seiji Ogo and Dr. Toshitaka Matsui for their kindly assistance and advice in constructing the mutants of myoglobin and on the various occasions. She is deeply indebted to Dr. Satoshi Takahashi (Kyoto univ.) for his aid in the construction of a rapid mixer. In particular, the rapid mixer was made up by the laborious and faithful work of Mr. Kazuhiro Kobayashi (IMS). Her thanks are also due to the staffs and friends of IMS for many advises and kindly encouragement.

This work would not have been possible without help of the members and coworkers in the Prof. Kitagawa's group. The author wishes to thank Mr. Shigenori Nagatomo, Dr. Takeshi Uchida, Dr. Kohji Yamamoto, Dr. Takehiro Ohta, Daichi Okuno and Akira Sato for their variable discussion and friendly encouragement. Acknowledgement is also made to Dr. Takashi Ogura (Tokyo Univ.) and Dr. Takeshi Tomita (Tohoku Univ.) for their helpful suggestion. She is thankful to Mrs. Emiko

Nomura for her office work and heartfelt kindness.

Finally, the author expresses her sincere gratitude to her patents for their supports, generous understanding, and affectionate encouragement.

September, 2001

Okazaki

Nami Haruta

CONTENTS

Chapter 1. General Background for Protein Dynamics and Ultraviolet Resonance Raman Spectroscopy

1-1. Dynamic Landscapes of Protein

Higher order structure of protein is responsible for the most of its properties. Although primary sequence of protein is only the sequence of twenty kinds of amino acids, the complicated and specific conformation of protein is formed according to this sequence (Figure 1-1.1). How does amino acid sequence specify a tertiary structure of protein? This is a folding problem. Tertiary structure of the protein causes long range interactions on the sequence, referred as topology, and yields tight packing with various interactions; hydrophobic, electrostatic and hydrogen bonding interactions. These interactions enable to provide the cooperative folding process and also cooperative conformational change in native structure. In general the native conformation is considered to be a unique structure that can be defined by X-ray crystallographic analysis, and it is distinct from a denatured conformation that is a generic name for huge amount of flexible conformations (Section 3-1). However, another aspect has been provided that the structure of native proteins also have different conformational states that can be described as a dynamic landscape like folding funnel (Section 3-1). The funnel shape implies the distribution and flexibility of protein conformations with different degrees of roughness, width and barrier height and they will change by environment. One of the dynamic features of proteins is seen for ligand binding (Figure 1-1.2) (*1-3*).

Ligand binding is an essential process of biological reactions. Highly complicated structures of proteins are required for their function inevitably through the binding; enzyme and substrate, protein-protein, -DNA, -peptide and -lipid and so on. In many cases, the ligand binding is accompanied with the conformational change. Figure 1-1.2 shows the energy landscape for dynamics in the two different environment.

In environment 1, for example unliganded state, the distribution of conformational substates is large, meaning the large fluctuation of structure, and substate A is dominant. Changing the environment to environment 2, that is ligand bound state, the distribution of conformational states changes to a narrow shape, and the dominant substate is substate B.

The ensemble of native structure has been demonstrated by various experiments of different techniques. X-ray crystallographic structures of homodimeric proteins show often that two subunits cannot be overlapped exactly on each other, although two subunits have same sequence and quite similar structures. For example, CooA, the homodimeric DNA binding protein, indicated that DNA binding domain takes an open form in one subunit and a closed form in the other subunit (Figure 2-1.6) (*4*). It seems that the crystal structure is a snap shot of a motion of protein driving in the fluctuation between two conformational substates. The structure analysis of protein by NMR, for which a sample is in a solution phase, determined many similar structures more than ten (*5*). Recently, nonlinear optical spectroscopy such as hole burning and photon echo in visible or infrared region have been applied to explore the protein dynamics in equilibrium state (*6-8*). The photon echo experiments of myoglobin at a low temperature indicates that the energy barrier of the transition to next conformational substate is between 3 and 10 kJ/mol and each conformational substate has furthermore lower conformational substates (*8*).

The fluctuation of the protein in the narrow balance with small energy difference is essential for the protein to change the conformation easily and to reflect the environment upon functioning. The character of the distribution of the substates may have a possibility to predict the direction of the reactions. Moreover, since the two-

state folding proteins allows to predict the folding rate from the topology of native structure (Section 3-1), the native conformational fluctuation may play a key role to answer the question how random coil searches can find the native topology. However, experimental studies are too small and diverse to achieve the consensus of the protein fluctuation. The experimental studies of protein fluctuation are increasingly important, especially related to the function, and it must ordered to hierarchies with clear definition.

1-2. Ultraviolet Resonance Raman Spectroscopy

Raman spectroscopy is an excellent tool to elucidate the structural features of proteins. Many colored proteins, especially heme proteins, have been analyzed by the Raman spectroscopy with visible lasers (*9*), but it is limited to apply the technique to many of colorless proteins. The recent progress of ultraviolet lasers has enabled us to carry out the Raman excitation in UV region, namely, the Raman excitation at absorption band of protein matrix. Consequently, ultraviolet resonance Raman (UVRR) spectroscopy excited in the region between 260 and 220 nm cause possible to monitor selectively aromatic residues (Figure 1-2.1,2) such as tryptophan, tyrosine and phenylalanine, and that excited below 220 nm can monitor the vibrations of amide bonds of polypeptide backbone.

Normal coordinate analysis of indole ring, that is a side chain of Trp residue, was provided by Takeuchi and Harada (*10, 11*). Figure 1-2.3 shows the 244-nm excited UVRR spectrum of horse apomyoglobin that has two Trp and Tyr residues. The Raman band derived from Trp residue is labeled with W and Tyr is with Y. Harada

and coworkers studied many model compounds of Trp under various solvents effect and pointed out that a selected Raman band of Trp residue serves as a good structural marker (*10*). W3 band around 1560 ~ 1540 cm^{-1} is due to the C_2-C_3 stretching mode and it shows a good correlation with torsion angle C_α-C_β- C_3-C_2 (Figure 1-2.1 and section 2-2). W7 band splits into two bands due to Fermi resonance between the N_1C_8 stretching fundamental (W7) and the combination of out-of-plane bending vibrations. The intensity ratio of the W7 doublet, I_{1358}/I_{1338}, is known to be sensitive to environments of the indole side-chain of Trp residue, being larger for more hydrophobic environment (*12*). W17 band around 885 ~ 870 cm^{-1} is a marker of hydrogen bonding. This band shows downshifts by the formation of the stronger hydrogen bonding.

The normal coordinate analysis of *p*-cresol, the side chain of Tyr residue was also carried out by Takeuchi and Harada (*13*). In UVRR spectrum excited at 244 nm, the ratio of Y7a (~1205 cm^{-1}) /Y9a (~1174 cm^{-1}) also serves as a good marker of hydrophobic and hydrophilic environment (Section 2-2) (*14*). Y8a (1619 cm^{-1}) and Y9a (~1174 cm^{-1}) bands are downshifted to 1600 and 1170 cm^{-1}, respectively, upon the deprotonation of phenoxy group of Tyr (Section 2-3) (*10*).

Among the vibrational modes of main chain amide group, CHR-CO-NH-CHR', amide I, amide II and amide III bands (Figure 1-2.4) are enhanced in the UV region below 220 nm. Asher and coworkers have studied the protein dynamics with UVRR spectroscopy excited at 206.5 nm and established the methodology for calculation the secondary structure contents (*15, 16*). Table 1-2.1 shows the wavenumbers characteristic of each secondary structure.

It is demonstrated in this thesis that UVRR spectroscopy has a strong power for detecting the conformational changes of proteins. The detection of the subtle

conformational change of myoglobin during the ligand binding and selective enhancement of ion-coordinated Tyr are such examples. Furthermore, clear Raman information about the structural change and environmental change of Trp side chain can be used to distinguish between intermediates in the kinetic folding process. A possibility that UVRR spectroscopy can provide information on faster protein dynamics such as initial steps of protein folding and conformational relaxation is expected.

Reference

1. Kumar, S., Ma, B. Y., Tsai, C. J., Sinha, N., and Nussinov, R. (2000) *Protein Science 9*, 10-19.
2. Ma, B. Y., Kumar, S., Tsai, C. J., and Nussinov, R. (1999) *Protein Engineering 12*, 713-720.
3. Tsai, C. J., Ma, B. Y., and Nussinov, R. (1999) *Proceedings of the National Academy of Sciences of the United States of America 96*, 9970-9972.
4. Lanzilotta, W. N., Schuller, D. J., Thorsteinsson, M. V., Kerby, R. L., Roberts, G. P., and Poulos, T. L. (2000) *Nat Struct Biol 7*, 876-80.
5. Osapay, K., Theriault, Y., Wright, P. E., and Case, D. A. (1994) *J Mol Biol 244*, 183-97.
6. Champion, P. M. (1992) *J. Raman Spectrosc. 23*, 557-567.
7. Mukamel, S. (1995) *Principle of nonlinear optical spectroscopy*, Oxford University Press, New York.
8. Leeson, D. T., Wiersma, D. A., Fritsch, K., and Friedrich, J. (1997) *Journal of Physical Chemistry B 101*, 6331-6340.

9. Spiro, T. G. (1988) *Biological Application of Raman Spectroscopy*, Vol. 3, John Willy and Sons, New York.

10. Harada, I., and Takeuchi, H. (1986) *Advances in Spectroscopy: Spectroscopy of biological System*, 113-175.

11. Takeuchi, H., and Harada, I. (1986) *Spectrochimica Acta 42A*, 1069-1078.

12. Harada, I., Miura, T., and Takeuchi, H. (1986) *Spectrochimica Acta 42a*, 307-312.

13. Takeuchi, H., Watanabe, N., and Harada, I. (1988) *Spectrochimica Acta 44A*, 749-761.

14. Haruta, N., Aki, M., Ozaki, S., Watanabe, Y., and Kitagawa, T. (2001) *Biocemistry 6956-6963.*

15. Chi, Z., and Asher, S. A. (1998) *Biochemistry 37*, 2865-72.

16. Chi, Z., Chen, X. G., Holtz, J. S., and Asher, S. A. (1998) *Biochemistry 37*, 2854-64.

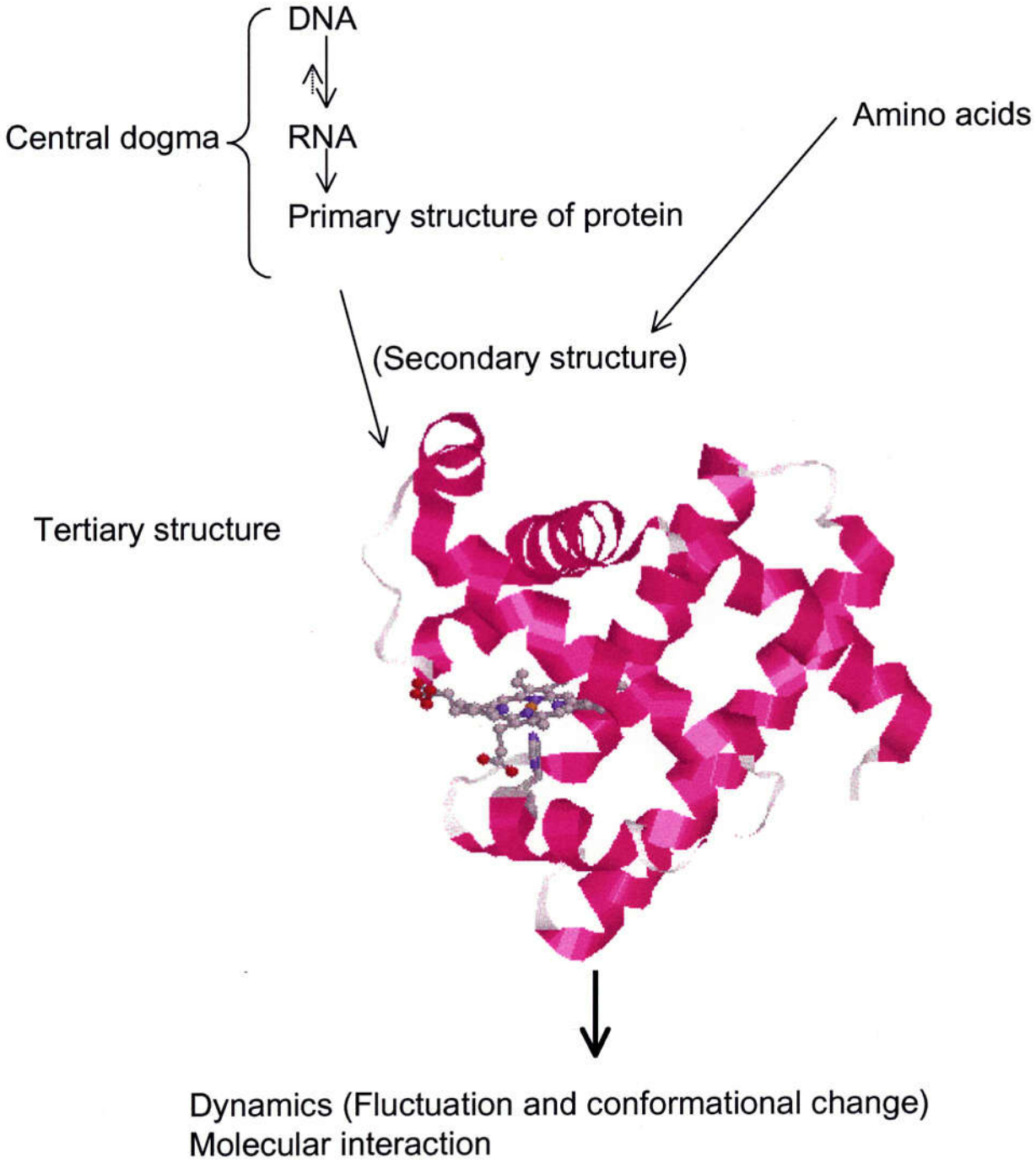


Figure 1-1.1. The flow chart from genomic information to the protein function. Protein structures are constructed by polypeptide chain of amino acids sequence according to DNA sequence. The central structure is drawn by ribbon diagram that traces the protein backbone.

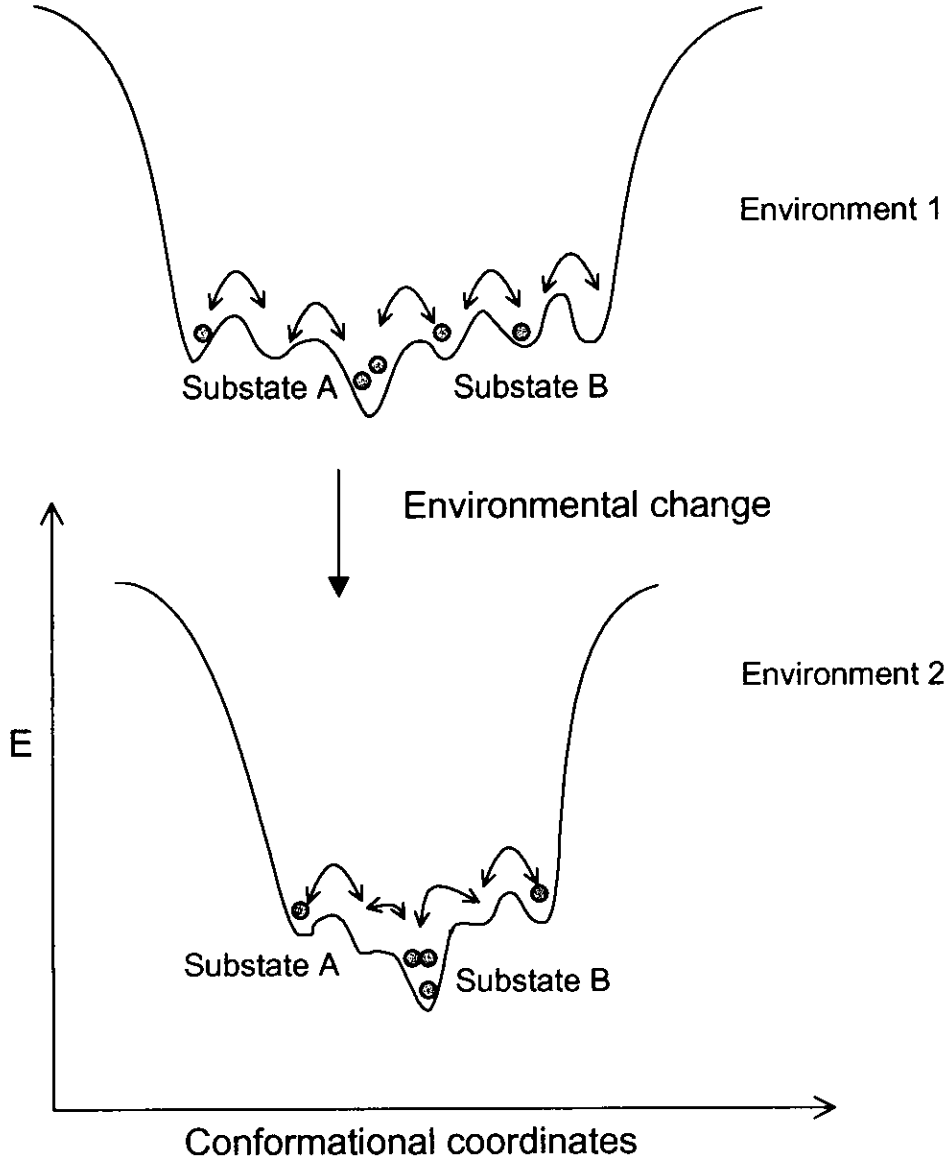


Figure 1-1.2. Scematic depiction of protein dynamic landscape. In the environment 1, a substate A is dominant and in environment 2, which has different distribution and different shape of landscape from 1, substate B is dominant.

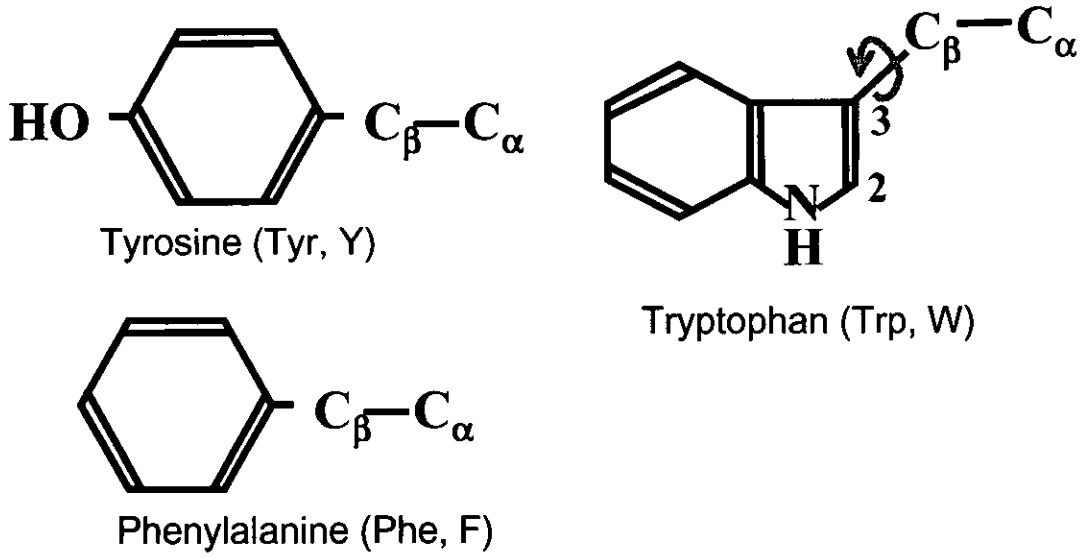


Figure 1-2.1. The structures of aromatic residues.

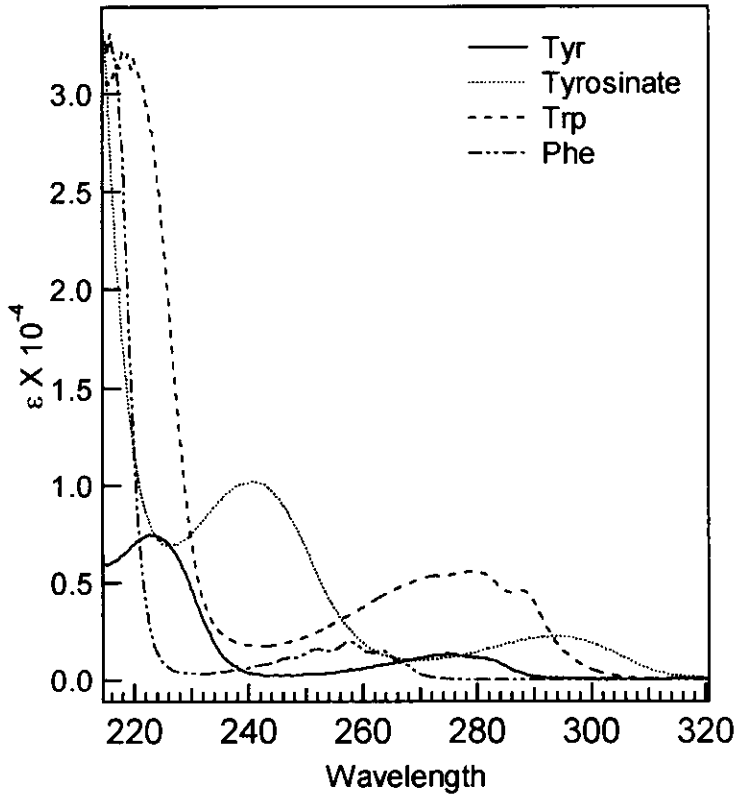


Figure 1-2.2. Ultraviolet absorbance spectra of the amino acids revised the concentration to 1 M in a 1 cm cell at 20 °C.

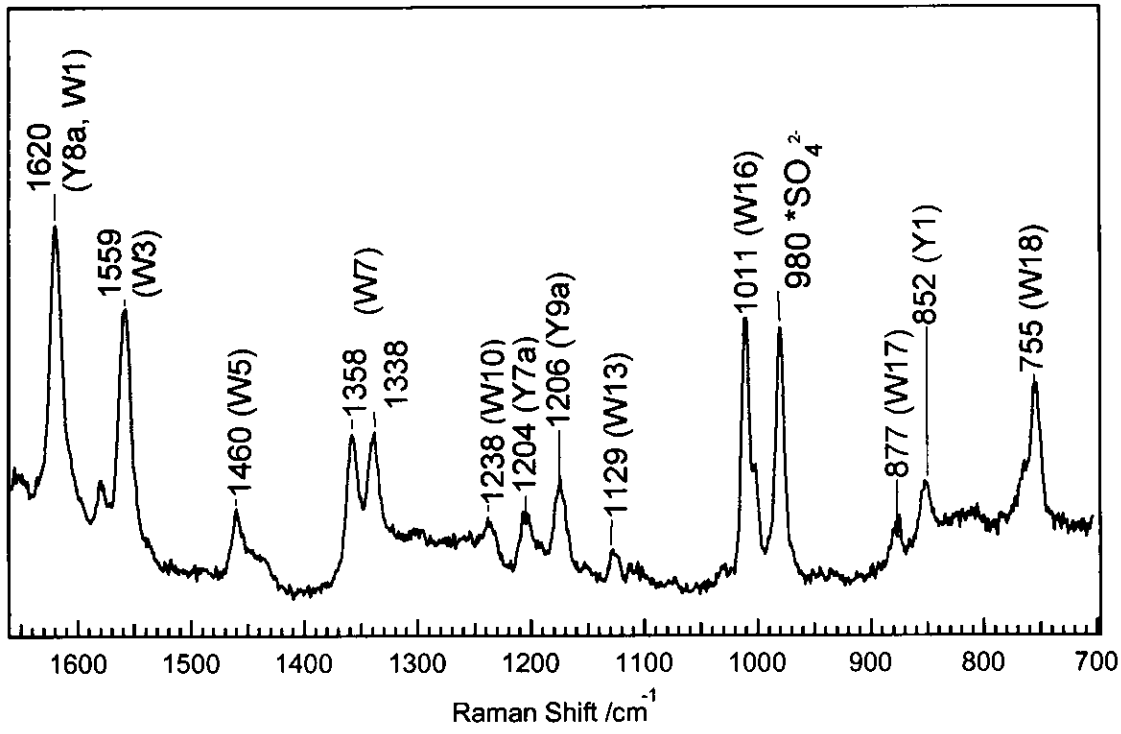


Figure 1-2.3. The UVRR spectrum of horse apomyoglobin excited at 244 nm. UVRR bands of Trp and Tyr are labeled by W and Y, respectively. Horse myoglobin has two Trp and Tyr residues.

Table 1-2.1. *The main structural maker band of main chain in UVRR spectroscopy*

Wavenumber	mode	The correlation of the structure	
Main chain (χ_{ex}= 206.5 nm)			
1685 ~ 1645	amide I (C=O str.)	Downshift by stronger hydrogen bonding	
		unordered	1665 cm^{-1}
		β sheet	1654 cm^{-1}
		α Helix	1647 cm^{-1}
1560 ~ 1540	amide II (C-N str.)	Downshift by stronger hydrogen bonding	
		unordered	1560 cm^{-1}
		β sheet	1551 cm^{-1}
		α Helix	1545 cm^{-1}
1300 ~ 1239	amide III (N-H bend + C-N str.)	Downshift by stronger hydrogen bonding	
		unordered	1267 cm^{-1}
		β sheet	1235 cm^{-1}
		α Helix	1299 cm^{-1}
ca.	amide S (CH bend coupled amide III)		
		unordered	1386 cm^{-1}
		β sheet	1386 cm^{-1}
		α Helix	-

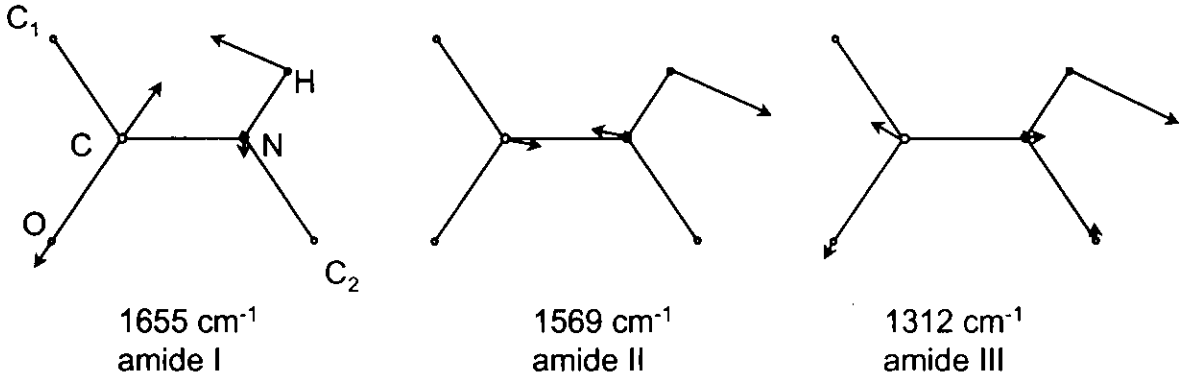


Figure 1-2.4. Atomic displacement vectors of amide modes of N-methylacetamide.

Chapter 2. Conformational Change of Myoglobin upon Ligand Binding Probed by Ultraviolet Resonance Raman Spectroscopy

2-1. Overview: Recognition of a Small Ligand and Heme-Linked Conformational Change of Heme Proteins

2-1.1 Structure and properties of the heme

2-1.2 Structure and property of the heme pocket

a. The role of the residues at the distal side

b. The role of the proximal side

2-1.3 The conformational change of heme proteins

a. Hemoglobin

b. Diatomic heme-based sensor proteins

Ligand binding is the essential reaction of protein functions. Many of molecules, such as diatomic gas molecules, amino acids, small peptides, lipids, nucleotide and protein, could induce a reaction in the cell as a ligand, to put it another way, as a substrate, a signal molecule, and so on. The most of the ligand binding reaction has accompanied with a conformational change of protein structure to control their reaction by itself or some molecules, or for transmission of the signal to another subunit or proteins. The essential factor in the response of the cell is the protein-protein interaction accompanied with the conformational change of the protein structure.

Heme protein is general term of proteins, which contains the heme, the iron complex of porphyrin, as a cofactor. The common function is the catch up a small gas molecule to bind with the heme iron atom, except few proteins such as cytochrome *c*, which functions as an electron transfer. The next step of reaction depends on each protein, which can be divided to three categories, enzyme activity (Catalase, cytochrome P-450 family, cytochrome c oxidase), gas molecule transport and storage (hemoglobin and myoglobin) and sensor for diatomic molecules (soluble guanylate cyclase, FixL, CooA, DOS). These manifold functions of heme proteins are derived from the property of the amino acids residues that surround the heme and construct the heme pocket and that coordinate to the heme iron atom.

2-1.1 Structure and properties of the heme

The structure of the heme, the iron porphyrin complex, is shown in the Figure 2-1.1. The four nitrogen atoms of pyrrole rings coordinate to the iron. In addition, an amino acid residue on the protein matrix and an exogenous ligand bind as a 5th and 6th axial ligands. The porphyrin ring has a highly conjugated π-electron system and has

two absorption bands in the visible and near UV regions, called as Q band and Soret band (B band), respectively. These absorption bands cause the vivid color of porphyrin, red, brown and green that changes according to the oxidation numbers and coordination states (*1*).

The residues that can coordinate to the heme iron are histidine, cysteine, methionine, tyrosine and proline. Proline was found only in CooA recently (*2*). In ordinary circumstances, a heme has either of two oxidation states, the oxidized state (Fe^{3+}, ferric, or met) and the reduced state (Fe^{2+}, ferrous, or deoxy) in stationary conditions. The heme in the oxidized state can bind a water molecule or a number of different anions such as N_3^-, CN^-, NO_2^-, SCN^- and F^-, and in the reduced state can bind O_2, CO, NO, aryl nitroso and alkyl isocyanide compounds (*3*). The spin state of heme iron changes depending on the axial ligand. The size of heme iron in the low spin state is fitted to porphyrin centered core, while the heme iron in the high spin state is larger than the core size of porphyrin ring. Therefore, iron atom is displaced from the heme plane and heme ring changes to a distorted shape like a dome. The displacement of the iron atom affects to the protein matrix through the axial ligated residue as described below. Resonance Raman spectroscopy excited at Soret band is very sensitive to these changes of shape and size of porphyrin ring, since the vibrational state and electronic state of porphyrin ring are dramatically changed by the spin state of iron (*4*).

2-1. 2 Structure and property of the heme pocket

Heme proteins have a heme pocket that contacts the iron porphyrin. The iron atom is connected to a specific residue by a covalent bond. In some proteins such as

cytochrome *c*, coordinate the heme iron with two residues from the opposite directions. However, most proteins coordinate to the heme iron with only one axial ligand and is this axial residue called as a proximal residue. The opposite side of the proximal site can be coordinated by an exogenous ligand as a sixth ligand and the surrounding of this side is called as a distal side. Since the sort of proximal residue and property of the distal side of the heme pocket control the function of each protein, I will discuss the roles of amino acid residues in the distal and proximal side of myoglobin (Mb), as a representative model of heme protein. Mb is a heme protein that transports and stores the oxygen molecule in aerobic muscle tissue of vertebrates (Figure 2-1.2).

a. The role of the residues at the distal side

The structure and electrostatic property of the distal side is significant to recognize an exogenous ligand molecule. Although the chemical and structural roles that define this interaction are unique to Mb, the underlying mechanisms have general applicability to all protein-ligand interactions.

An unhindered pentacoordinate model heme in organic solvents binds O_2 and CO with a ratio that prefer CO to O_2 by as much as 30000- to 100000- folds (*5*). When heme is embedded in the protein matrix of Mb, that ratio is changed to approximately 30. Therefore, Mb recognized these chemically and structurally similar molecules at the atomic level and selectively discriminating against the binding of CO in favor of O_2. Many studies of site directed mutagenesis targeted to the residues of distal side have been succeeded to reveal the electrostatic and steric role of each residue and the mechanism of the preference of O_2 over CO. Figure 2-1.3 shows the X ray crystal structure focused on the heme pocket of sperm whale myoglobin (*6*). The selective

target residues of mutagenesis, including His64 (E7), Val68 (E11), Phe43 (CD1), Phe46 (CD4) and Leu29 (B10) are labeled. Since His64 is the closest residue to the iron coordinated ligand, called as distal His, its substitution result in the significant influence on the affinity of CO and O_2. In many detailed studies of X-ray crystallography, there is one water molecule stored in the heme pocket of deoxy unliganded state by the hydrogen bond with distal His. Since the initial step of CO and O_2 binding may be the displacement of this water, the substitution of distal His resulted in the increase of the overall association rate constant of CO to the mutants without the water molecule (*7, 8*). However, in O_2 binding, the distal His substitution significantly affected the dissociation rate constant. The site directed mutagenesis and high resolution crystallography indicated that coordinated O_2 to iron is stabilized by the hydrogen bonds with distal His and the mutation effect came from the disruption of this hydrogen bond rather than other factors such as steric hindrance or electrostatic fields. Therefore, the displacement of water must also occur in the case of O_2 binding, but is compensated by an even stronger hydrogen-bonding interaction to bound O_2.

The substitutions of other residues of the distal side are less effective on the affinity of CO and O_2 than that of distal His. Results for the series mutation of Val68, Phe43 and Leu29 seemed that the steric hindrances of these residues are small. In the kinetic studies combined with various mutants, it is indicated that Val68 and L29 are located along the pathway of the exogenous ligand and these mutants change the ratio of the geminate species after photodissociation (*9*).

b. The role of the proximal side

The selection of the proximal side residue such as His, Tyr and Cys is absolutely

based on the function for influencing the binding and activity of trans ligand. If the axial ligands are the same residue between two kinds of heme proteins, the property of the heme is influenced by the protein matrix through proximal residues, as in Hb and soluble guanylate cyclase (sGC), and it will be mentioned in the next section. However, since the large structure change is affected to the heme iron through the proximal residue, the small local mutation around the proximal residue, except just the proximal residue, is smaller influence on the function such as specificity of a ligand than the substitution of the distal side (*10*). In other hands, the substitution of the residue on the protein surface to cut off the subunits-subunits interaction is undoubtedly effective to the cooperative function of oligomeric protein(*11*).

2-1.3 The conformational change of heme proteins

Some proteins show the clear correlation between ligand binding on the heme and conformational change of the protein matrix for regulation of its function. Here, I will mention about two categories of proteins, hemoglobin and heme-based sensor proteins. Hemoglobin has a long history of the studies and it is one of the first protein whose X ray crystal structures were reported as same era as myoglobin (*12, 13*). In contrast, heme-based sensor proteins were found in the relatively recent studies, and now many kinds of new proteins have been found continually in this category. Thus, a few proteins that have been advanced in the studies are given as a good example.

a. Hemoglobin

Hemoglobin (Hb) is found to be packed at high concentrations (~20mM) in red

blood cells. Hb transports the oxygen from a lung to tissues of vertebrates. Hb is an $\alpha_2\beta_2$ heterotetramer and each subunit, α and β, has a folding structure similar to monomer of Mb. This higher order structure of heterotetramer, quaternary structure, is specified and is the function of Hb advanced more than that of Mb, regarding cooperativity of O_2 binding.

Figure 2-1.4 shows the oxygen binding curve of Hb and Mb. The Mb curve is a simple rectangular hyperbola as expected from the mass action law, while the Hb curve is sigmoid, called "cooperative binding". The oxygen affinity increases with increasing oxygen saturation and this phenomenon is important for the oxygen transport function that Hb binds oxygen in the lung and releases in the tissue. Cooperative binding is often expressed in terms of the empirical Hill equation,

$$Y / (1\text{-}Y) = K (P_{O2})^{n}$$

Where Y is the fractional saturation of binding sites, K is an apparent binding constant, P_{O2} is the partial pressure of oxygen, and n is Hill coefficient. The value of n is proportional to the cooperativity. If the value of n is near 1, it indicates the no cooperative binding and the binding curve shows as the curve of Mb in Figure 2-1.4. The value of $n > 1$ means the positive cooperativity that the ligand binding increases the binding affinity, and the max value of n is the number of binding sites. Typical value of the Hill coefficient for normal human Hb is 2.8. This Hill equation can be applied to other oligomeric proteins and the proteins that have n value larger than unit are referred as allosteric enzymes.

In 1960's, Perutz succeeded the X ray crystal analyses of Hb and revealed the difference of the structures between the deoxy state and CO bound state (*13*). On the base of this information, he proposed a mechanism to explain the cooperative oxygen

binding (*14*). Hb has two different structural states of tetrameric Hb, which differ not in crystal structure but in the solubility, ligand affinity, and other properties. One state is R (relax) state, in which the affinity is high and the structure is characteristic of all liganded form of normal Hb (Figure 2-1.5). Another is T (tense) state, in which affinity is low and the structure is characteristic of unliganded form of normal Hb. All heme irons are in the five coordinated state. Perutz suggested that the tetrameric structure of T state has strong interaction between the subunits by the electrostatic interactions and that the iron atom of heme is strained by the protein matrix through the proximal His. The ligand binding to the trans site of proximal His cause the movement of the iron atom to the heme plane and induces the conformational change of Hb from T state to R state (*14*). It is difficult to demonstrate the change of the bonding strength between the iron and nitrogen due to the proximal change by the X ray crystallographic studies. The bond strength may control the affinity of oxygen in Perutz mechanism.

However, the resonance Raman spectroscopic studies excited at 441.6 nm can detect the iron-histidine (Fe-His) stretching mode and gave strong support for the Perutz mechanism. The Fe-His stretching mode in the R state of Hb and deoxy Mb appeared at 221 and 220 cm^{-1}, respectively, but in the T state of Hb at 215 cm^{-1}. The low frequency in the T state indicates that the bonding of Fe-His is weaker in the T than R state and the bond length is longer by the strain of protein matrix (*11, 15*).

b. Diatomic heme-based sensor proteins

Recently, many heme proteins have been found to function as a sensor protein for diatomic molecules including CooA (transcriptional factor) for CO (*16*), FixL (histidine kinase)(*17*), *hem*AT (*18*) and Dos protein (*19*) for O_2 and soluble guanylate cyclase

(sGC, GTP → cGMP) for NO (*20*). These heam-based sensor proteins play an important role in biological signaling, and their structures consist of the heme domain (or subunit) as a sensor domain and functional domain such as catalytic site for other substrate or DNA binding. The heme-based sensor proteins recognize a specific diatomic molecule by its binding to the heme iron, and subsequent conformational changes of the protein result in the control of the functional reaction at the catalytic site.

CooA is a CO sensing transcriptional factor from *Rhodospririllum rubrum* that grows using the energy evolved from oxidation of CO to CO_2 as its sole energy source. When CO binds to the heme, DNA binding domain of CooA can bind to the DNA and induces the multicomponent CO oxidation systems (*21*). The heme ligation states of CooA are changed dramatically, and recent works of mutagenesis (*22-24*) and X ray crystallographic analysis (*2*) explain each ligation states (Figure 2-1.6). The astonishing result of X ray crystal structure of 6-coordinated ferrous states is that the 6 th axial ligand is proline residue, that is the first example of Pro as a ligand of metalloproteins, and moreover that this Pro comes from the other subunit. We can expect that the Pro residue as the sixth ligand will be kicked out when CO coordinates to heme iron (*25*) and it induced the global conformational change through subunit-subunit interaction to increase the affinity to DNA.

NO is a significant molecule as a signal of physiological important reactions and sGC is an NO receptor protein in a series of signal transduction. The activity of sGC, the conversion from GTP to cGMP, is increasing 150-fold by the NO binding (*26*). The Fe-His bond in ferrous sGC is weakest of all known His bound heme proteins, as shown by Fe-His stretching mode at 204 cm^{-1}. The NO binding at the trans site induced the disruption of the bond between iron and His to yield a five coordinated low spin ferrous

nitrosyl heme. It is easy to image that this cleavage of Fe-His linkage induces the conformational change of GTP binding site. Furthermore, since the presence of the substrate influenced on the Raman band of Fe-NO and FeN-O streching vibrations, the heme domain must have strong interaction with catalytic domain in a highly cooperative way (*27*).

The studies of heme based proteins, especially about the structure and dynamics, are insufficient now and it is anticipated that the increase of the conformational information will explain the specific interaction for regulation of their functions.

References

1. Adar, F. (1978) in *The Porphyrins* (Dolphin, D., Ed.) pp 167-209, Academic Press, New York.
2. Lanzilotta, W. N., Schuller, D. J., Thorsteinsson, M. V., Kerby, R. L., Roberts, G. P., and Poulos, T. L. (2000) *Nat Struct Biol 7*, 876-80.
3. Yu, N., and Kerr, E. A. (1988) in *Biological Application of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 39-95, John Willy and Sons, New York.
4. Spiro, T. G., and Li, X. (1988) in *Biological Application of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 1-37, John Willy and Sons, New York.
5. Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., and Olson, J. S. (1989) *J Biol Chem 264*, 3057-60.
6. Kachalova, G. S., Popov, A. N., and Bartunik, H. D. (1999) *Science 284*, 473-6.
7. Quillin, M. L., Arduini, R. M., Olson, J. S., and Phillips, G. N., Jr. (1993) *J Mol Biol 234*, 140-55.

8. Springer, B. A., Sligar, S. G., Olson, J. S., and Phillips, G. N., Jr. (1994) *Chem. Rev. 94*, 699-714.

9. Olson, J. S., and Phillips, G. N., Jr. (1996) *J Biol Chem 271*, 17593-6.

10. Shiro, Y., Iizuka, T., Marubayashi, K., Ogura, T., Kitagawa, T., Balasubramanian, S., and Boxer, S. G. (1994) *Biochemistry 33*, 14986-92.

11. Matsukawa, S., Mawatari, K., Yoneyama, Y., and Kitagawa, T. (1985) *J. Am. Chem. Soc. 107*, 1108-1113.

12. Kendrew, J. C. (1960) *Nature 185*, 422-427.

13. Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. (1968) *Nature 219*, 131-9.

14. Perutz, M. F. (1970) *Nature 228*, 726-39.

15. Kitagawa, T. (1988) in *Biological Application of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 97-131, John Willy and Sons, New York.

16. He, Y., Gaal, T., Karls, R., Donohue, T. J., Gourse, R. L., and Roberts, G. P. (1999) *J Biol Chem 274*, 10840-5.

17. Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. (1991) *Nature 350*, 170-2.

18. Hou, S., Larsen, R. W., Boudko, D., Riley, C. W., Karatan, E., Zimmer, M., Ordal, G. W., and Alam, M. (2000) *Nature 403*, 540-4.

19. Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M. A. (2000) *Biochemistry 39*, 2685-91.

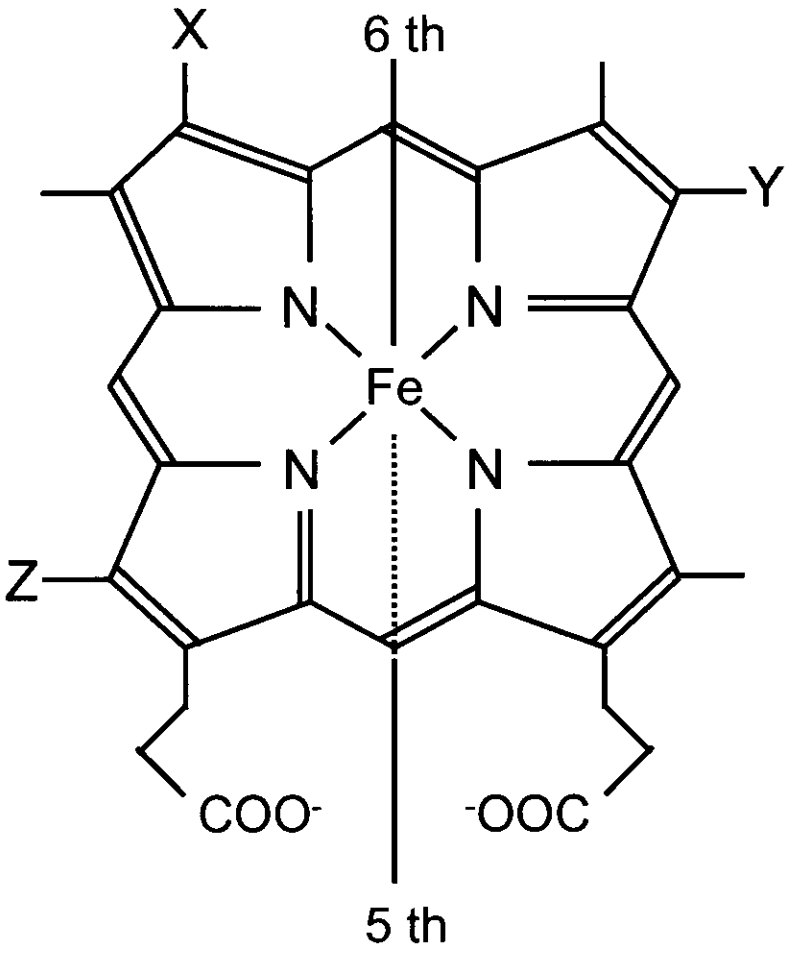
20. Stone, J. R., and Marletta, M. A. (1996) *Biochemistry 35*, 1093-9.

21. Shelver, D., Kerby, R. L., He, Y., and Roberts, G. P. (1997) *Proc Natl Acad Sci U S A 94*, 11216-20.

22. Aono, S., Ohkubo, K., Matsuo, T., and Nakajima, H. (1998) *J Biol Chem 273*,

25757-64.

23. Thorsteinsson, M. V., Kerby, R. L., Conrad, M., Youn, H., Staples, C. R., Lanzilotta, W. N., Poulos, T. J., Serate, J., and Roberts, G. P. (2000) *J Biol Chem 275*, 39332-8.
24. Nakajima, H., Honma, Y., Tawara, T., Kato, T., Park, S. Y., Miyatake, H., Shiro, Y., and Aono, S. (2001) *J Biol Chem 276*, 7055-61.
25. Uchida, T., Ishikawa, H., Ishimori, K., Morishima, I., Nakajima, H., Aono, S., Mizutani, Y., and Kitagawa, T. (2000) *Biochemistry 39*, 12747-52.
26. Tomita, T., Tsuyama, S., Imai, Y., and Kitagawa, T. (1997) *J Biochem (Tokyo) 122*, 531-6.
27. Tomita, T., Ogura, T., Tsuyama, S., Imai, Y., and Kitagawa, T. (1997) *Biochemistry 36*, 10155-60.



heme *a* (cyt *c* oxidase)
X= $-C_{17}H_{27}O$
Y= $-CH=CH_2$
Z= -COH

heme *b* (Protoporphyrin XI)
(Mb, Hb)
X, Y= $-CH=CH_2$
Z= $-CH_3$

heme *c* (cyt *c*)
X, Y= $-CH(CH_3)$-(S-Cys)
Z= $-CH_3$

Figure 2-1.1. The schematic diagram of hem. Myoglobin has one protoporphyrin IX in one molecule and linkaged to 5th coordinate site of the iron atom with proximal His (His93).

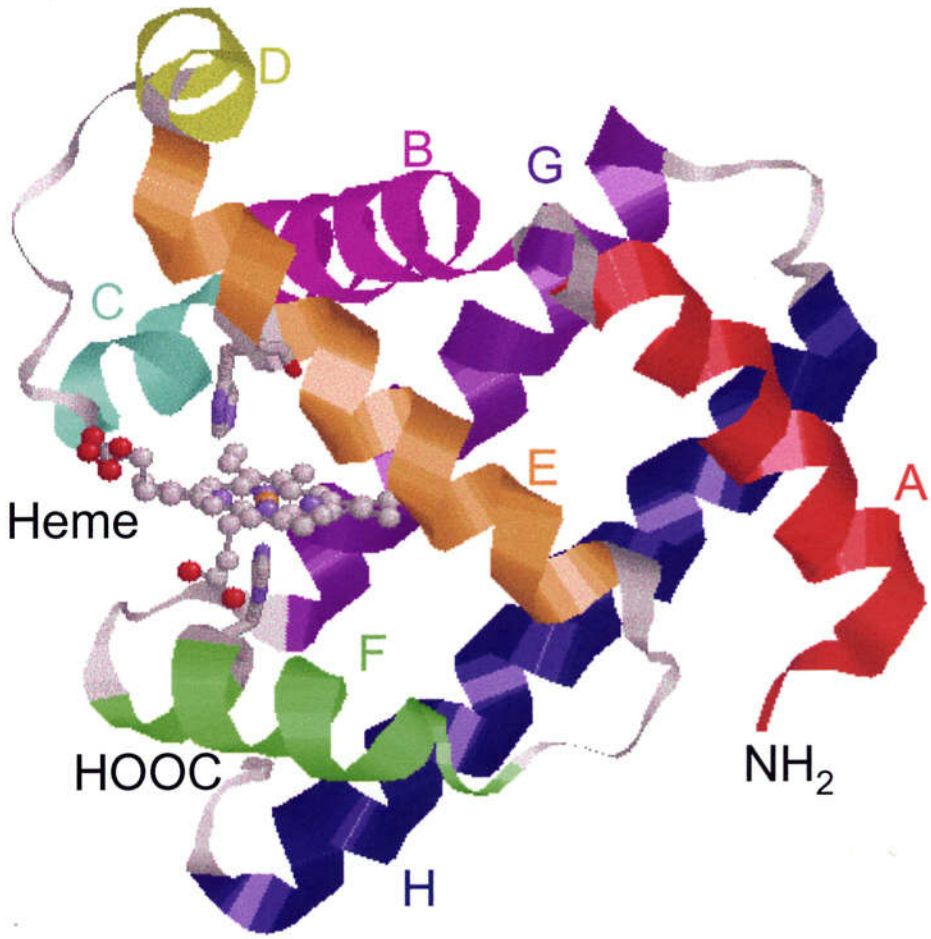


Figure 2-1.2. The ribbon diagram of the sperm whale Mb (1bzr). The 8 helices are labeled from A to H and heme embedded between the E (orange) and F (green) helix. His 93 on the F helix is coordinated to the heme iron (proximal His).

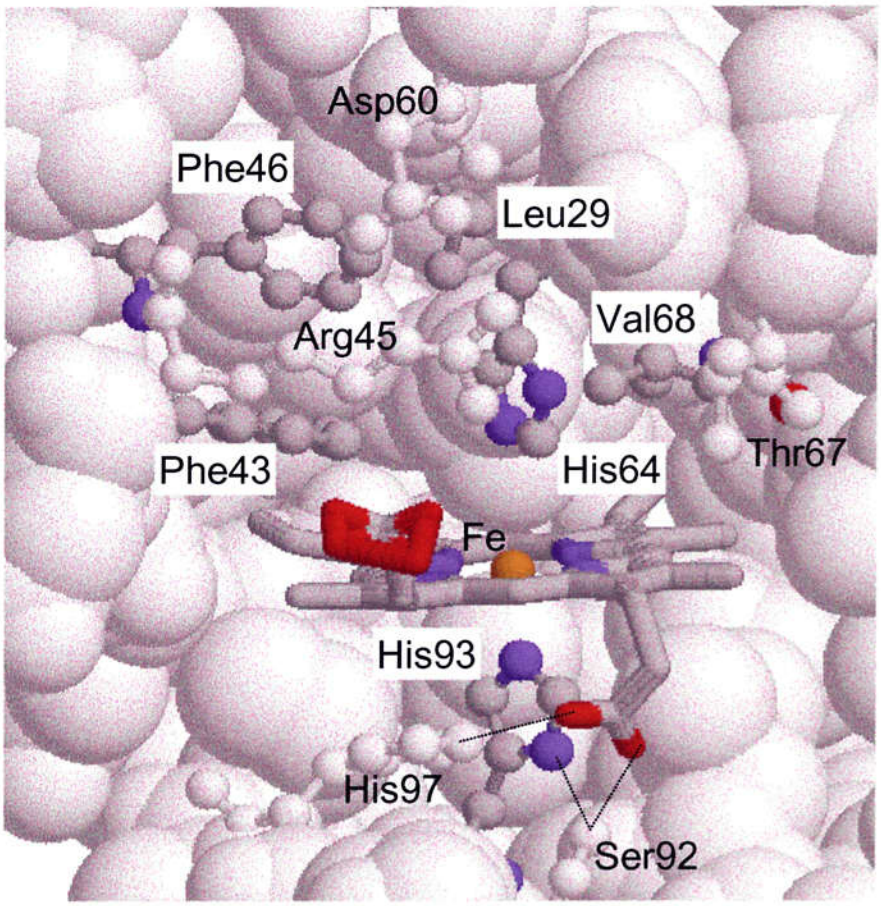


Figure 2-1.3. The X-ray crystal structure around the heme pocket of sperm whale MbCO (1bzr). The main residues are drawn by the ball and sticks with labeling the name, and others are drawn by spacefill style.

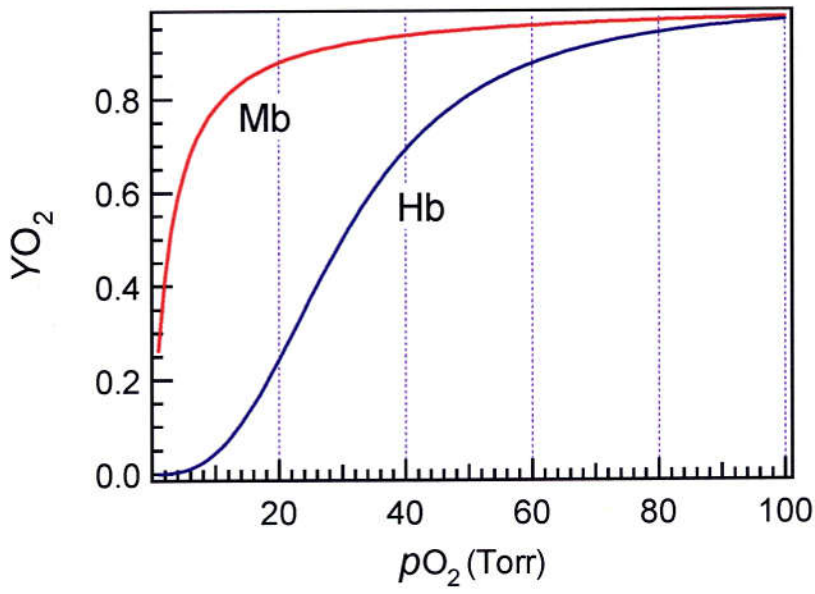


Figure 2-1.4. Oxygen binding curve of Mb and Hb fitted with Hill equation, $YO_2 = [pO_2]^n / ([pO_2]^n + P_{50}{}^n)$; Mb, n=1 and P_{50}=2.8 ; Hb, n=2.8 and P_{50}=30.

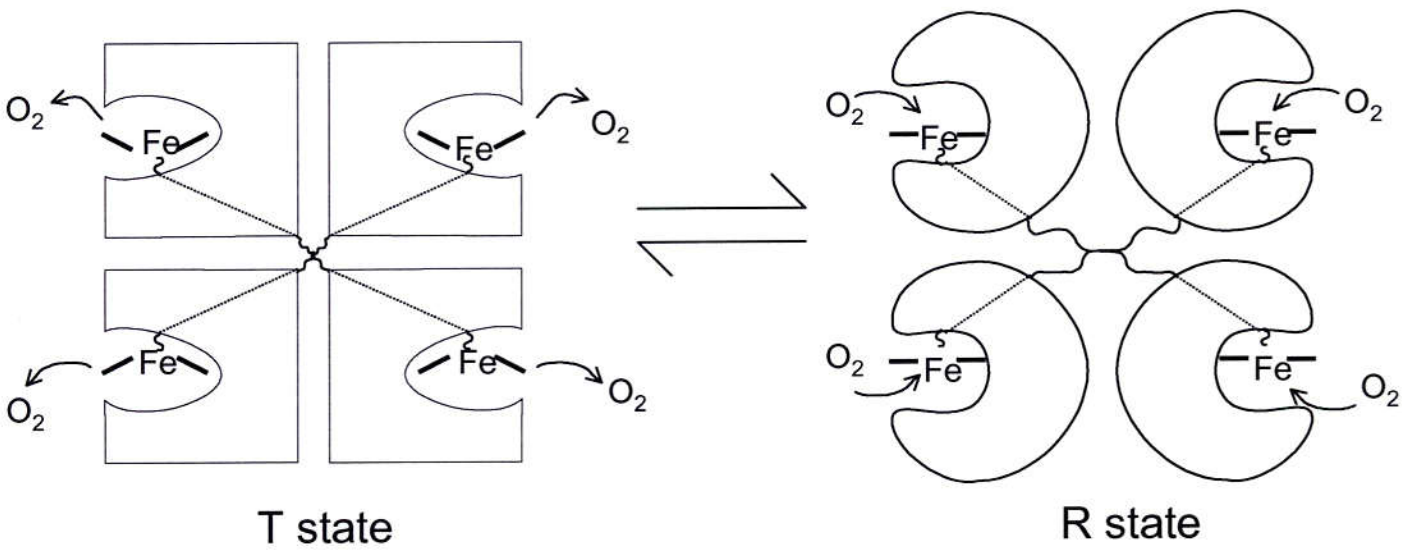


Figure 2-1.5. The schematic diagram of the two states of Hb according to the Perutz mechanism.

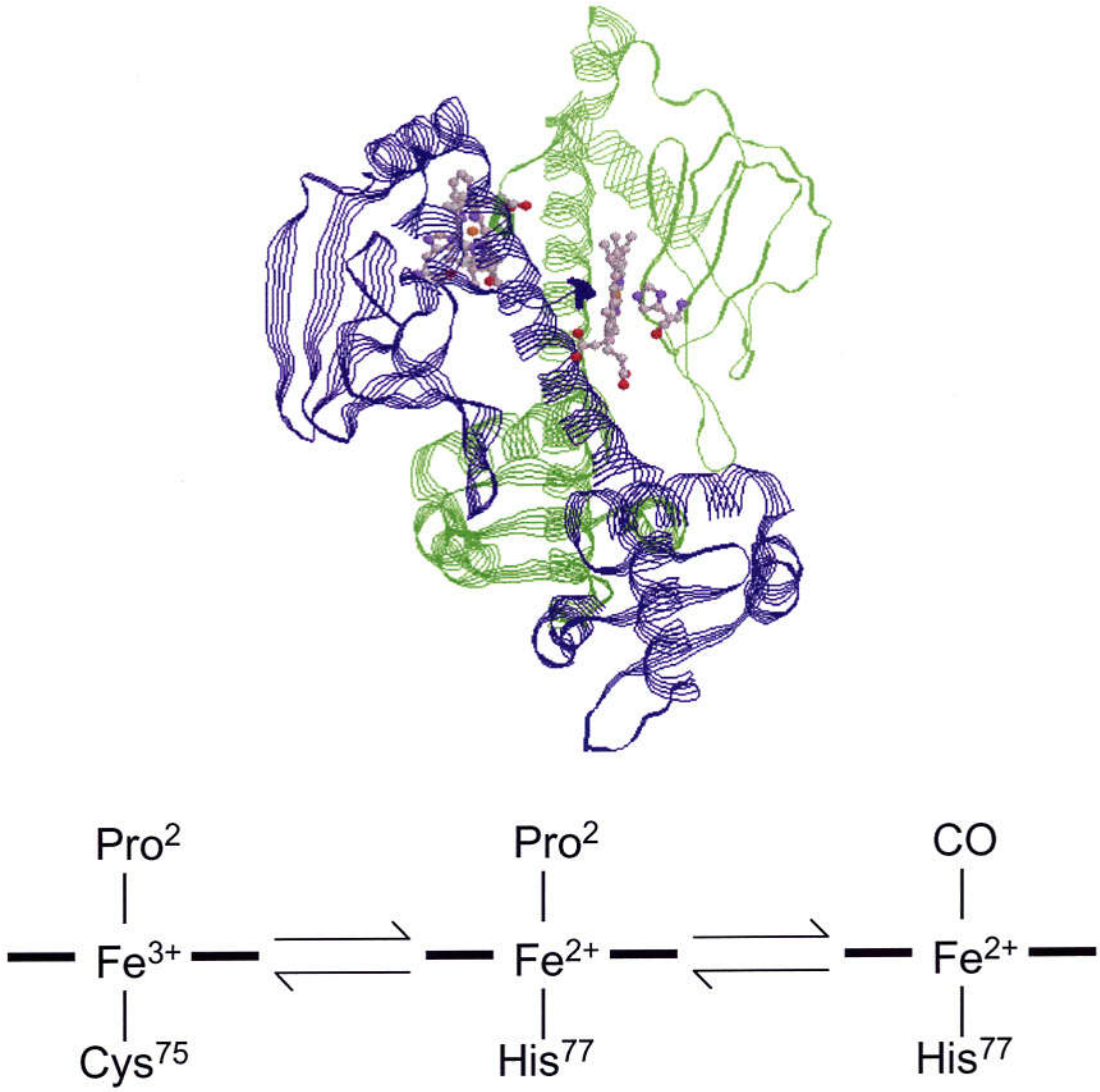


Figure 2-1.6. The Ribbon diagram of CooA (*2*) and its ligation states. CooA is homodimeric protein (green chain and blue chain) and N-terminal Pro2 from opposite subunit provides the 6th ligand.

2-2. Conformational Changes upon Ligand Binding of Sperm Whale Myoglobin and Its Mutants

(Contents of this section has been published in Biochemistry 40, 6956-6963)

Abstract

Conformational change of myoglobin (Mb) accompanied by binding of a ligand was investigated with the 244 nm-excited ultraviolet resonance Raman Spectroscopy (UVRR). The UVRR spectra of native sperm whale (sw) and horse (h) Mbs and W7F and W14F swMb mutants for the deoxy and CO-bound states enabled us to reveal the UVRR spectra of Trp7, Trp14, and Tyr151 residues, separately. The difference spectra between the deoxy and CO-bound states reflected the environmental or structural changes of Trp and Tyr residues upon CO binding. W3 band of Trp7 near the N-terminus exhibited a change upon CO binding, while Trp14 did not. Tyr151 in the C-terminus also exhibited a definite change upon CO binding, but Tyr103 and Tyr146 did not. The spectral change of Tyr residues was characterized through solvent effects of a model compound. The corresponding spectral differences between CO- and *n*-butylisocyanide-bound forms were much smaller than those between the deoxy and CO-bound forms, suggesting that the conformation change in the C- and N- terminal regions is induced by the proximal side of heme through the movement of iron. Although the swinging up of His64 upon binding of a bulky ligand is noted by X-ray crystallographic analysis, UVRR spectra of His for *n*-butylisocyanide bound form did not detect the exposure of His64 to solvent.

Abbreviations: Mb: myoglobin, Hb: hemoglobin, RR: resonance Raman, UVRR: ultraviolet resonance Raman, Fe(PP): iron-protoporphyrin-IX, CTAB: cetyl-

trimethylammonium bromide, 2-MeIm: 2-methylimidazole、HMPA: hexamethylphosphoric triamide. MES: 2-(N-Morpholino)ethanesulfonic Acid, Fe(II)(PP)(2-MeIm): 2-MeIm adduct of Fe(II)(PP), COFe(II)(PP)(2-MeIm): CO adduct of Fe(II)(PP)(2-MeIm).

Introduction

Signal transduction in a biological system becomes increasingly important and elucidation of its structural mechanism is a current topic in basic biochemistry (*1*). Many sensor proteins for diatomic molecules including CooA (transcriptional factor) for CO (*2*), FixL (histidine kinase) for O_2 (*3*), soluble guanylate cyclase (GTP $\rightarrow$ cGMP) for NO (*4*) and Dos protein for O_2 (*5*) are heme proteins which recognize a specific diatomic molecule by its binding to the heme iron, and subsequent conformational changes of a protein result in a functional reaction at the catalytic site. Therefore, thorough elucidation of a structural change of protein upon ligand binding to heme is essential to interpret mechanisms of these proteins. We have selected myoglobin (Mb)[1] as a model molecule for exploring this problem, because Mb binds O_2, CO, and NO, and there are many physicochemical studies along this line including the applications of X-ray crystallographic analysis (*6*), and of NMR (*7*), IR (*8*), and resonance Raman (*9*) techniques to various mutant proteins as well as native one.

Kinetic studies and X-ray crystallographic analysis combined with mutagenesis pointed out that the distal side residues including H64, V68, L29, and F43 control the affinity of exogenous diatomic ligand (*10, 11*). General concept for recognition of an exogenous molecule in the ligand binding have been worked out (*12*), but there are only a few studies about overall conformation changes induced by ligand binding (*13-15*). A pathway of an exogenous ligand from solvent to the buried binding site is formed when distal His is swung up outward. This structure is referred as the open conformation which can be stationary seen upon protonation of distal His (*16, 17*), binding of a bulky ligand (*18-20*), and mutation of F46 (*21*) and transiently upon photodissociation of CO (*22*).

Mb is a typical α-helical protein (~80%) with eight helices labeled by A through H. A single heme is embedded between E and F helices and connected directly to the protein through a covalent bond between His93 (proximal His) and the heme iron. The heme iron

moves out of the heme plane towards His93 by 0.3 Å (*6*) in the deoxy state, while it stays in the porphyrin plane in the ligand-bound form. This movement of the iron atom is considered to serve as a trigger of the T-R transition of hemoglobin (Hb), because it accompanies movement of proximal His (F8) and accordingly induces the movement of F helix containing F8 His. Then, A helix, which is hydrogen bonded with E helix and contains the N-terminus, is dislocated. (*23*). Conversely, the strain of protein produced by the intersubunit interactions is reflected by the strength of Fe-N (F8) bond, which has good correlation with the O_2 affinity (*24*). Consequently, the Fe-His stretching frequency can serve as a marker of quaternary structure of Hb (*25*).

The picosecond IR study by Causprove and Dyer (*13*) demonstrated that a change of protein backbone structure probed by amide I almost is completed in 50 ps after photolysis of CO. The geminate recombination rate of CO to the photodissociated unrelaxed species is distinctly faster than the rate to the equilibrium deoxyMb (*14*), indicating that the small conformational change is functionally significant. Although a fluorescence decay time of Trp7 is reported to be different between deoxyMb and COMb (*15*), little is known about their structural differences.

Resonance Raman (RR) spectroscopy is a powerful tool to monitor a structure of chromophore molecule in the protein (*25*). Particularly UV resonance Raman (UVRR) spectra are useful to monitor protein structural changes (*26*). Upon excitation around 220-250 nm, side chain vibrations of tyrosine and tryptophan residues are selectively intensity enhanced, while amide modes reflecting the secondary structure are dominant upon excitation around 200 nm (*27*). UVRR bands of Trp and Tyr residues reflect not only their hydrophobic/hydrophilic surroundings but also hydrogen bonding (*28*). Some modes of protonated histidine can also be enhanced by 244-nm excitation only for D_2O solutions (*29, 30*). Accordingly, in this study, we applied UVRR spectroscopy to Mb and its mutants to

explore structural change upon ligand binding through spectral changes of Tyr and Trp bands and tried to detect the open form with the signal of protonated His64 in D_2O solution.

Materials and Methods

Procedures for Protein Sample Preparation Native sperm whale Mb (swMb, Biozyme) and horse Mb (hMb, Sigma) were dissolved in 25 mM potassium phosphate (K-P) buffer, pH 6.0, containing potassium ferricyanide. After gel filtration through Sephadex 25G (Amersham Pharmacia Biotech), the Mb solution was loaded on a CM52 column (Whatman), which had been equilibrated with 25 mM phosphate buffer, pH 6.0, and eluted by a linear gradient with 50 mM K-P buffer, pH 9.0. The buffer was changed finally to 50 mM K-P buffer, pH 7.0. DeoxyMb and COMb were prepared by adding sodium dithionite (final concentration, 2 mM) after replacement of inside air of the cell with N_2 and CO, respectively. To obtain the *n*-butyl isocyanide (n-BuIC) adduct of Mb (n-BuICMb), n-BuIC (Aldrich Chem. Co.) was added to the 370 μM Mb solution in the 50 mM K-P buffer (pH 7.0) to make the final concentration of n-BuIC 2.4 mM. To get a D_2O solution of Mb, lyophilized hMb (Sigma) was dissolved in the 50 mM D_2O solution of MES (Dojin Chemicals).

The expression and purification of W7F and W14F mutants were performed according to the method described by Springer (*31*), with some modifications. The mutation of a sequence coding Mb between Kpn I and Pst I restriction site of pUC19 vector was introduced by the PCR-based technique. The cell pellet incubated over night at 37 °C was disrupted by sonication, and a red crude extract was applied to ammonium sulfate fractionation, anion exchange chromatography with DE52 (Whatman) and finally cation exchange chromatography (CM52) as described above.

Preparation of Iron-porphyrin Complexes and Others The ferrous five- and six-coordinate iron-protoporphyrin-IX [Fe(II)(PP)] were prepared by dissolving hemin (Sigma) in 0.1 M NaOH solution followed by dilution to the final porphyrin concentration of 0.5 mM with 50 mM K-P buffer, pH 7.0, containing 1 % cetyltrimethylammonium bromide (CTAB) and 20 mM 2-methylimidazole (2-MeIm). Reduction of this solution with sodium dithionite under N_2 and CO atmosphere yielded the five-coordinated high-spin Fe(II)(PP)(2-MeIm) and six-coordinated low-spin COFe(II)(PP)(2-MeIm), respectively. The formations of these complexes were confirmed by Q band of their absorption spectra.

N-acetyl-tyrosine ethyl ester (Aldrich) and hexamethylphosphoric triamide (HMPA, Aldrich) were used as purchased. N-acetyl-tyrosine ethyl ester was solved into propanol to yield 100 mM solution which was used as a stock solution. For every use the stock solution was diluted with cyclohexane or distilled water at various ratios. Raman bands of propanol and cyclohexane were used to normalize the spectral intensity. HMPA was added to yield 2 % (v/v) solution.

UVRR Measurements The UVRR excitation source at 244 nm was generated by frequency-doubling of the 488-nm fundamental line of an Ar^+ ion laser, which was mode-locked at 80 MHz. About 150 μl aliquot of sample solution was put into a spinning cell with a stirring function (32). Raman scattered light at right angle was collected with a UV microscope objective lens, dispersed with a 126 cm single monochromator (Spex1269) equipped with a 3600 grooves/mm holographic grating, and detected by an intensified charge coupled detector (ICCD, Princeton Instruments, model ICCD-1024MG-E/1). Details of the measurement system are explained elsewhere (32). The spectral resolution was 6.9 cm^{-1} for 244 nm excitation. The laser power at the sample point was 0.12-0.2 mW and the total exposure time to get one spectrum was about one hour or slightly longer.

The integrity of sample after exposure to UV laser light was confirmed by comparing the visible absorption spectra measured before and after the UVRR measurements. To standardize the Raman intensity, the Mb solution containing 100 mM Na_2SO_4 was measured repeatedly. It was found through such measurements that the intensity of the sulfate band at 982 cm^{-1} always reproduced its peak height among the samples from the same solution when the curve slopes at both sides of the spectral region (back ground) were adjusted. To avoid possible salt effects on Mb structures, the intensity standardization for various Mb solutions were carried out using the curve slope adjustment at both sides for the spectra obtained under a certain fixed instrumental condition.

Results

Figure 2-2.1 shows the 244-nm excited UVRR spectra of sperm whale Mb in the deoxy- (A), CO-bound (B), and n-BuIC-bound forms (C). The spectra (D) and (E) delineate the difference spectra, that is, deoxyMb – COMb and deoxyMb – n-BuICMb. Since n-BuIC has a bulky alkyl chain and X-ray crystallographic results indicated the swung up geometry of His64 for it (*20*), some differences are expected if steric interactions at the distal pocket is the origin of structural changes. The 244-nm excited Raman spectra are dominated by contributions from Trp and Tyr residues. The Raman bands arising from Trp and Tyr residues are assigned according to Harada and co-workers (*28*) and labeled by W and Y, respectively. Trp bands are observed at 1619 (W1 overlapped with Y8a), 1559(W3), 1460(W5), 1358-1338(W7; tryptophan doublet) and 1237 cm^{-1} (W10). The UVRR bands of Tyr are observed at 1619 (Y8a), 1206(Y7a) and 1177 cm^{-1} (Y9a).

The difference spectrum between deoxy and CO-bound states (Fig 2-2.1D) indicates that the intensity of the W1 and/or Y8a and W3 bands of Trp become weaker upon binding of

a ligand, and the frequency of Y9a of Tyr is shifted. The spectral changes are qualitatively same between traces (D) and (E), suggesting that the structural change is not caused by steric repulsion between the bound ligand and nearby residues. In both difference spectra, there are some peaks that cannot be assigned to Trp and Tyr bands. To clarify the origin of such bands, UVRR spectra of Fe(II)(PP)(2-MeIm) and COFe(II)(PP)(2-MeIm) were examined.

Figure 2-2.2 shows UVRR spectra of five-coordinated high-spin Fe(II)(PP)(2-MeIm) (A), six-coordinated low-spin COFe(II)(PP)(2-MeIm) (B) and the solvent (C) containing 2-MeIm and 1 % CTAB. Fe(II)(PP)(2-MeIm) and COFe(II)(PP)(2-MeIm) can be regarded as model compounds of the heme moiety of deoxyMb and COMb, respectively. 2-Methylimidazole is presumably under preresonance upon excitation at 244 nm and accordingly, gave many Raman bands (at 1564, 1519, 1492, 1402, and 1357 cm^{-1}) at pH 8.3. The relative intensity of two strong bands at 1492 and 1519 cm^{-1} were sensitive to pH, because the former and the latter were derived from an imidazole and an imidazolium ion, respectively.

The difference spectrum (Fig.2-2.2D), five-coordinate minus six-coordinate complexes, exhibits positive peaks at 1622 and 1304 cm^{-1} and negative ones at 1496 and 1370 cm^{-1}. Since the contributions from free 2-MeIm and CTAB should be completely canceled in the difference spectrum, these difference peaks must arise from the heme. Accordingly, the negative peaks at 1496 and 1370 cm^{-1} and positive peaks at 1622 and 1304 cm^{-1} in trace D are assigned to the bands of CO-bound and deoxy forms of iron porphyrin, respectively. It is clear that the difference peak at 1560 cm^{-1} in Figure 2-2.1 arises from changes of Trp residues, but it is necessary to subtract the difference spectrum of heme (Fig.2-2.2D) from the observed deoxyMb-minus-COMb UVRR spectra to abstract the contribution of the globin moiety for bands overlapped with the porphyrin bands. Note that Tyr bands are not overlapped with porphyrin bands.

Although swMb has three Tyr residues (Tyr103, Tyr146 and Tyr151), most of mammalian Mbs have two Tyr residues, that is, Tyr103 and Tyr146. To determine the Raman contribution of Tyr151, the UVRR spectra of hMb were examined. Figure 2-2.3(A) and (B) show the 244-nm excited RR spectra of deoxyMb and COMb, respectively, and traces C, D and E delineate the difference spectra, deoxy minus CO forms of swMb, hMb and the model compound, respectively. Apparently, the deoxy-minus-CO difference spectra of hMb and swMb are alike regarding bands at 1620, 1560, 1497 and 1370 cm^{-1} bands, to which the heme moiety also seems to contribute. The inset of Figure 2-2.3 shows the expanded spectra of swMb and hMb in the frequency region of 1230-1140 cm^{-1}, where Y7a (1208 cm^{-1}) and Y9a (1177 cm^{-1}) bands appear.

The solid and broken lines in the upper two spectra of the inset show the raw spectra of deoxy and CO-bound forms, respectively, and the lower two curves represent the deoxy-minus-CO difference spectra of swMb (SW) and hMb (Horse). It is clear that the differential peak at 1177 cm^{-1} (Y9a) is present for swMb but absent for hMb, indicating that the Tyr difference band of swMb arises from Tyr151. Except for this differential peak, spectra C and D are quite similar. It means that Tyr146 undergoes no environmental change upon ligand binding despite of its proximity to Tyr151. Although the X-ray structural analyses of swMb and hMb pointed out the presence of four different regions between the two proteins (*33*), those regions seem to have no influence on the vibrational spectra of Tyr and Trp residues. To get insight into the structural change of Tyr151, seen for swMb upon CO binding in Fig. 2-2.3C, solvent dependence of UVRR spectrum of a model compound of Tyr residue was investigated.

Figure 2-2.4a shows the UVRR spectra in the 1150-1250 cm^{-1} region of N-acetyl-tyrosine-ethyl ester, which is a model compound of Tyr residue. This compound is soluble in water-organic mixed as well as neat organic solvents, and accordingly solvent effects of RR

spectra could be examined. In Figure 2-2.4a, hydrophobicity of solvent decreases from A toward E. Although three Tyr bands including Y9a (1177 cm^{-1}), Y7a (1205 cm^{-1}) and Y7a' (~1250 cm^{-1}), are expected to appear in this region, the intensity of Y7a' is very weak upon excitation at 244 nm. The absolute intensities of Y7a and Y9a bands excited at 244 nm decrease and their peaks shift to higher frequencies as hydrophobicity decreases. Moreover, their relative intensities, Y9a/Y7a exhibit a good correlation with the hydrophilic/hydrophobic environment as plotted in Fig.2-2.4b. It is noted, however, that N-acetyl-tyrosine-ethyl ester was not soluble in pure cyclohexane to the extent to yield a Raman band but it became soluble in cyclohexane mixed with a small amount of propanol, giving rise to intense Y9a band (spectrum A). When a strong proton acceptor, HMPA, was added to this mixed solvent, the Y9a band became further stronger as depicted by spectrum A'. Thus, it would be more precise to address that Y9a band becomes stronger when phenolic hydrogen serves as a hydrogen-bond donor in hydrophobic environments.

Intensities of Y9a band of hMb in the deoxy and CO-bound states, shown in Figure 2-2. 3A and B, respectively, are fairly strong. According to the solvent effect of the Tyr model compound, Tyr103 and Tyr146 are deduced to stay in hydrophobic environments and to serve as a hydrogen bond donor. Since the intensities of Y7a and Y9a were nearly the same in the swMb-minus-hMb difference spectrum (data not shown), the environment of Tyr151 is inferred to be hydrophilic on the basis of Figure 2-2. 4b. The study of pH titration with absorption spectra (*34*) and X-ray crystal structures (*6*) indicate that Tyr151 is exposed to solvent. Consequently, the UVRR results for a solution is consonant with the X-ray results for crystals.

As noted for Figure 2-2.3, the intensity of the 1560 cm^{-1} band of Trp changes upon CO binding. SwMb has two Trp residues, Trp7 and Trp14, which are both on the A-helix, and are highly conserved. To determine which Trp residue undergoes environmental changes upon

ligand binding, we tried to isolate a spectrum of individual Trp residues by preparing two mutants, W7F and W14F. These two mutants had correct folding, since their absorption spectra were the same as that of native Mb, and the 244-nm excited UVRR spectra were quite similar to that of native Mb except for the bands derived from Trp residues. Figure 2-2.5a displays the raw UVRR spectra of deoxy- (solid lines) and CO-bound forms (broken lines) of native (A), W14F (B), and W7F (C) swMb mutants. The ordinate scales of all spectra are normalized with Tyr bands; Y7a and Y9a, which are different between deoxyMb and MbCO but are the same among the three spectra.

Accordingly, in their deoxy-minus-CO difference spectra delineated in Figure 2-2.5b, the magnitudes of differential patterns at Y9a of traces (A), (B) and (C) are alike. Nevertheless, the magnitudes of difference peaks of Trp at 1561 cm^{-1} are different among them. Note that spectra (B) and (C) in Figure 2-2.5b represent the environmental changes of Trp7 and Trp14, respectively, while spectrum (A) reflects sum of both. Spectrum (D) depicts the deoxy-minus-CO difference spectrum of the heme model compound, and it does not give any difference peak around 1560 cm^{-1}. To demonstrate the protein contribution more clearly, the heme contributions in traces (A-C) were removed by subtracting trace (D) from them. The double difference spectra are delineated in Fig.2-2.5c. It is evident that the most part of the change of Trp W3 band of native Mb comes from Trp7. The W3 band of Trp7 (W14F mutant) is slightly shifted to a lower frequency and exhibits intensity reduction upon ligand binding. In addition, the difference peak around 1620 cm^{-1} involves the contribution not only from the vinyl group of heme but also from Tyr and Trp residues.

More detailed examination of the W3 bands of W7F and W14F mutants showed appreciable difference of their peak frequencies. Figure 2-2.6 displays the expanded raw spectra of the CO-bound forms of native (solid line), W14F (dotted line) and W7F mutants (broken line) in the W3 (left panel) and W7 regions (right panel). The W3 band of Trp7

(W14F mutant) appeared at 1557 cm^{-1} whereas that of Trp14 (W7F mutant) did at 1560 cm^{-1}. Since the digital sum of the spectra of Trp7 and Trp14 well reproduced the observed spectrum of native Mb, this difference of frequency is considered to reflect the nature of each Trp residue.

The I_{1358}/I_{1338} intensity ratio of the W7 doublet is known to be sensitive to environments of the indole side-chain of Trp residue, being larger for more hydrophobic environments. This is owed to the fact that the 1358/1338 cm^{-1} doublet arises from Fermi resonance between the N_1C_8 stretching fundamental (W7) and the combination of out-of-plane bending vibrations (*35*). As demonstrated in the right Figure of Figure 2-2.6, the I_{1358}/I_{1338} ratio of Trp 14 is 1.3, which is larger than 1.0 for Trp7, consistent with the fact that Trp 14 is buried in more hydrophobic environment than that of Trp7. Since the intensity ratio exhibits little change between the deoxy and CO-bound states as shown in Figure 2-2.5c, the difference between Trp14 and Trp 7 seems to be retained upon ligand binding/dissociation. Because of appreciable contribution from the heme moiety to the difference spectrum in this frequency region, it would be more reliable to use W3 rather than W7 band to discuss the environments of Trp residues when a heme is present in the protein.

The pKa of distal His of Mb is thought to be 4.4 (*16*). This value is extremely lower than an ordinary value (pKa = 6.65) (*30*) and might be ascribed to the hydrophobic environments of heme pocket. On the other hand, the structure of n-BuICMb has been solved with X-ray crystallography, reported in the data bank (*36*), which indicated that the distal His is swung up to protein surface. If this position is exposed to solvent, distal His at an appropriate pH might be protonated at the exposed position but not protonated in the heme pocket. Only a protonated histidine is known to yield a UVRR band around 1410 cm^{-1} in D_2O (*29, 30*). Accordingly, we examined UVRR spectra of COMb and n-BuICMb in D_2O between pH 6.4 and 4.6, and the results are displayed in Figure 2-2.7. Deuterated

histidinium in Mb gave a Raman band at 1407 cm^{-1} and its intensity increased as decrease of pH as shown by spectra (A-D) in Figure 2-2.7. Although the intensity of the 1407 cm^{-1} band of free histidine exhibited no pH dependence below pD 5, that of Mb still increased there due to the presence of His residues with different pKa. Nevertheless, the difference spectra between COMb and n-BuICMb delineated by traces (E-F) gave no peak at 1407 cm^{-1} between pD 6.4 and 4.6, while differences derived by heme were seen at 1500 and 1370 cm^{-1}.

Discussion

Trp Conformation In the 244-nm excited UVRR spectra of Mb, some RR bands of heme unexpectedly appeared in addition to RR bands of Tyr and Trp residues. Since they change between the deoxy and ligand-bound states, it is important to get rid of the contribution of the heme from the protein spectra. On the basis of the difference spectra between the deoxy and CO-bound states of the model complex shown in Figure 2-2.2, some difference peaks shown in Figure 2-2.1 were attributed to the heme moiety; the positive band at 1622 cm^{-1} is assigned to the C=C stretching of vinyl group of heme side-chain (*37-40*), and the negative bands at 1495 and 1371 cm^{-1} arise from the ν_3 and ν_4 modes of porphyrin skeleton, respectively. Although Trp doublet around 1358/1338 cm^{-1} were overlapped with the heme bands, W3 band of Trp and Y9a band of Tyr have no overlapping with heme bands.

It has been pointed out by Miura et al. (*41*) that the W3 frequency of Trp residue has a good correlation with torsion angle (θ) of the C_2-C_3-C_β-C_α part of indole side-chain. The two frequencies thus observed are plotted against the X-ray determined torsion angles (*6*) and compared with the expected ones in Figure 2-2.8, where the solid line denote the empirical curve newly fitted with cos(3θ) as a variable. While the W3 frequency of Trp7 falls on the empirical curve, the frequency of Trp14 is higher than the expected value. Trp14 is buried in

the folded protein and its environment seems to be more hydrophobic than that of Trp7 as discussed below, but this abnormal frequency and intensity of W3 band of Trp14 cannot be explained satisfactorily with our current knowledge. One possibility is that a hydrogen bond to the indole nitrogen atom affects this frequency and intensity, and the other is that the torsion angle of C_3-C_β-C_α-N, which might influence on the W3 frequency, is different from those of the Trp model compounds used in the plots (*41*). It is emphasized that the present study provides the first separate detection of individual spectra of Trp 7 and Trp 14.

Relation with X-ray Results High resolution X-ray crystallographic data are available for various states of native and mutant Mbs (*6, 17, 42-44*). Although the location of Tyr 151 is different among several analyses, it can be summarized that Tyr 151 has two positions; one is directed towards the F-G loop (inward) and the other is directed towards F helix (outward). In the recent data with near-atomic resolution, Bartunik and coworkers (*6*) addressed that Tyr151 pointed to F helix in the CO-bound form but had two directions in the deoxy state, while Phillips and co-workers (*43*) pointed out that Tyr151 had two directions in both states, occupation density of which was changed in the two states. Common tendency in these two analyses is that Tyr151 is predominantly directed inward in the deoxy state and outward in the CO-bound state. The flexibility and fluctuation of C-terminus may explain these difference and similarity.

The Y9a band of swMb exhibited the deoxy-minus-CO difference peak, but that of hMb did not (Figure 2-2.3). Therefore, this difference peak is attributed to Tyr151, indicating that Tyr151 of swMb is flexible in the solution and placed in different environments between the deoxy and CO-bound states. Its frequency is lower and intesity is higher in the deoxy state than in the CO-bound state. This means that Tyr151 is placed in more hydrophobic environments and works as a hydrogen bond donor in the deoxy than CO-bound state on the

basis of the results shown in Figure 2-2.4. Probably the hydrophobic environments cause a red shift of the electronic La absorption band, approaching to the Raman excitation wavelength, and as a result, resonance enhancement of Raman intensity becomes larger upon excitation at 244 nm. Although there was a question whether lattice forces in crystals might prevent the C-terminal flexibility in solution from the tertiary conformational change, the UVRR results for solution is compatible to the X-ray results for crystals, if the inward and outward locations of Tyr151 correspond to the more hydrophobic and hydrophilic environments, respectively, suggesting the absence of large lattice constrain. It is reported that the Y9a frequency depends on the torsion angle of phenolic C-O-H group of tyrosine side chain (*45*). Since the distance of Tyr151 directed towards the G-F loop allows to make a hydrogen bond with an oxygen atom of main chain Lys98, the torsion angle of hydroxyl group would be changed through the movement of Tyr151.

Importance of the movements of the C terminal residues have been anticipated from the studies of Hb cooperativity, in which removal of the C terminal residues such as Arg141α and His146β from HbA raised oxygen affinity and made the oxygen binding equilibrium non-cooperative *(46,47)*. In fact, it was demonstrated by the UVRR study of Hb quaternary structure *(48)* that the penultimate tyrosines (Tyr140α and Tyr145β) exhibit clear spectral changes upon ligand binding. The present result is consistent with these features, which seem to be a common basic property of heme proteins.

Open/Close Forms A pathway of an external ligand from solvent to heme could not be found in the stationary structures of deoxy- and CO-bound forms of Mb. However, X-ray crystallographic studies pointed out that distal His was rotated and swung up toward the protein surface under extreme conditions and that a pathway of a ligand was opened. Such a structure has been found for swMb with bulky ligand such as imidazole (*18, 46*),

phenylhydrazine (*19*), and alkyl isocyanide (*20*), and are regarded as a model of the open form of the pathway. Yang and Phillips Jr. (*17*) found that a similar open form was produced with COMb at pH 4, owing to the electrostatic interactions between the protonated distal His in hydrophobic environment of pocket and bound CO. Due to this structural change, the Fe-CO stretching band of COMb shifts from 507 cm^{-1} at pH 7 to 488 cm^{-1} at pH 4 (*16*). Furthermore, a transient open form was detected in the time resolved UVRR experiments using CO photolysis of H64Y mutant (*22*). These studies supported the idea that the movement of distal His is involved in the open/close changes of a ligand pathway. The UVRR spectrum of isocyanide adduct of Mb shown in Figure 2-2.1C, which is considered to adopt the distal His swung up structure, is unexpectedly very close to that of CO-bound form (Fig.2-2.1B). This means that the movement of His64 is localized around heme pocket and scarcely affects the overall protein structure.

N-deuterated imidazolium ring of histidine residue gave a clear UVRR band at 1407 cm^{-1} for Mb upon excitation at 244 nm. We expected that the open form could be detected if pKa's of His 64 was higher in the open form than in the closed form. The stationary open form is expected to generate when a large ligand is bound to the heme iron (*18-20, 49*), although it is not expected for COMb above pH 4.4. Accordingly, UVRR spectra of the CO- and n-BuIC-bound Mbs were examined at various pD values from 6.4 to 4.6 as shown in Figure 2-2.7 (A-D). Although intensity of the 1407 cm^{-1} band of deuterated histidyl imidazolium increased at lower pH, no peak was recognized in the difference spectra between the CO- and n-BuIC bound forms as depicted by traces (E-H) in Figure 2-2.7. This would mean that the pKa of His64 within the heme pocket is not so different from that in the position of the open form, or alternatively the S/N ratio in the present experiment was not high enough to extract the contribution of one His residue from the twelve His residues included in swMb.

Conformational Change of Mb upon Ligand Binding The proximal His is known to play a significant role in the T-R transition of Hb, and much attention has been paid to it (*25*). In contrast, studies on proximal side of Mb are much less. Although structural changes of Mb upon ligand binding might be smaller than those of Hb, the present study demonstrated that Trp7 and Tyr151 in the N- and C-terminal regions, respectively, undergo appreciable changes. Mb consists of eight helices and the heme is embedded between E and F helices, while Trp7 and Tyr151 are contained in A and H helices, which are interacting with E and F helices, respectively. Therefore, some changes in helix-helix interactions must be involved in the ligand biding. Seno and Go (*50*) carried out normal coordinate analysis of deoxyMb and on the basis of 151 modes below 40 cm^{-1} interpreted the atomic displacements of globin moiety upon ligand binding. The largest displacements were contained in the region from F helix to the beginning of G helix, and secondary regions are A and C helices, CD-corner, E helix, and C-terminal side of H helix. All the Trp and Tyr residues are contained in the secondary sensitive regions. The appearance of the ligand binding effects only on Trp7 and Tyr151 residues might be partly due to the fact that these side chains are located near the surface of the protein and accordingly favorable for the detection of slight conformational change, and partly to that the slight structural change of the active site is amplified in proportion to the distance.

Causprove and Dyer (*13*) measured time-resolved FTIR spectra of photolyzed MbCO, pointing out that the change of amide I band with the time constants of 6-8 ps was mainly due to the proximal side, through F-helix. They deduced that the contribution to the structural change from the distal side was smaller than that from the proximal side. Accordingly, some rearrangements of helix topology caused by movement of the heme iron would be transmitted to Tyr151 in the C-terminus and Trp7 near the N-terminus of the polypeptide chain. This is

compatible to the result from the time resolved fluorescence study, which pointed out appreciable difference in the quenching time of the Trp7 fluorescence between the deoxy- and CO-forms (*15*).

The present results demonstrate that a structural change of heme upon ligand binding is communicated to the terminal parts of polypeptide chain. While the environmental and structural changes of Trp7 and Tyr151 imply an overall conformational change of protein, it is difficult to address how this subtle conformational change is connected to the function of Mb. However, it became clear from the present study that a small structural change in the buried portion of a protein causes a rather large conformational change at the surface and can communicate it to an adjacent subunit to regulate its function such as signal transduction.

Reference

1. Rodgers, K. R. (1999) *Curr Opin Chem Biol 3*, 158-67.
2. He, Y., Gaal, T., Karls, R., Donohue, T. J., Gourse, R. L., and Roberts, G. P. (1999) *J Biol Chem 274*, 10840-5.
3. Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. (1991) *Nature 350*, 170-2.
4. Stone, J. R., and Marletta, M. A. (1996) *Biochemistry 35*, 1093-9.
5. Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M. A. (2000) *Biochemistry 39*, 2685-91.
6. Kachalova, G. S., Popov, A. N., and Bartunik, H. D. (1999) *Science 284*, 473-6.
7. Osapay, K., Theriault, Y., Wright, P. E., and Case, D. A. (1994) *J Mol Biol 244*, 183-97.
8. Causgrove, T. P., and Dyer, R. B. (1993) *Biochemistry 32*, 11985-91.
9. Sage, J. T., Schomacker, K. T., and Champion, P. M. (1995) *J. Phys. Chem. 99*, 3394-

3405.

10. Springer, B. A., Sligar, S. G., Olson, J. S., and Phillips, G. N., Jr. (1994) *Chem. Rev. 94*, 699-714.

11. Olson, J. S., and Phillips, G. N., Jr. (1996) *J Biol Chem 271*, 17593-6.

12. Olson, J. S., and Phillips Jr., G. N. (1997) *J. Biol. Inorg. Chem. 2*, 544-552.

13. Causgrove, T. P., and Dyer, R. B. (1996) *J. Phys. Chem. 100*, 3273-3277.

14. Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J., and Eaton, W. A. (1994) *Biochemistry 33*, 5128-45.

15. Gryczynski, Z., and Bucci, E. (1998) *Biophysical Chem. 74*, 187-196.

16. Ramsden, J., and Spiro, T. G. (1989) *Biochemistry 28*, 3125-8.

17. Yang, F., and Phillips, G. N., Jr. (1996) *J Mol Biol 256*, 762-74.

18. Bolognesi, M., Cannillo, E., Ascenzi, P., Giacometti, G. M., Merli, A., and Brunori, M. (1982) *J Mol Biol 158*, 305-15.

19. Ringe, D., Petsko, G. A., Kerr, D. E., and Ortiz de Montellano, P. R. (1984) *Biochemistry 23*, 2-4.

20. Johnson, K. A., Olson, J. S., and Phillips, G. N., Jr. (1989) *J Mol Biol 207*, 459-63.

21. Lai, H. H., Li, T., Lyons, D. S., Phillips, G. N., Jr., Olson, J. S., and Gibson, Q. H. (1995) *Proteins 22*, 322-39.

22. Mukai, M., Nakashima, S., Olson, J. S., and Kitagawa, T. (1998) *J. Phys. Chem. 102*, 3624-3630.

23. Hu, X., Rodgers, K. R., Mukerji, I., and Spiro, T. G. (1999) *Biochemistry 38*, 3462-7.

24. Matsukawa, S., Mawatari, K., Yoneyama, Y., and Kitagawa, T. (1985) *J. Am. Chem. Soc. 107*, 1108-1113.

25. Kitagawa, T. (1988) in *Biological Applications of Raman Spectroscopy* (Spiro, T. G.,

Ed.) pp 97-131, Wiley, New York.

26. Austin, J. C., Rodgers, K. R., and Spiro, T. G. (1993) in *Method in Enzymology* (Riordan, J. F., and Vallee, B. L., Eds.) pp 374-396, Academic press inc., San Diego.
27. Kitagawa, T. (1992) *Prog. Biophys. Molec. Biol. 58*, 1-18.
28. Harada, I., and Takeuchi, H. (1986) *Advances in Spectroscopy: Spectroscopy of biological System*, 113-175.
29. Lord, R. C., and Yu, N. T. (1970) *J Mol Biol 51*, 203-13.
30. Tasumi, M., Harada, I., and Takamatsu, T. (1982) *J Raman Spec. 12*, 149-151.
31. Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., and Olson, J. S. (1989) *J Biol Chem 264*, 3057-60.
32. Aki, M., Ogura, T., Shinzawa-Itoh, K., Yoshikawa, S., and Kitagawa, T. (2000) *J. Phys. Chem. 104, 10765-10774.*
33. Evans, S. V., and Brayer, G. D. (1988) *J Biol Chem 263*, 4263-8.
34. Uyeda, M., and Peisach, J. (1981) *Biochemistry 20*, 2028-35.
35. Harada, I., Miura, T., and Takeuchi, H. (1986) *Spectrochimica Acta 42a*, 307-312.
36. Smith, R. D., Olson, J. S., and Phillip Jr., G. N. (1997) X-ray structure of n-BuICMb was referred from Protein Data Bank, PDB code 104M.
37. Choi, S., Spiro, T. G., Langry, K. C., Smith, K. M., Budd, D. L., and La Mar, G. N. (1982) *J. Am. Chem. Soc. 104*, 4345-4351.
38. Choi, S., Spiro, T. G., Langry, K. C., and Smith, K. M. (1982) *J. Am. Chem. Soc. 104*, 4337-4344.
39. Devito, V. L., and Asher, S. A. (1989) *J. Am. Chem. Soc. 111*, 9143-9152.
40. Devito, V. L., Cai, M., Asher, S. A., Kehres, L. A., and Smith, K. M. (1992) *J. Phys. Chem. 96*, 6917-6922.
41. Miura, T., Takeuchi, H., and Harada, I. (1989) *J. Raman Spectrosc. 20*, 667-671.

42. Vitkup, D., Petsko, G. A., and Karplus, M. (1997) *Nat Struct Biol 4*, 202-8.

43. Quillin, M. L., Arduini, R. M., Olson, J. S., and Phillips, G. N., Jr. (1993) *J Mol Biol 234*, 140-55.

44. Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R. M., and Schlichting, I. (1999) *Biophys J 77*, 2153-74.

45. Takeuchi, H., Watanabe, N., Satoh, Y., and Harada, I. (1989) *J. Raman Spectrosc. 20*, 233-237.

46. Kilmartin, J. V., Hewitt, J. A. and Wooton, J. F. (1975) *J. Mol. Biol. 93*, 203-18.

47. Fung, L. W.-M. and Ho, C. (1975) *Biochemistry 14*, 2526-35.

48. Nagai, M. Wajcman, H., Lahary, A., Nakatsukasa, T., Nagatomo, S., and Kitagawa, T. (1999) *Biochemistry, 38*, 1243-51.

49. Lionetti, C., Guanziroli, M. G., Frigerio, F., Ascenzi, P., and Bolognesi, M. (1991) *J Mol Biol 217*, 409-12.

50. Seno, Y., and Go, N. (1990) *J Mol Biol 216*, 111-26.

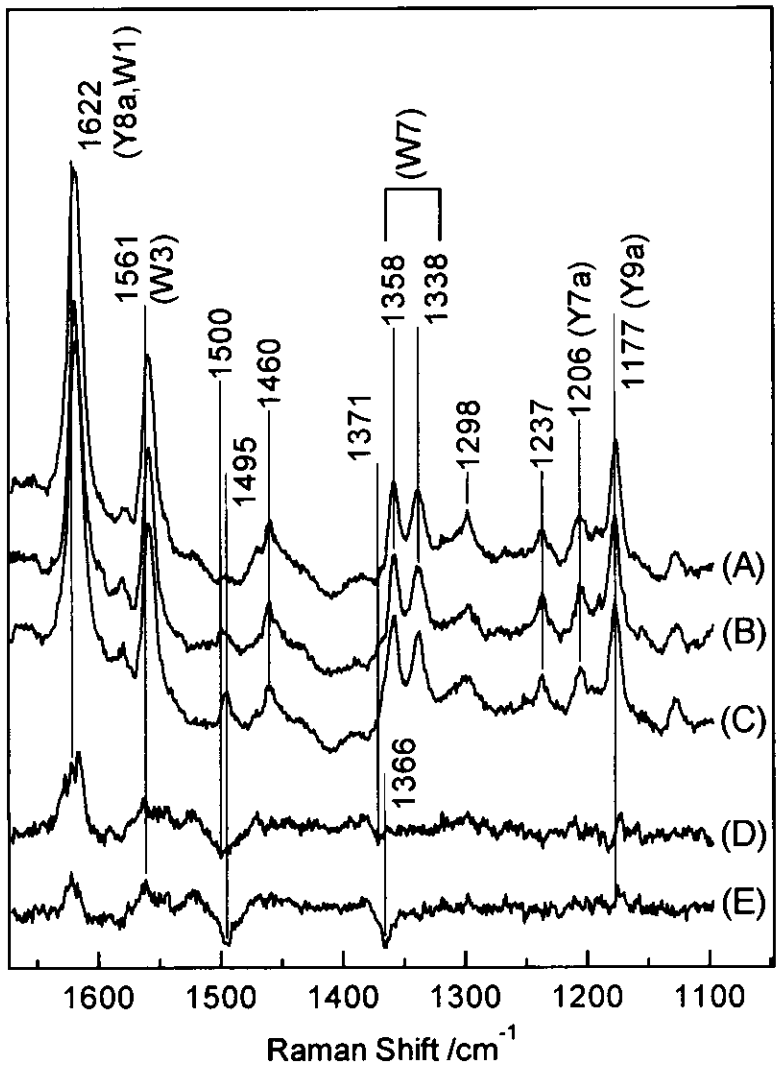


Figure 2-2.1. The 244-nm excited UVRR spectra of deoxy Mb (A), COMb (B), and *n*-BuICMb (C) of native swMb. The spectra D and E are differences, (A) minus (B) and (A) minus (C), respectively. The protein concentration was 370 μM in 50 mM potassium phosphate buffer, pH 7.0. *n*-BuICMb was derived by adding *n*-butylisocyanide to the deoxyMb solution to yield its final concentration of 2.4 mM at pH 7.0. The laser power at sample point was 0.12 mW and the total exposure time was 120 min.

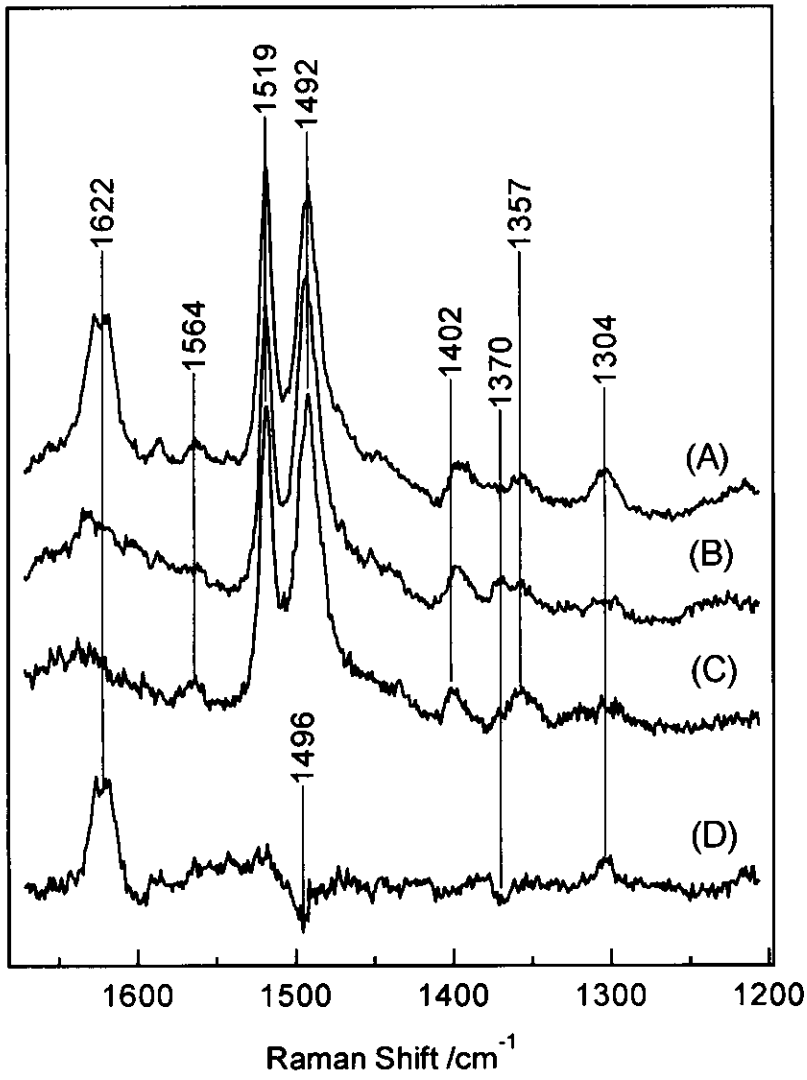


Figure 2-2.2. UVRR spectra of iron(II)-protoporphyrinIX 2-methylimidazole complexes in 50 mM borate buffer, pH 8.0, containing 1% CTAB and 20 mM 2-MeIm; five-coordinate deoxy form (A), six-coordinate CO-bound form (B) and their difference (D = A – B). Spectrum C denotes the spectrum of solvent containing 1% CTAB and 20 mM 2-MeIm. The concentration of iron-porphyrin was 500 μM.

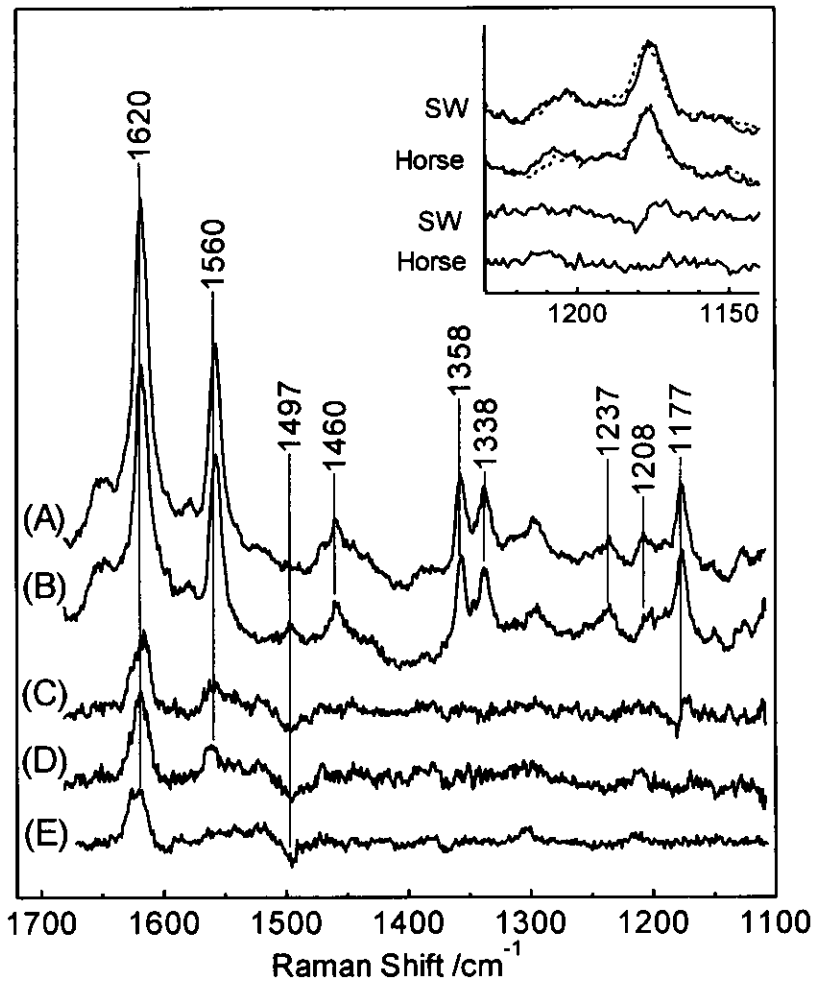


Figure 2-2.3. The UVRR spectra of the deoxy (A) and CO-bound (B) forms of hMb excited at 244 nm. Spectra C, D, and E represent the differences, deoxy-minus-CO spectra of swMb, hMb, and 2-MeIm adduct of iron porphyrin, respectively. The inset shows the expanded spectra in the Y7a and Y9a region of swMb and hMb. The top and second traces delineate the raw spectra of swMb and hMb, respectively, in which solid lines and broken lines denote the deoxy and CO-bound forms, respectively. The third and fourth traces indicate the deoxy-minus-CO difference spectra of swMb and hMb, respectively.

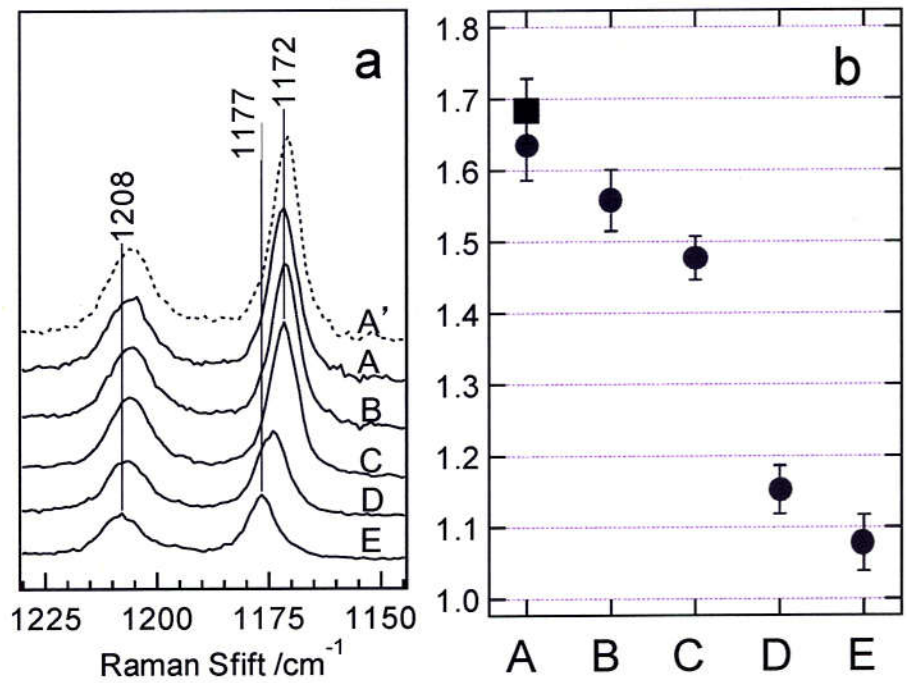


Figure 2-2.4. Solvent effects on the Y9a and Y7a bands of N-acetyl-tyrosinethyl ester (panel a) and the Y9a/Y7a intensity ratio (panel b). Solvents; A) cychohexane:propanol = 9:1, B) cyclohexane:propanol = 1:1, C) neat propanol, D) propanol:water = 1:1, E) propanol:water = 1:9. Spectrum A' in panel a and square in panel b denote the case in which a strong proton acceptor (HMPA) was added to the mixture A with the concentration of 2.0 % (v/v).

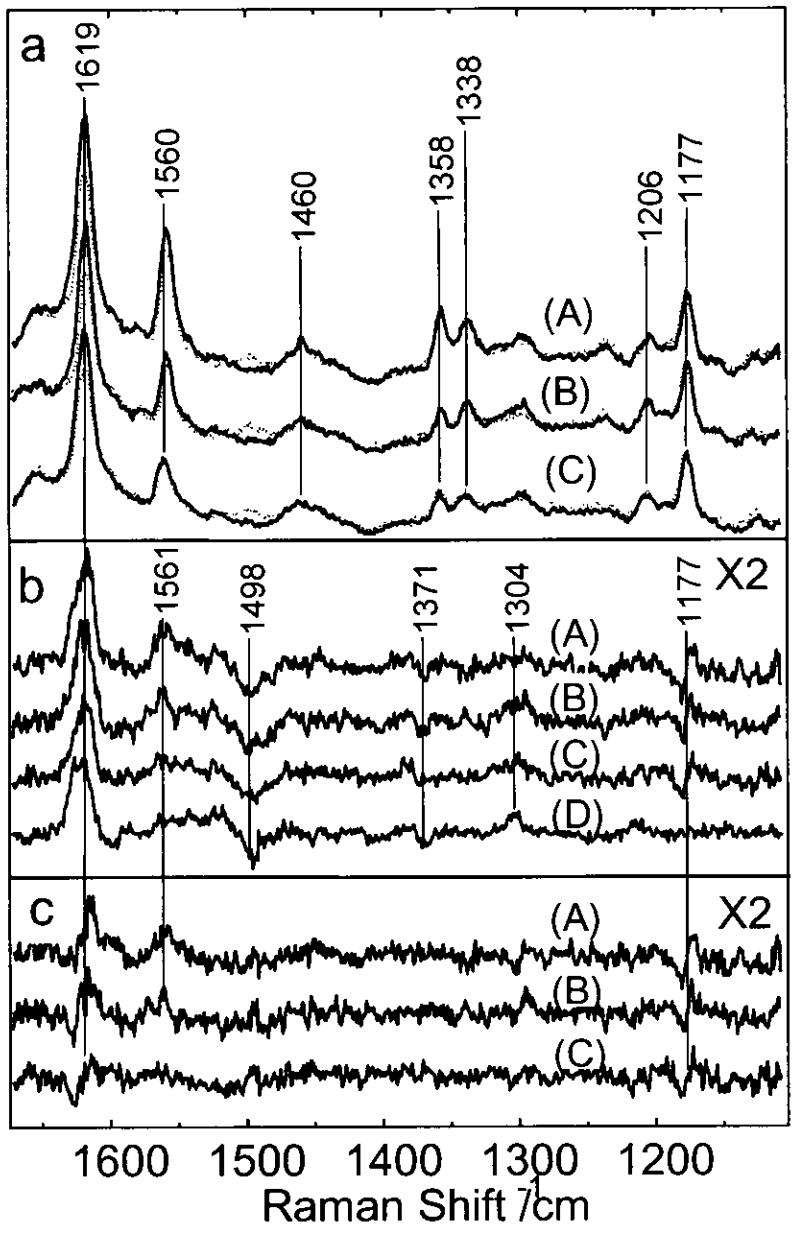


Figure 2-2.5. Panel a: UVRR spectra of deoxy (solid lines) and CO-bound forms (broken lines) of native swMb (A), W14F mutant (B), and W7F mutant (C). panel b: The deoxy-minus-CO difference spectra of native swMb (A), W14F (B), W7F (C) and 2-MeIm adduct of iron-porphyrin (D). The ordinate scales are expanded by a factor of 2 compared with those in panel a. Panel c: The double difference spectra, spectra (A-C) minus spectrum (D) of panel b. The ordinate scales are expanded by a factor of 2 compared with those in panel a.

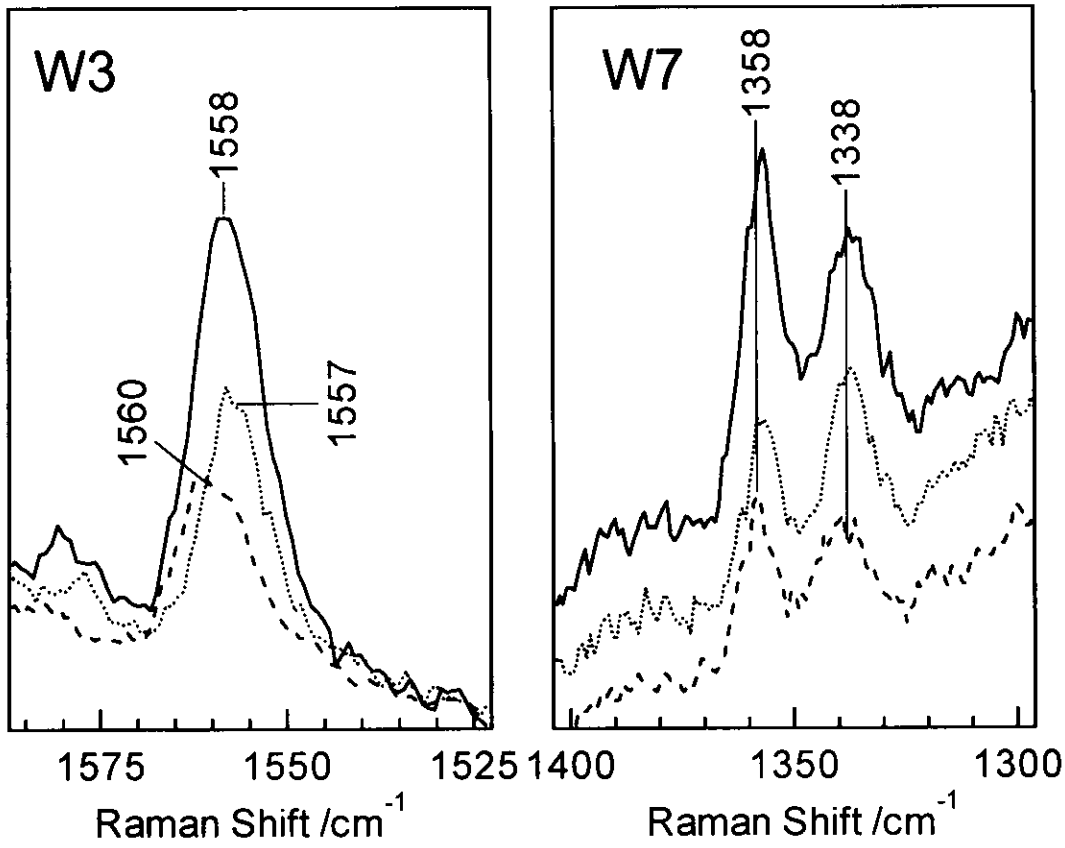


Figure 2-2.6. The expanded spectra in the W3 and W7 regions of CO-bound native swMb (solid line), W14F mutant (dotted line) and W7F mutant (broken line).

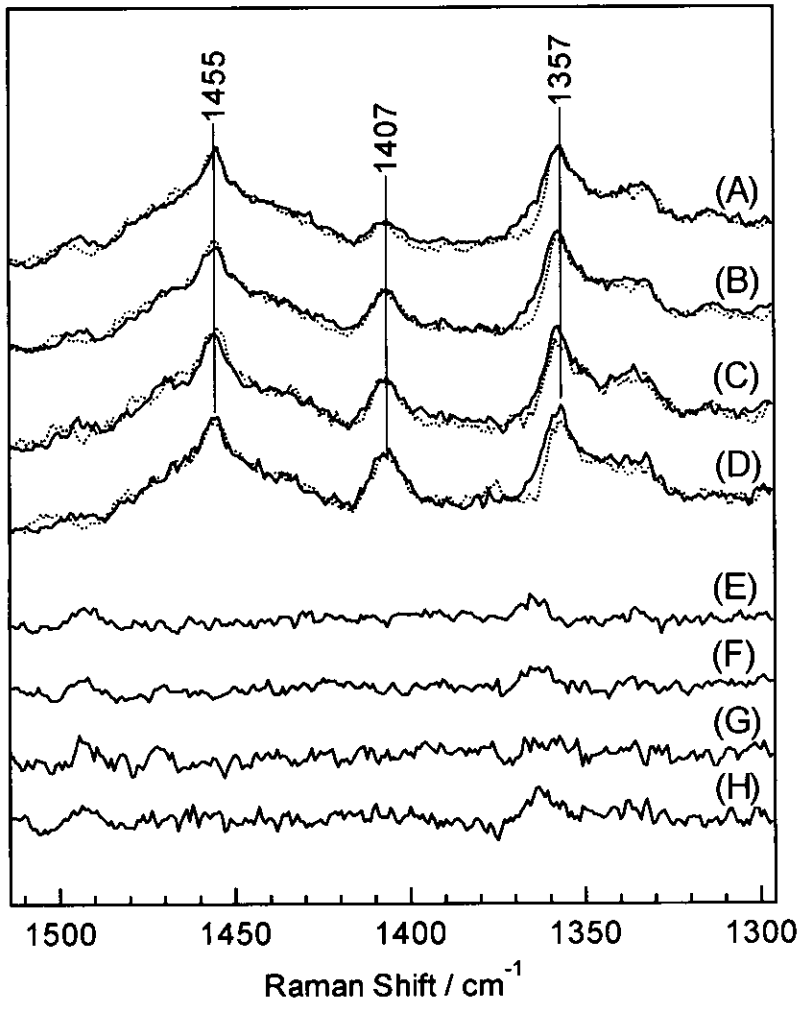


Figure 2-2.7. UVRR spectra of *n*-BuICMb (solid lines) and COMb (broken lines) at various pD values in D_2O. A) pD 6.4, B) pD 5,2, C) pD 4.9, and D) pD 4.6. The difference spectra, *n*-BuICMb minus COMb, of spectra (A-D) are delineated by traces (E-H), respectively. The ordinate scales of (E-H) are the same as those of (A-D). The pD values are regarded as direct reading values on the pH meter.

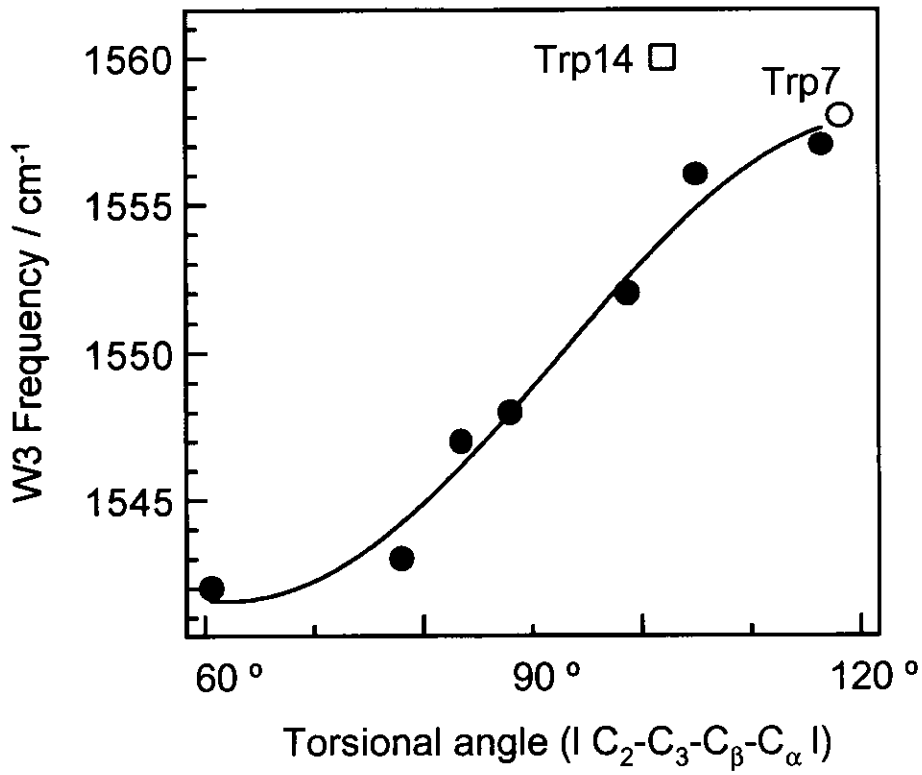


Figure 2-2.8. The correlation between W3 frequencies and C_2-C_3-C_β-C_α torsional angle (θ) of tryptophan residues found by Miura et al. *(41)*. Closed circles denote the original plot *(41)* and open marker indicate the data obtained in the present study. The solid line shows the curve best fitted in this study using $\cos(3\theta)$ as a variable.

2-3. Role of Distal Residues: Characterization of H64Y and F64V/H64Y Mutants

Abstract

The mutation of the distal histidine to tyrosine resulted in unique and interesting properties, that the oxidized form is stabilized. The 244 nm excited ultraviolet resonance Raman studies detected the band of tyrosinate that coordinated to the heme iron in H64Y. This tyrosinate band of H64Y did not change between pH 8.7 and 5.7, while that of F46V/H64Y mutant diminished at pH 5.7. The visible resonance Raman spectroscopy excited at 488 nm indicated that the two mutants have individually two coordination states of distal Tyr that yield two sets of $\nu_{Fe\text{-}O}$ / $\nu_{TyrC\text{-}O}$ bands at 598cm^{-1} /1289cm^{-1} and 578cm^{-1}/1301 cm^{-1}. The environment of heme pocket in CO-bound form are not largely different between two mutants on the basis of the observed of $\nu_{C\text{-}O}$ bands. Therefore, it is assumed that the abnormal pK of tyrosinate in H64Y and rapid autooxidation resulted from stabilization of the ferric 6-coordinated forms due to the steric hindrance by surrounding residues of distal Tyr.

Abbreviations: Mb: myoglobin, Hb: hemoglobin, UVRR: ultraviolet resonance Raman, RR: resonance Raman, EPR: electron paramagnetic resonance, EXAFS: extended X-ray absorption fine structure, XANES: X-ray absorption near-edge structure, Tyr^{-}: tyrosinate

Introduction

Myoglobin (Mb) is a good system to study about the recognition of a substrate at atomic level and various site-directed mutagenesis experiments target to distal residues, especially distal His (H64) (*1*). The substitution of distal His with Tyr (H64Y) shows an interesting and unique character that the autooxidation is abnormally rapid and stabilizes the met form with green color (*2*). Hemoglobin (Hb) M are classes of hemoglobins that have oxidized form in either α or β subunit, and in most of mutants the distal or proximal His is replaced to the Tyr residue and oxygen is not bound; these mutants include Hb M Boston [His(αE7)Tyr], Hb M Saskatoon [His(βE7)Tyr], Hb M Iwate [His(αF8)Tyr], Hb M Hyde Park [His(βF8)Tyr] and Hb M Milwaukee [Val (βE11)Glu](*3*).

The X-ray crystallographic analysis of H64Y suggested that distal Tyr was coordinated to the heme iron, thus 6 coordinated (Figure 2-3.1) (*4*). The substitution resulted in some distortion of heme plane and affected the adjacent residues, Leu29 and Lue104. Some spectroscopic studies of H64Y, such as visible resonance Raman spectroscopy (*5*), EPR (*4, 6*), EXAFS (*7*) and XANES (*6*) also indicated that 6-coordinated high spin in met form and two coordination isomers. These spectroscopic methods mainly give the information on the states of heme iron and porphyrin ring, but little about coordinated Tyr residue, except for visible resonance Raman spectroscopy that can detect the (Tyr)O-Fe and C-O stretching mode of Tyr residue and ring modes.

The heme cofactor is surrounded by the protein matrix for selective recognition of an exogenous ligand and protection from the autooxidaton. The pathway of an exogenous ligand from solvent to the heme iron may be made by the movement of distal His that swings up toward outside, that is open form, while the distal His with ordinary conformation is referred as the closed form. The open form of Mb can be produced by the

ligation of bulky ligand, such as imidazole (*8, 9*), phenylhydrazine (*10*), and alkyl isocyanide (*11*), and by the low pH owing to the electrostatic repulsion of protonated distal His (*12*). Additionally, the mutation of Phe46 (Figure 2-3.1) to Val, which is in contact with distal His, resulted in the open form (*13*).

In this section, ultraviolet resonance Raman (UVRR) spectroscopy is used to obtain more information about the heme coordinated Tyr residue, since the excitation at 244 nm selectively enhances the Raman band of a tyrosinate (Tyr^-). According to the expectation, the new band of Tyr^- can be detected clearly and the pK value of Tyr^- is largely different between the H64Y and F46V/H64Y. F46V/H64Y is double mutant in which distal Tyr will be more flexible than that of H64Y.

Materials and Methods

Preparation of the Samples H64Y was constructed by J.S.Olson and coworkers at Rice University (*14*). The double mutant, F46V/H64Y was constructed as described in the section 2-2. The purification and preparation of H64Y and F46V/H64Y follow the same procedure as those in the section 2-2. The sample solutions were prepared with the 25 mM borate buffer for pH 8.7, with 50 mM potassium phosphate buffer for pH 7.0 and with 25 mM citric buffer for pH 5.7.

Measurement of UV Resonance Raman Spectroscopy The measurements of 244-nm excited UVRR spectra were performed essentially the same instruments and procedures as these described in section 2-2. The laser power of 244 nm light was 0.12 mW at the sample point and measurement was carried out at 20 °C. The total exposure time is 30 min for one

spectrum, involving the replacement of the sample solution with new one every 5 min. The Raman frequencies were calibrated with cyclohexane and Raman intensity was normalized the W3 band of Trp residue at 1559 cm^{-1}.

Measurement of Visible Resonance Raman Spectroscopy The Visible resonance Raman spectra were excited with the 441.6 nm line from a He/Cd laser (Kinmon Electric, model CD4805R) and with the 415.4 nm line from a Kr^+ gas laser (Spectra-Physics, Beamlok 2060) with a power of 3 mW at the sample point. Raman scattered light was dispersed with a 100-cm single polychromator (Ritsu Oyo Kagaku, MC-100DG) equipped with a liquid nitrogen-cooled CCD detector (Princeton Instruments, UV/CCD-1340/400). All experiments were carried out at 20 °C with a quartz spinning cell rotated at 2200 rpm. Raman shifts were calibrated with the indene and CCl_4 and analyzed with the Igor Pro software (WaveMetrics).

Results

Figure 2-3.2 shows the 244 nm excited UVRR spectra of native, H64Y (dotted) and F46V/H64Y (broken) mutants of sperm whale metMb (solid) at pH 8.7. Raman bands were dominated by the contribution of Trp and Tyr residues and assigned to Trp and Tyr according to Harada and co-workers (*15*) which are labeled by W and Y, respectively. The UVRR spectra of H64Y, F46V/H64Y and native met Mbs can be overlapped at the W3 band of Trp residue, and it indicated those overall conformations are alike. New bands appeared in the spectra of two mutants at low frequency side of the strongest band at 1619 cm^{-1} (W1, Y8a). The difference spectra obtained by subtracting the native spectrum from

that of H64Y (B) and F46V/H64Y (C) are illustrated in Figure 2-3.2 and they gave this new band at 1600 cm^{-1}. At the same time, there are two bands at 1200 cm^{-1} and 1163 cm^{-1} (H64Y) or 1168 cm^{-1} (F46V/H64Y) in the difference spectra between each two mutants and native Mb. These new bands in the mutants are derived from tyrosinate (Tyr^-), phenoxyl anion, and we assign the band at 1600 cm^{-1} to Y8a , the band at 1205 cm^{-1} to Y7a and bands at 1163 and 1168 cm^{-1} to Y9a. The absorption spectrum of Tyr^- has a strong peak around 240 nm, while that of Tyr has a valley around 240 nm between the La and Bb bands. Therefore, the Raman band of Tyr^- excited at 244 nm is resonance enhanced and has strong intensity than that of Tyr residue. This is the first spectrum to detect the selectively the Tyr^- band for the protein matrix and serves as the direct evidence for Tyr coordination to the heme iron in the deprotonated form.

Although the normal Tyr residue has pK = 10 for conversion to the phenoxyl anion, H64Y has abnormally low pK, which was reported to be near pH 5.6, but the titration curve of pK was not completed (*6*). Figure 2-3.3 shows the UVRR spectra of two mutants at pH 8.7 (A, C) and 5.7 (B, D). The spectrum of H64Y mutant has the remaining Tyr^- band even at pH 5.7 and difference spectrum between the pH 8.7 minus pH 5.7 (spectrum E) indicated almost no peak. On the other hands, the spectrum of F46V/H64Y at pH 5.7 (D) diminishes the bands of Tyr^- (1600 cm^{-1}), and the pH difference spectrum (F) yielded a positive peak, indicating that the pK value of distal Tyr in F46V/H64Y is higher than H64Y. The pH dependent absorption changes of Q band (not shown) indicated that the pK of F46V/H64Y is located in neutral pH.

To consider the reason for the difference in pK between H64Y and F46V/H64Y, the visible resonance Raman (RR) spectra were occurred to detect the band of Fe-O stretching mode and (Tyr)C-O stretching mode in met form and Fe-C band in CO bounded state for

the comparison of the heme pocket. Figure 2-3.4 shows the visible RR spectra of native and two mutants of metMb excited at 488 nm. The Raman bands were labeled mainly according to Egeberg and coworkers (*5*) and Hu and coworkers (*16*). In 488 nm excitation, the band of Fe-O (Tyr) stretching mode and other two bands related to the Tyr residue are enhanced (*5, 17, 18*). The Fe-O stretching mode, $\nu_{Fe\text{-}O}$, was dissolved into two bands at 598 and 578 cm^{-1} for both mutants and this results was the same with the results of H64Y reported by Egeberg and coworkers (*5*). Although previous studies of H64Y reported only one broad band of the $\nu_{TyrC\text{-}O}$ mode at 1301 cm^{-1}, the $\nu_{TyrC\text{-}O}$ band in the spectrum of F46V/H64Y mutant exhibited apparently two separate bands at 1301 and 1289 cm^{-1}.

The frequency of ν_{10} and ν_2 band have a relation to the core size of porphyrin ring (*19*). The lower frequencies of two mutants compared with that of native state indicated that heme iron is high spin and 6 coordinated state. Thus, the ligand field of Tyr^{-} may not be so strong.

Figure 2-3.5 shows the visible resonance Raman spectra of CO bound form of native, H64Y, F46/V/H64Y and F46V excited 415.4 nm. The Fe-C stretching mode, $\nu_{Fe\text{-}C}$, appearing around 500 cm^{-1} is very sensitive to the environment of CO, that is environment of the heme pocket (*20*). The F46V mutant has a broader $\nu_{Fe\text{-}C}$ band and it could be deconvoluted into two bands; one is same as native state (507 cm^{-1}) and the other has a peak at 491 cm^{-1}. The protonation of distal His at low pH has an open conformation in which distal His swings up toward out side because of electrostatic repulsion (*12*) and the $\nu_{Fe\text{-}C}$ band of open form was shifted to 491 cm^{-1} (*21, 22*). Since the distal His of F46V mutant swing up in crystallographic study (*13*), the $\nu_{Fe\text{-}C}$ band of F46V indicated that it has an equilibrium between open and closed conformations, dominant in open conformation in

some degree.

On the other hand, the ν_{Fe-C} band of F46V/H64Y appears at the same position as that of H64Y, that is, at 488 cm^{-1}, (the H64Y spectrum B is overlaid on spectrum C of F46V/H64Y with a broken line). It indicated that distal Tyr of F46V/H64Y does not swing up unlike F46V and the environments of F46V/H64Y and H64Y around the heme pocket are not so different.

Discussion

Two Forms of Tyr Coordination to the Heme Iron The visible resonance Raman studies are consistent with the results of other studies that H64Y mutant has two structures with different coordination geometries to the heme iron (*4-7*). The ESR spectroscopy of H64Y characterized the two structures as follows; the main component has strong Fe-O-C(Tyr) bond with heme distortion and the second has a different bond length than main species and less distortion of heme plane (*4*). Although the spectrum of H64Y has not separated band of ν_{TyrC-O} mode, the additional studies of F46V/H64Y revealed that not only the ν_{Fe-O} band but also the ν_{TyrC-O} band separated to two bands. The studies of Hb M mutants with visible RR spectroscopy excited at 488 nm in the met form demonstrated the presence of two kinds of tendency that ν_{Fe-O}/ν_{TyrC-O} are 578 or 588 cm^{-1} / 1300 cm^{-1} (Hb M Saslkatoon, Hb M Hyde Park) and 603 cm^{-1}/ 1278 cm^{-1} (Hb M Boston). In the case of Hb M, these bands were single set in each mutant unlike Mb distal mutants. The empirical correlation of these two bands can be explained as that one species with the strong binding to heme, has the stronger Fe-O (603 cm^{-1}) bond and weaker TyrC-O (1278 cm^{-1}) bond, and the other with weak binding with weaker Fe-O (578 or 588) bond and stronger TyrC-O bond (1300)

(*18, 23*). Therefore, it seemed that two types of the Tyr-Fe bond in Mb mutant coexists, stronger and weaker have the $\nu_{Fe\text{-}O}/\nu_{TyrC\text{-}O}$ bands to 598cm^{-1} /1289 cm^{-1} and 578cm^{-1} /1301 cm^{-1}, respectively, but the intensity ratio of $\nu_{Fe\text{-}O}$ band to $\nu_{TyrC\text{-}O}$ band are opposite. The ratio of these two coordination types is slightly different between the two mutants.

Tyr Coordinated to the Heme Iron as a Tyrosinate The UVRR spectrum of H64Y and F46V/H64Y excited at 244 nm gave a new band derived from tyrosinate (Tyr^-). Figure 2-3.6 shows the spectra of Tyr^-(A), Tyr (B) and Trp (C) residues. The spectrum D is calculated the sum of Tyr, Tyr^- and Trp with the population ratio in the mutant swMb (3:1:2), while the spectrum E is the result of the same calculation of D but the intensity of Tyr^- is reduced to one-fifth. These calculated spectra indicated that the intensity ratio of Tyr^- / Tyr of Y8a band in protein matrix is smaller than that for an aliquot solution. Since the intensity of Tyr band is smaller in polar environment than nonpolar, if the intensity of Tyr^- has opposite effect by the environment, it may be possible to explain this difference of intensity between in the free solution and protein. Another possibility is that only one of the two Tyr^- coordination forms gives a Tyr^- UVRR band, but another doesn't.

The Y9a band has also large change and it shows the different frequencies among H64Y, F46V/H64Y and free Tyr solutions at 1163, 1168 and 1175 cm^{-1}. The Y9a band of Tyr is very sensitive and low frequency shift due to the environment change toward polar condition (*24*). Since the environment of Tyr residue F46V/H64Y seems to be more polar than that of H64Y due to the open conformation, it is considered that the frequency of Y9a band of Tyr^- is also shifted by the influence from the environment.

The Environment of Heme Pocket Although the distal His in H46V mutants can have

two directions that is, open form and closed form, the distal Tyr in CO bounded form of F46V/H64Y has no difference from that of H64Y and shows only a single ν_{C-O} band at 488 cm^{-1}. It indicated that there is little difference in electrostatic field between the two mutants. Since the coordination bond of the Tyr to heme iron is not so strong, distal Tyr residue of F46V/H64Y has higher pK than that of H64Y mutant. Therefore, the main reason for the difference of pK may come from the structural hindrance of Tyr residue. The structure with the neutral Tyr in the distal pocket would be distorted by the heme and surrounding residues, and it is disadvantage in free energy than tyrosinate coordinated structure. If Phe46 was substituted to the Val, this distortion will be decreased and neutral Tyr residue can easily shift toward the Phe 46 position. Since it is easily imaged that the ferrous ligand bound state of H64Y, such as CO-bound or O_2-bound form, has large perturbations among the distal Tyr, the ligand and other residues, it destabilizes the ligation of the ligand and accelerates toward the oxidized form that has smallest repulsion. In many cases of proteins, especially around the catalytic site, there are found many residues with abnormal pK and it is essential of enzyme catalysis. The coordination of Tyr in these Mb mutants or Hb M are lethal for its function, whereas that of catalase is essential for its catalysis (*25*). In this study we showed that not only the electronic field but also the steric hindrance control the pK values of the residues.

Reference

1. Springer, B. A., Sligar, S. G., Olson, J. S., and Phillips, G. N., Jr. (1994) *Chem. Rev. 94*, 699-714.
2. Hargrove, M. S., Singleton, E. W., Quillin, M. L., Ortiz, L. A., Phillips, G. N., Jr., Olson, J. S., and Mathews, A. J. (1994) *J. Biol. Chem. 269*, 4207-14.

3. Nagai, M., and Yoneyama, Y. (1983) *J. Biol. Chem. 258*, 14379-84.

4. Maurus, R., Bogumil, R., Luo, Y., Tang, H. L., Smith, M., Mauk, A. G., and Brayer, G. D. (1994) *J. Biol. Chem. 269*, 12606-10.

5. Egeberg, K. D., Springer, B. A., Martinis, S. A., Sligar, S. G., Morikis, D., and Champion, P. M. (1990) *Biochemistry 29*, 9783-91.

6. Pin, S., Alpert, B., Cortes, R., Ascone, I., Chiu, M. L., and Sligar, S. G. (1994) *Biochemistry 33*, 11618-23.

7. Tang, H. L., Chance, B., Mauk, A. G., Powers, L. S., Reddy, K. S., and Smith, M. (1994) *Biochim. Biophys. Acta. 1206*, 90-6.

8. Bolognesi, M., Cannillo, E., Ascenzi, P., Giacometti, G. M., Merli, A., and Brunori, M. (1982) *J. Mol. Biol. 158*, 305-15.

9. Lionetti, C., Guanziroli, M. G., Frigerio, F., Ascenzi, P., and Bolognesi, M. (1991) *J. Mol. Biol. 217*, 409-12.

10. Ringe, D., Petsko, G. A., Kerr, D. E., and Ortiz de Montellano, P. R. (1984) *Biochemistry 23*, 2-4.

11. Johnson, K. A., Olson, J. S., and Phillips, G. N., Jr. (1989) *J. Mol. Biol. 207*, 459-63.

12. Yang, F., and Phillips, G. N., Jr. (1996) *J. Mol. Biol .256*, 762-74.

13. Lai, H. H., Li, T., Lyons, D. S., Phillips, G. N., Jr., Olson, J. S., and Gibson, Q. H. (1995) *Proteins 22*, 322-39.

14. Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., and Olson, J. S. (1989) *J. Biol. Chem. 264*, 3057-60.

15. Harada, I., and Takeuchi, H. (1986) *Advances in Spectroscopy: Spectroscopy of biological System*, 113-175.

16. Hu, S., Smith, K. M., and Spiro, T. G. (1996) *J. Am. Chem. Soc. 118*, 12638-12646.

17. Que, L. J. (1988) in *Biological Application of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 491-521.

18. Nagai, M., Yoneyama, Y., and Kitagawa, T. (1989) *Biochemistry 28*, 2418-22.

19. Spiro, T. G., and Li, X. (1988) in *Biological Application of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 1-37, John Willy and Sons, New York.

20. Phillips, G. N., Teodoro, M. L., Li, T. S., Smith, B., and Olson, J. S. (1999) *J. of Phys. Chem. B 103*, 8817-8829.

21. Ramsden, J., and Spiro, T. G. (1989) *Biochemistry 28*, 3125-8.

22. Sage, J. T., Morikis, D., and Champion, P. M. (1991) *Biochemistry 30*, 1227-37.

23. Nagai, K., Kagimoto, T., Hayashi, A., Taketa, F., and Kitagawa, T. (1983) *Biochemistry 22*, 1305-1311.

24. Haruta, N., Aki, M., Ozaki, S., Watanabe, Y., and Kitagawa, T. (2001) 40, 6956-6963

25. Sharma, K. D., Andersson, L. A., Loehr, T. M., Terner, J., and Goff, H. M. (1989) *J. Biol. Chem. 264*, 12772-9.

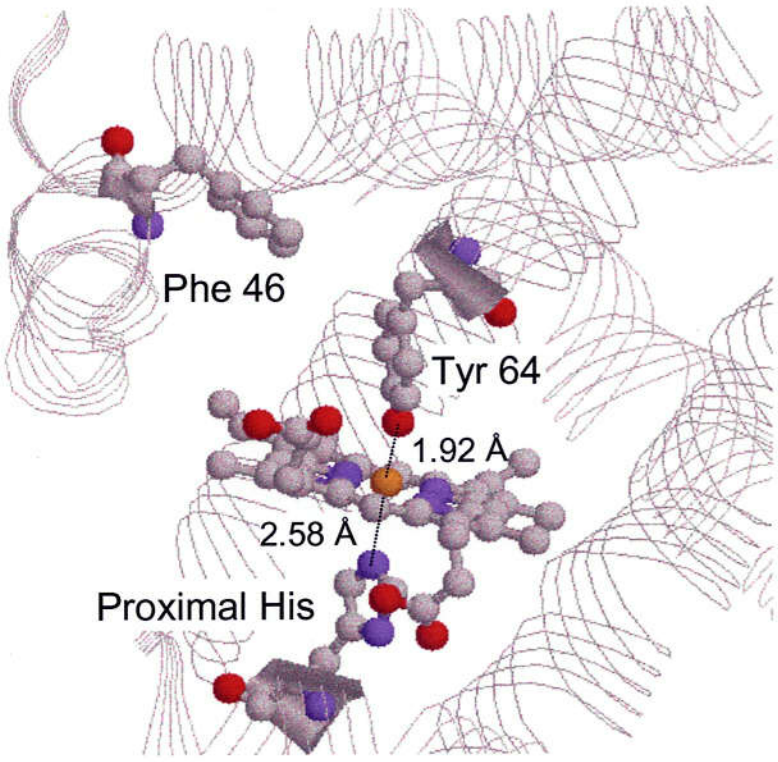


Figure 2-3.1. The structure of H64Y mutant at met state around the heme pocket *(4)*. Tyr64 is coordinated to the heme iron from opposite side of proximal His.

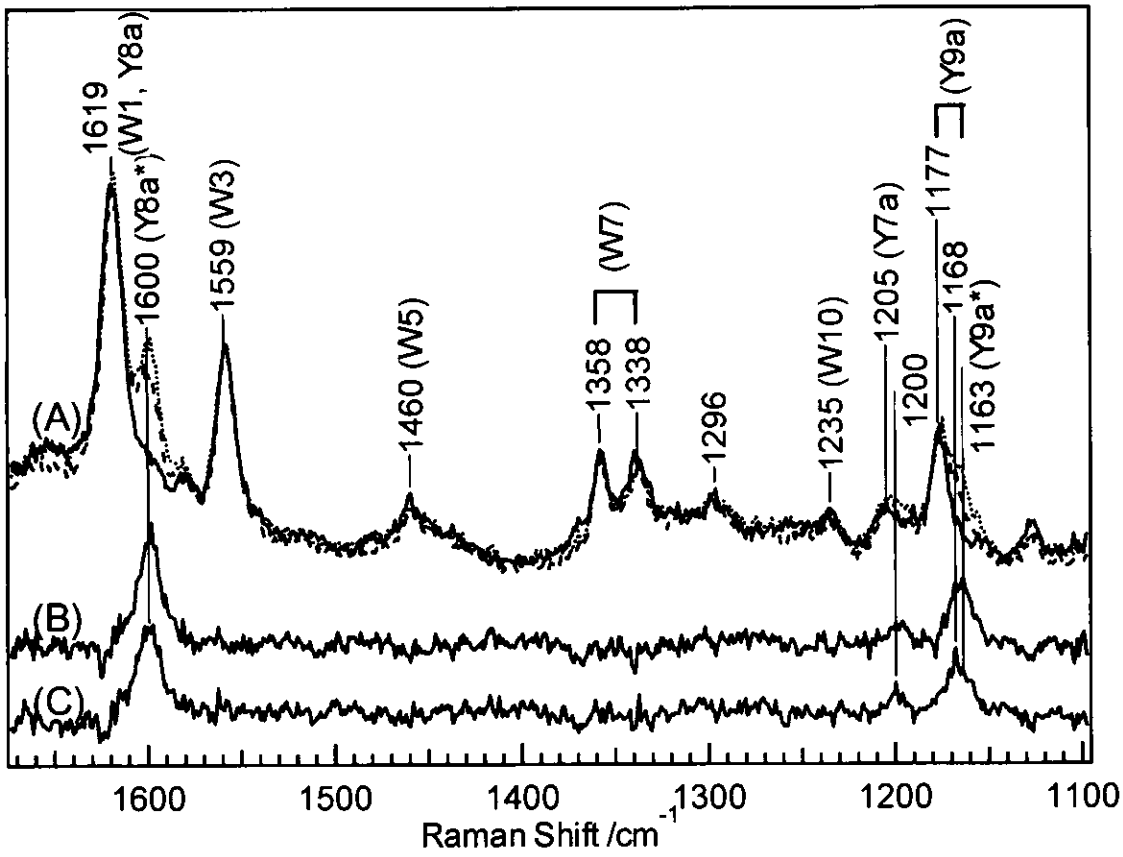


Figure 2-3.2. The 244-nm excited UVRR spectra of native (solid), H64Y (dotted) and F46F/H64Y (broken) mutants of sperm whale metMb at pH 8.7 (A). The spectra B and C are differences, subtracted by the spectrum of native Mb from H64Yand F46V/H64Y mutants, respectively. The total exposure time is 30 min with changing the sample each 5 min at 20 °C. The bands derived from Tyr- are labeled with *.

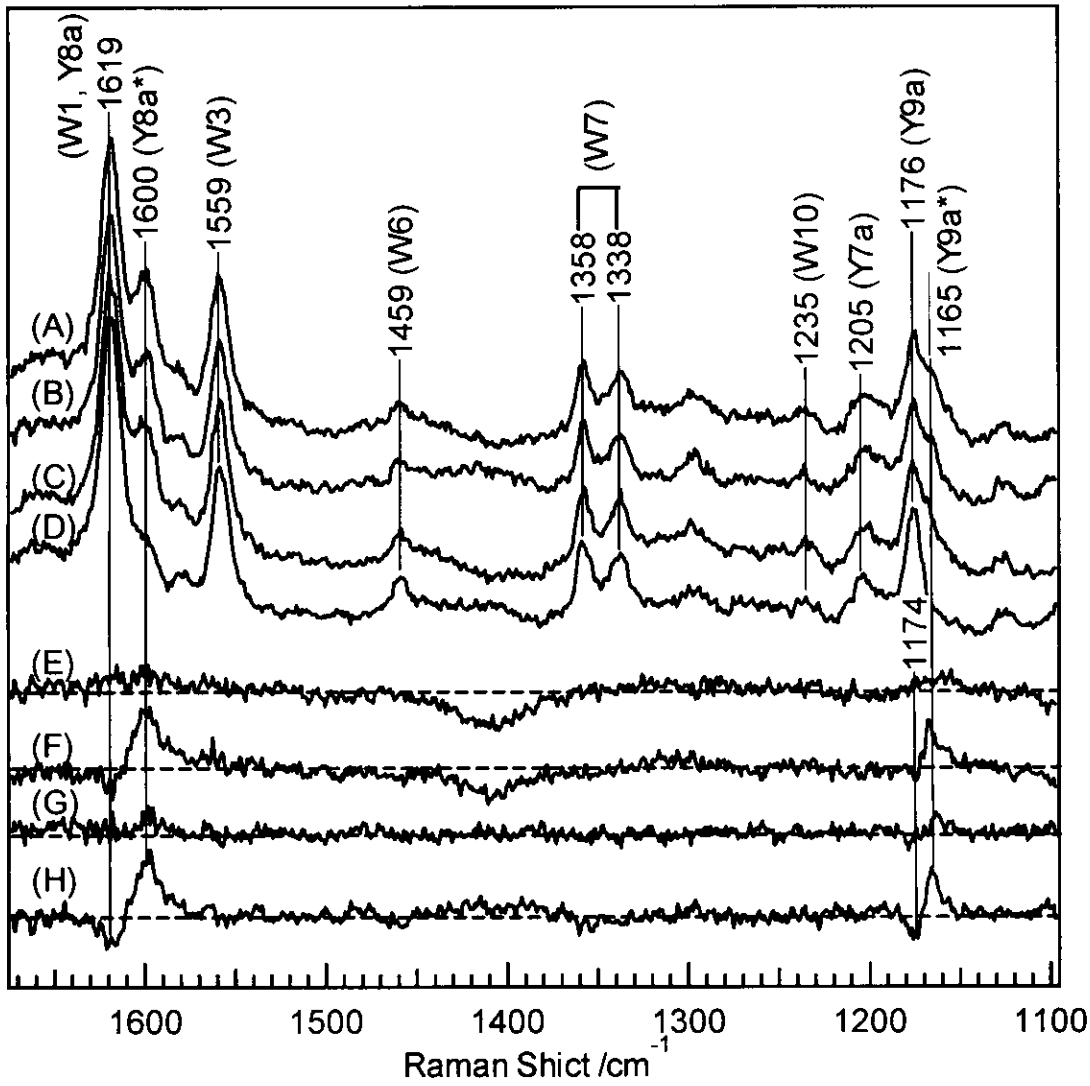


Figure 2-3.3. The 244-nm excited UVRR spectra of H64Y (A,B) and F46F/H64Y (C,D) mutants of sperm whale metMb at pH 8.5 (A, C) and pH 5.7 (B, D). The difference spectra are between pH 8.5 and pH 5.7 of H64Y (E=A-B) and F46F/H64Y(F=C-D), and between the mutants at pH 8.7 (G=A-C) and pH 5.7 (H=B-D).

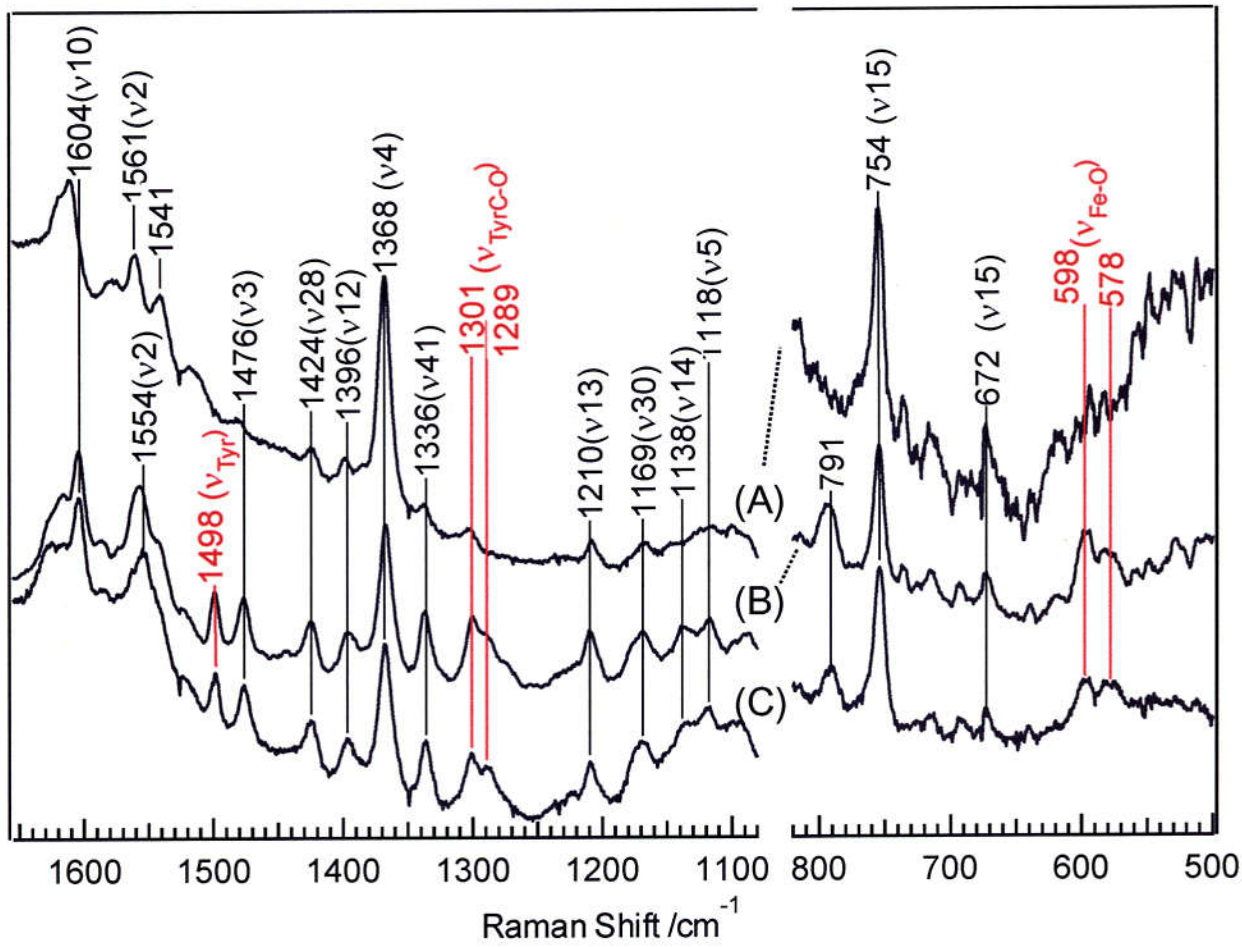


Figure 2-3.4. Visible resonance Raman spectra of native (A), H64Y (B) and F46F/H64Y (C) mutants of sperm whale metMb excited at 488 nm. The bands labeled with red color are relation bands of coordinated Tyr residue.

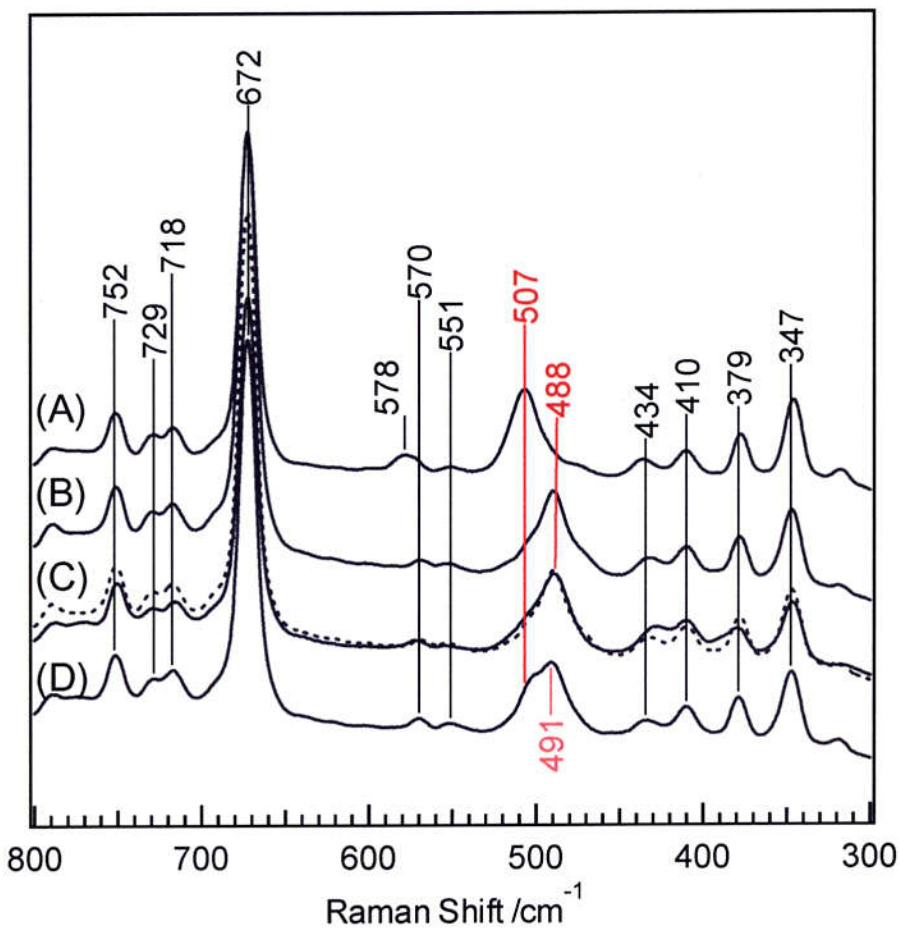


Figure 2-3.5. Visible resonance Raman spectra of native (A), H64Y (B), F46F/H64Y (C) and F46V (D) mutants of sperm whale Mb in the CO bounded forms excited at 415.4 nm. The broken spectrum overlapped with the spectrum C is copy of spectrum B for comparison of the spectra of H64Y with that of F46V/H64Y. The bands labeled with red color are FeC-O stretching mode.

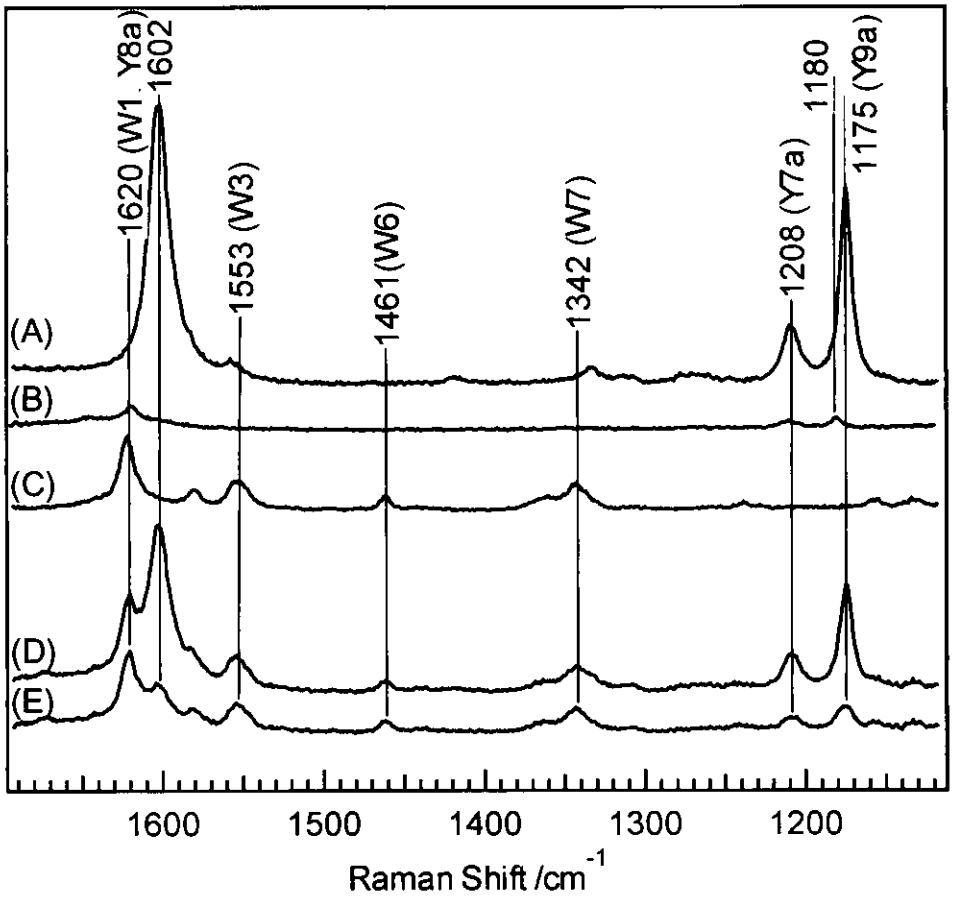


Figure 2-3.6. The UVRR spectra of free Tyr and Trp solution; A) Tyr at pH 13, B) Tyr at pH 7, C) Trp at pH 7. The intensity is normalized by the band of ClO_4^- at 933 cm^{-1} (not shown). The spectrum D is calculated the sum of Tyr, Tyr^- and Trp with the population ratio in mutant sperm whale Mb (3:1:2), while the spectrum E is the result from the same calculation as D but the intensity of Tyr^- is reduced to one-fifth.

Chapter 3. Studies on Protein Folding of Apomyoglobin Probed by Ultraviolet Resonance Raman Spectroscopy

3-1. Overview: Mechanism of Protein Folding and Experimental Technique

3-1.1 Old and new view of protein folding

3-1.2 Phenomenological models of protein folding

a. Framework model and diffusion-collision model

b. Hydrophobic collapse model

c. Nucleation-condensation model

3-1.3 Two-state folding proteins

a. ϕ value analysis

b. The folding rate and contact order

3-1.4 Aspects of stationary states

a. Denatured state

b. The equilibrium intermediate

c. Native state

3-1.4 Experiment techniques for kinetic folding reaction

a. The spectroscopic method of folding studies

b. The trigger of folding reaction

Protein is working as a molecular machine regulating the chemical reaction by changing their complicated tertiary structures. Although the flexibility and complexity of the protein structure are essential for their function, it is made up by the polypeptide chain consisting of the order of only 20 kinds of amino acids residues according to the DNA sequence. Therefore, a large gap between the tertiary structure of the protein and the primary sequence of DNA or RNA is attractive theme of science. The goal of protein folding study is to understand the rules of the formation of tertiary structures and to predict tertiary structures from primary sequence. When it is accomplished, a new protein will be constructed artificiality for desired function.

3-1.1 Old and new view of protein folding

Protein is synthesized as a primary polypeptide chain according to a gene sequence and constructs a tertiary structure. Since Anfinsen showed that polypeptide chain folds automatically to a unique tertiary structure *in vitro* (*1*), studies of folding have been trying to understand the principle of protein architecture. Levinthal pointed out that the random searching of all possible structures requires an immense amount of time, due to astronomic number of conformations that would need to be searched (*2*). However, protein actually folds very fast; most of small protein folds completely in a few minutes. Therefore, it is assumed that protein folding has a directed process. To represent the folding process, two views have been proposed, so-called old view and new view.

The old view has been defined as describing folding as parallel to chemical reactions involving distinct intermediates and transition states along a single, well

defined folding 'pathway'. On the other hand, a new view describes protein folding in terms of statistical ensembles of states and is focused on the general features of folding on a complex multidimensional potential energy function, so-called 'folding funnel' (*3-6*). Figure 3-1.1 shows the schematic diagram of folding funnel. The width of funnel corresponds to configurational entropy, and depth depicts a free energy function and the fraction of native contacts. Thus, the funnel implies that protein folding is a process, which decreases in energy and concomitantly lose entropy by increasing the native contact. This aspect of the new view has been supported by the result of calculations with simplified lattice model that assumes the protein as a string and beads positioned as sites on a lattice. In these models, the motion of lattice is simulated with a dynamic Monte Carlo algorithm (*3-6*).

3-1.2 Phenomenological models of protein folding

There are many experimental and theoretical studies about protein folding which use various small proteins. Three classes of folding reactions may be distinguished on the basis of the appearance of kinetic and equilibrium intermediates, except kinetically complex folding reactions such as the proline isomerization, heme ligand binding or disulfide bond formation. The first class of folding reaction, at least to a good approximation, is a two-state model. Small proteins, which include CspB (Cold-shock protein from *Bacillus subtilis*) (*7*), SH3 domain (*8, 9*), IgG binding domain of Protein L (*10*) and CI2 (chymotrypsin inhibitor 2) (*11*), show neither equilibrium nor kinetic folding intermediate during fording and unfolding reactions. The second class shows two-state folding reactions in the equilibrium folding/refolding system, but complex

kinetics of refolding that indicated the presence of kinetic intermediate. Barnase (*12*), ubiquitin (*13*), RNase H (*14*) and lysozyme (*15*) are listed as representative proteins for this class. The third class, which includes apomyoglobin, cytochrome *c* and α-lactalbumin, shows equilibrium intermediates that are relatively stable than unfolded state under most of conditions and stable than native state under specified conditions. These intermediates appeared in the kinetic refolding reaction (*16*). It is not clear whether the two-state folding and multikinetic folding with one or more intermediates have quite different mechanism or not. However, one may explain that the difference is not fundamental, but rather simply reflect differences in stability of intermediate along the pathway (*17*).

The simple folding models are proposed to explain the mechanism of the first step of folding process. Framework model and hydrophobic model allow the presence of the intermediate, but nucleation-condensation model is based on the two-state folding (Figure 3-1.2).

a. Framework model and diffusion-collision model

The framework model explains the appearance of the folding intermediate. The formation of isolated secondary structure is very fast and automatic and depends on the sequence. Framework model emphasizes the formation of native local secondary structure independently that of tertiary structure (*18*). The structure of the intermediate has some secondary structures but not tertiary contacts that resembles the molten globule state. The local structured elements would diffuse until they collide, successfully adhere, and coalesce to give a tertiary structure. Diffusion-collision model focuses on this diffusion process (*19*).

b. Hydrophobic collapse model

The hydrophobic interaction is one of the important factors for the formation and stabilization of the structure. The central idea of the hydrophobic collapse model is the tendency of hydrophobic residues to be excluded from the aqueous environment (*20*). The property of the intermediate in this model is, precisely, the hydrophobic collapse with nonspecific interaction. The advantage of the collapse is considered to reduce the conformational space for searching the correct and nearly activationless pathway. Therefore, the collapse is kinetically favored, and its formation is significantly more rapid under conditions favoring the native state.

c. Nucleation-condensation model

The nucleation-condensation model is proposed for explanation of two-state folding reaction by Fersht (*21*). In this mechanism, the rate-limiting step of the reaction is the formation of a single nucleus with diffusion at the determined place, and development of nucleus, that is condensation process, occurs at the transition state. If a protein has more than two nucleus positions, this protein might be folded with an intermediate according to the framework model. Such a single domain protein with single folding nucleus can be treated as the folding module, or foldon (*22*).

3-1.3 Two-state folding proteins

In many years of folding studies, the characterization of successive intermediates along the pathway has been thought to be the best way of establishing the folding pathway. In this decade, however, an increasing number of small proteins that fold

rapidly by simple two-state kinetics, without the accumulation of an intermediate and with only one kinetically important transition state have been found. CI2 (chymotrypsin inhibitor 2) (*11*), SH3 domain (*8, 9*), acylphosphatase (*23*), CspB (Cold-shock protein from *Bacillus subtilis*)(*7*) and IgG binding domain of Protein L (*10*) are repoted as two-state folding proteins.

a. ϕ value analysis

Since the structure of transition states of two-state folding protein cannot be directly analyzed, kinetic and equilibrium measurements on suitable mutants have been carried out at microscopic levels. Fersht have proposed the method that analyses the ϕ value, which was determined by suitable mutants (*24*). ϕ value is defined by $\Delta\Delta G^*$ /$\Delta\Delta G$, where $\Delta\Delta G^*$ is a change in free energy of activation and $\Delta\Delta G$ is that of equilibrium (Figure 3-1.3). Since the structure of transition state can be indicated by ϕ value, Fersht and coworkers implied that the structure of transition state resembles native structure around the structural nucleus in the analysis of CI protein (*25, 26*).

The ϕ value analysis requires the hard work because of preparations of a large amount of mutants. Recently, ϕ value analysis of some proteins that fold with two-state are reported and studies of these different proteins suggested a similar result that the native topology plays an important role and the transition state has native like topology (*8, 9, 23*).

b. The folding rate and contact order

From finding and analyzing the two-state folding proteins, the theory of protein folding is developed and a new concept has arisen. Baker proposed the easily

calculated parameter, the 'relative contact order', which is simply an average distance in sequence between interacting residues, normalized by the chain length. Surprisingly, the contact order exhibits a good correlation with the folding rate of small proteins that proceed in the two-state folding (*27-30*). It indicated that native topology determines mainly the folding rate and mechanism rather than its inter-atomic interactions. However, there is little knowledge about how a given protein finds the native topology, or what determines the native topology.

3-1.4 Aspects of stationary states

The studies about the character and property of protein in the denatured state give us abundant information to understand the folding reaction as well as the structure of native one. Since the stability of protein structure is calculated by the difference of free energy between the native state and the denatured (unfolded) state, the thermodynamics properties of denatured state are also important for the argument of protein structure and stability.

a. Denatured state

Protein denaturation occurrs by changing physical and chemical condition apart from its physiological environment. Protein can be denatured by cooled and heated temperatures, low or high pH, pressure and high concentration of denaturants such as urea and guanidinium chloride (GuHCl). The structure and feature of denatured state depend on these physical and chemical solution conditions. The denaturation by physical change of environment alone, such as temperature and pH, is relatively

moderate, and partially the some of native structures are retained. In contrast, the denatured state with strong denaturants seems almost unfolded and the structure of this state resembles random coil.

The denatured state of protein is not described as a single state, but is treated as the ensemble of many states that have very similar energies and can convert easily one another, called as microstates. However, this ensemble of conformational heterogeneity behaves as a single, more-or-less continuous distribution of microstates. Therfore, we can treat the two-state approximation between a native state and a denatured state (*31, 32*). It is difficult to characterize the structure of denatured state, since it has highly flexible and dramatic character and heterogeneity. The gross property of denatured state such as hydrodynamic volume and radius of gyration may be indicative of the average of compactness.

b. The equilibrium intermediate

There are many cases in which the denaturation curve of protein can not be used as the approximation of two-state model. It shows the non-cooperative transitions among the results of different spectroscopic measurements such as near and far UV CD. These results indicated the presence of intermediate, which may be related to the so-called *molten globule* state (*16*). This is the one of the denatured state that has a partially folded structure but is not unfolded completely. Although the molten globule state has heterogeneous conformations, it has some common properties which was clarified by many spectroscopic methods; far UV CD for detection of the secondary structure, near UV CD and fluorescence of intrinsic chromophore for tertiary structure, the binding of ANS for hydrophobic core and X-ray small scattering for the

compactness of the protein molecule. Consequently, the common structural characteristics of molten globule state are as follows: 1) a pronounced amount of secondary structure is present; 2) most of the specific tertiary structure produced by tight packing of side chain is absent; 3) protein molecule has compactness which radius of gyration is 10 to 35 % larger than that of native state; 4) there is loosely packed hydrophobic core that increases the hydrophobic surface accessible to solvent.

Since the characterization of the structure of protein in the molten globule state can not be clear due to its flexibility and heterogenety, it is difficult to determine whether the equilibrium intermediate is same as kinetic folding intermediate or not. However, the recent development of the detection and trigger of reaction technique such as pulsed-hydrogen NMR gave us the great information about the structure of intermediate. Some proteins that have been numerously studied in detail, such as apomyoglobin (apoMb) (*33*) and cytochrome *c* (cyt *c*) (*34*), indicated that the equilibrium intermediate present during the acid pH unfolding or induced by addition of anion has very similar structure to those of sequential kinetic intermediate.

c. Native state

The structure of native state can be determined by the X-ray crystallographic analysis or NMR. The structure of native state is rigid and compact with tight packing. In general, the hydrophobic residues are located inside of protein, whereas the polar residues tend to be exposed to solvent.

Although the structure of native state has a minimum Gibbs free energy, the difference of free energy between the native state and unfolded state is only 5 to 10 kcal/mol (20-40 kJ/mol). This surprisingly small value has the same order of

magnitude as a typical hydrogen bond. It is indicated that the tertiary structure of native state results from the balance of many oppositely working large forces; hydrophobic interaction, electrostatic interaction and repulsion, and hydrogen bonds. We must consider not only the enthalpy and entropy of the protein itself but also the interaction with water solvent. The "iceberg" structure of the water around the hydrophobic residues has significance in consideration of the dramatic temperature dependence of the stability and explanation of cold denaturation (*24*).

3-1.4 Experiment techniques for kinetic folding reaction

Many kinds of techniques have been applied to the protein folding reactions and yielded the variety aspects of the protein structures. Triggers of folding reaction are also important because they determine often the time resolution, and the techniques of the trigger have been progressed with the development of detection techniques.

a. The spectroscopic method of folding studies

The structural change of protein during the folding reaction is large. Many spectroscopic methods have been applied to detect the dynamics of protein folding from various aspects (Table 3-1.1).

The most common method for the detection of the secondary structure such as α helical and β sheet structure is near UV circular dichroism (CD) spectroscopy (*35*) and it is routinely used in laboratories in combination with stopped-flow method. Although the secondary structure is well estimated by many studies, it is the total contents of overall protein structure and it could not determine the position of the folded

and unfolded structure. The hydrogen exchange NMR is a strong tool for determination of the folded position, and recent work that developed the labeling system succeeded in detection the transient α helix formation in a few milliseconds (*33, 36*). Secondary structure is also detected by IR and ultraviolet resonance Raman (UVRR) spectra with below 220 nm excitation and these vibrational spectroscopy are often combined with the laser injection trigger (pump-probe method) that performs high time resolution nearly nanoseconds (*37-39*).

Tertiary structure is difficult to detect directly without NMR and X-ray crystallographic techniques, which can not use for the rapid reactions and requires the critical conditions. Therefore, in general, we use the intrinsic aromatic residue, such as Trp and Tyr, as a sensitive probe for conformational changes. Fluorescence of Trp (*40-42*) is most popular, and fluorescence polarization anisotropy (*40, 41*), far UV CD and UVRR spectroscopy (*43*) excited from 230 to 260 nm can also detect Trp residues.

The external probes for fluorescence also give the various information of the structure. ANS is known to bind to hydrophobic surface and its fluorescence signal dramatically increases on binding, while it has a low signal when free in solution. Therefore, ANS binding method is used to characterize the structural features of folding intermediate (*42*). Other informations obtainable from fluorescence studies are the distance between the energy transfer donor-acceptor pairs positioned at different sites and the accessibility of fluorophores to solute quenchers (*40*).

Small-angle X-ray scattering can measure the radius of gyration and it indicates the compactness of overall protein structure (*44, 45*). Since most of these experiments give us only one or two information from each technique, we must apply many techniques of the detection to the same reaction condition of same protein. Indeed, it

is hard. The time-resolutions of these spectroscopic methods are almost dependent on the trigger method described below except NMR.

b. The trigger of folding reaction

The trigger of the folding reaction often determines the time limitation rather than the detection. Table 3-1.2. shows main methods of the trigger applied to the folding reaction. The most popular and routine method is the stopped-flow method. Since this method uses basically the mixing of two solutions, this can be applied to most of the reactions that are changed by the physical and chemical condition. The amount of sample that required for measurement is small, only 50 to 200 μl from each. However, the dead time is as long as milliseconds in which the major change in the folding process of many proteins has already finished .

The continuous-flow method is recently developed the dead time from milliseconds of stopped flow method to few tens microseconds. Roder and coworkers designed the continuous-flow apparatus with a small ball of platinum at the mixing point of two solutions and achieved the dead time of about 50 μs (*46*). They could observe the formation of initial collapse of cyt *c* without the burst phase by the change of Trp fluorescence with time constant of ~50 μs. Takahashi and coworkers constructed another type of mixing apparatus that efficiently mixed two solutions by the collision from opposite side through the T shaped flow path (*47*) (Figure 3-3.1). They combined this apparatus with many spectroscopic detection methods, such as resonance Raman spectroscopy (*47*), CD (*48*), IR and X-ray small angle scattering. Although this method has greatly improved the time resolution, it may be difficult to reach the dead time shorter than microsecond in which nucleation process of secondary structure

occurs. Furthermore, the significant disadvantage of this apparatus is that the large amount of sample is required. Thus, only proteins, such as cyt *c* and myoglobin, which are relatively easier to purify or purchase are used for this system.

The trigger for the fast dynamics is to use a pulsed laser in the pump-probe method. This is the conventional method for analysis of the fast dynamics of chemical reaction and these time resolution reached femtoseconds. For application of folding reactions, the appropriate time resolution is prospected to be nanoseconds to microseconds. The pump-probe method is divided into two classes. One is the change of the solvent condition by pump light, such as heating of water by tuning the wavelength of pump light to the absorption of the water (*38, 39, 49, 50*). Since this temperature-jump can perform only one direction, that is heating, many experiments are to observe heat denaturation from native state or refolding from cold-denaturation. Recently, the pH-jump was also performed with the caged compound that releases the proton by the laser injection (*51*).

Another class is the change of the property of the solute itself. Gray and coworkers found that the stability of protein is lower in the oxidized form than reduced form. Therefore, the cyt *c* was modified with Ru(2,2-bipyridine)$_3^{2+}$(Ru^{2+}) that supplies an electron by the photo injection and triggered the unfolding reaction by change the redox state. Since this method is very specific to metalloproteins whose stability might be changed by the redox, only the study of cyt *c* was reported (*52*).

Reference

1. Anfinsen, C. B. (1973) *Science 181*, 223-30.

2. Levinthal, C. (1968) *J. Chim. Phys. 65*, 44-45.

3. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) *Proteins 21*, 167-95.

4. Dill, K. A., and Chan, H. S. (1997) *Nat Struct Biol 4*, 10-9.

5. Dobson, C. M., and Karplus, M. (1999) *Curr Opin Struct Biol 9*, 92-101.

6. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) *Trends Biochem Sci 25*, 331-9.

7. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat Struct Biol 2*, 663-73.

8. Martinez, J. C., and Serrano, L. (1999) *Nat Struct Biol 6*, 1010-6.

9. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. I., and Baker, D. (1999) *Nat Struct Biol 6*, 1016-1024.

10. Scalley, M. L., Yi, Q., Gu, H., McCormack, A., Yates, J. R., 3rd, and Baker, D. (1997) *Biochemistry 36*, 3373-82.

11. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10428-35.

12. Dalby, P. A., Oliveberg, M., and Fersht, A. R. (1998) *J Mol Biol 276*, 625-46.

13. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) *Nat Struct Biol 3*, 193-205.

14. Raschke, T. M., and Marqusee, S. (1997) *Nat Struct Biol 4*, 298-304.

15. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature 358*, 302-7.

16. Arai, M., and Kuwajima, K. (2000) *Adv Protein Chem 53*, 209-82.

17. Miranker, A. D., and Dobson, C. M. (1996) *Curr Opin Struct Biol 6*, 31-42.

18. Kim, P. S., and Baldwin, R. L. (1990) *Annu Rev Biochem 59*, 631-60.

19. Karplus, M., and Weaver, D. L. (1994) *Protein Sci 3*, 650-68.

20. Dill, K. A. (1985) *Biochemistry 24*, 1501-9.

21. Fersht, A. R. (1997) *Curr Opin Struct Biol 7*, 3-9.

22. Panchenko, A. R., Luthey-Schulten, Z., and Wolynes, P. G. (1996) *Proc Natl Acad Sci U S A 93*, 2008-13.

23. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M., and Dobson, C. M. (1999) *Nat Struct Biol 6*, 1005-9.

24. Fersht, A. (1998) *Structure and Mechanism in Protein Science*, W. H. Freeman and Company, New York.

25. Fersht, A. R., Itzhaki, L. S., elMasry, N. F., Matthews, J. M., and Otzen, D. E. (1994) *Proc Natl Acad Sci U S A 91*, 10426-9.

26. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) *J Mol Biol 254*, 260-88.

27. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) *J Mol Biol 277*, 985-94.

28. Goldenberg, D. P. (1999) *Nat Struct Biol 6*, 987-90.

29. Munoz, V., and Eaton, W. A. (1999) *Proc Natl Acad Sci U S A 96*, 11311-6.

30. Baker, D. (2000) *Nature 405*, 39-42.

31. Smith, L. J., Fiebig, K. M., Schwalbe, H., and Dobson, C. M. (1996) *Fold Des 1*, R95-106.

32. Shortle, D. (1996) *Faseb J 10*, 27-34.

33. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892-6.

34. Colon, W., and Roder, H. (1996) *Nat Struct Biol 3*, 1019-25.

35. Chen, Y. H., and Yang, J. T. (1971) *Biochem Biophys Res Commun 44*, 1285-91.

36. Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y., and Roder, H. (2001) *Nat Struct Biol 8*, 151-5.

37. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1996) *Biochemistry 35*, 691-7.

38. Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1997) *proc.Natl.Acad.Sci.USA 94*, 3709-3713.

39. Lednev, I. K., Karnoup, A. S., Sparrow, M. C., and Asher, S. A. (1999) *J. Am. Chem. Soc. 121*, 4076-4077.

40. Eftink, M. R. (1991) in *Methoed of Biological Analysis* (Clarence, Ed.) pp 127-205, John Willy and Sons, Inc.

41. Jiskoot, W., Hlady, V., Naleway, J. J., and Herron, J. N. (1995) *Pharm Biotechnol 7*, 1-63.

42. Eftink, M. R., and Shastry, M. C. (1997) *Methods Enzymol 278*, 258-86.

43. Harada, I., and Takeuchi, H. (1986) *Advances in Spectroscopy: Spectroscopy of biological System*, 113-175.

44. Nishii, I., Kataoka, M., Tokunaga, F., and Goto, Y. (1994) *Biochemistry 33*, 4903-9.

45. Eliezer, D., Jennings, P. A., Wright, P. E., Doniach, S., Hodgson, K. O., and Tsuruta, H. (1995) *Science 270*, 487-8.

46. Shastry, M. C., and Roder, H. (1998) *Nat Struct Biol 5*, 385-92.

47. Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) *Nat Struct Biol 4*, 44-50.

48. Akiyama, S., Takahashi, S., Ishimori, K., and Morishima, I. (2000) *Nat Struct Biol 7*, 514-20.

49. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *nature structural biology 3*, 923-926.

50. Yamamoto, K., Mizutani, Y., and Kitagawa, T. (2000) *Biophysical Journal 79*, 485-495.

51. Abbruzzetti, S., Crema, E., Masino, L., Vecli, A., Viappiani, C., Small, J. R.,

Libertini, L. J., and Small, E. W. (2000) *Biophys J 78*, 405-15.

52. Pascher, T., Chesick, J. P., Winkler, J. R., and Gray, H. B. (1996) *Science 271*, 1558-60.

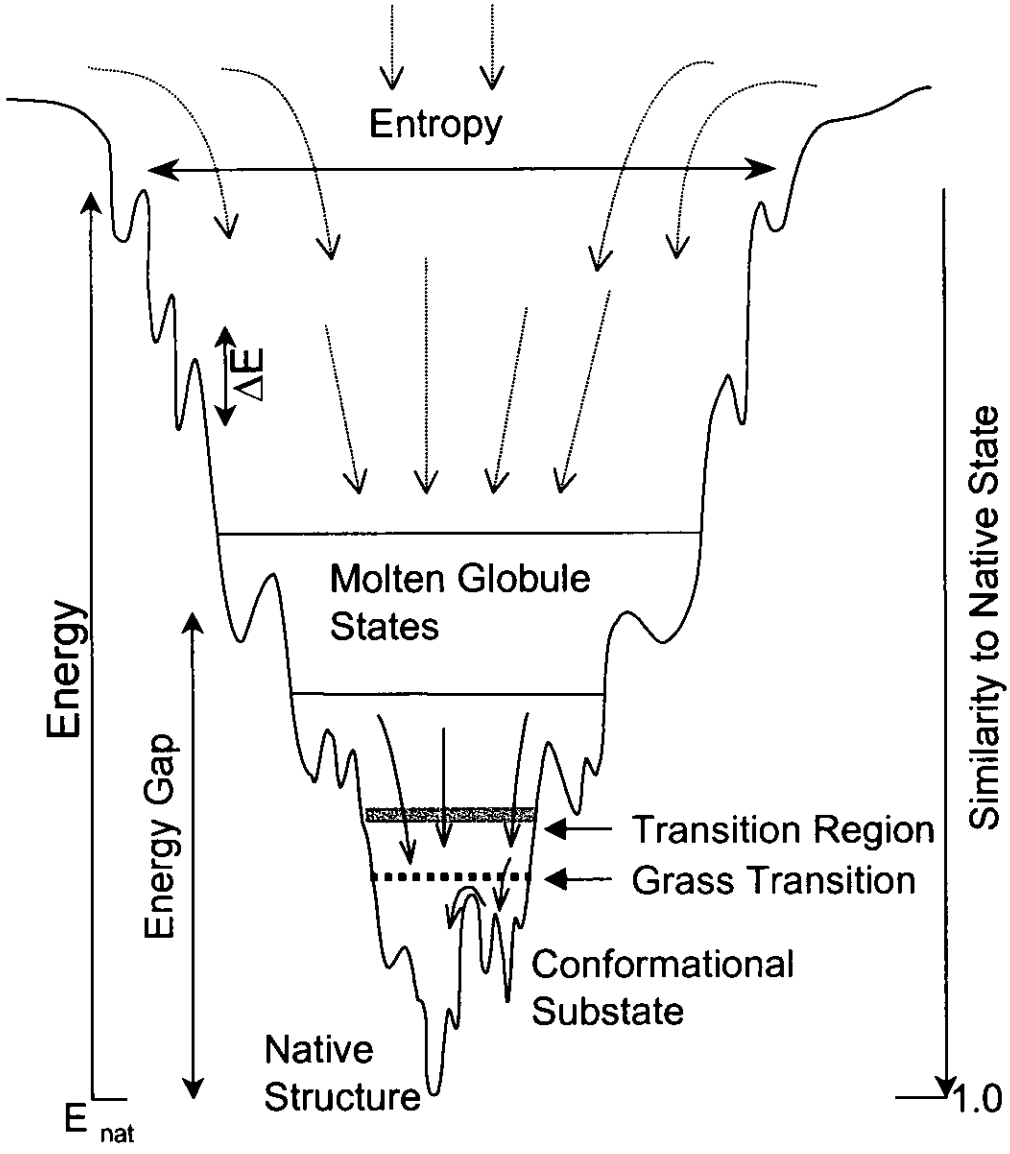


Figure 3-1.1. A schematic representation of the energy landscape for a minimally frustrated heteropolymer as required in protein folding. The width of the funnel corresponds to configurational entropy and depth corresponds to a free energy function and the fraction of native contacts.

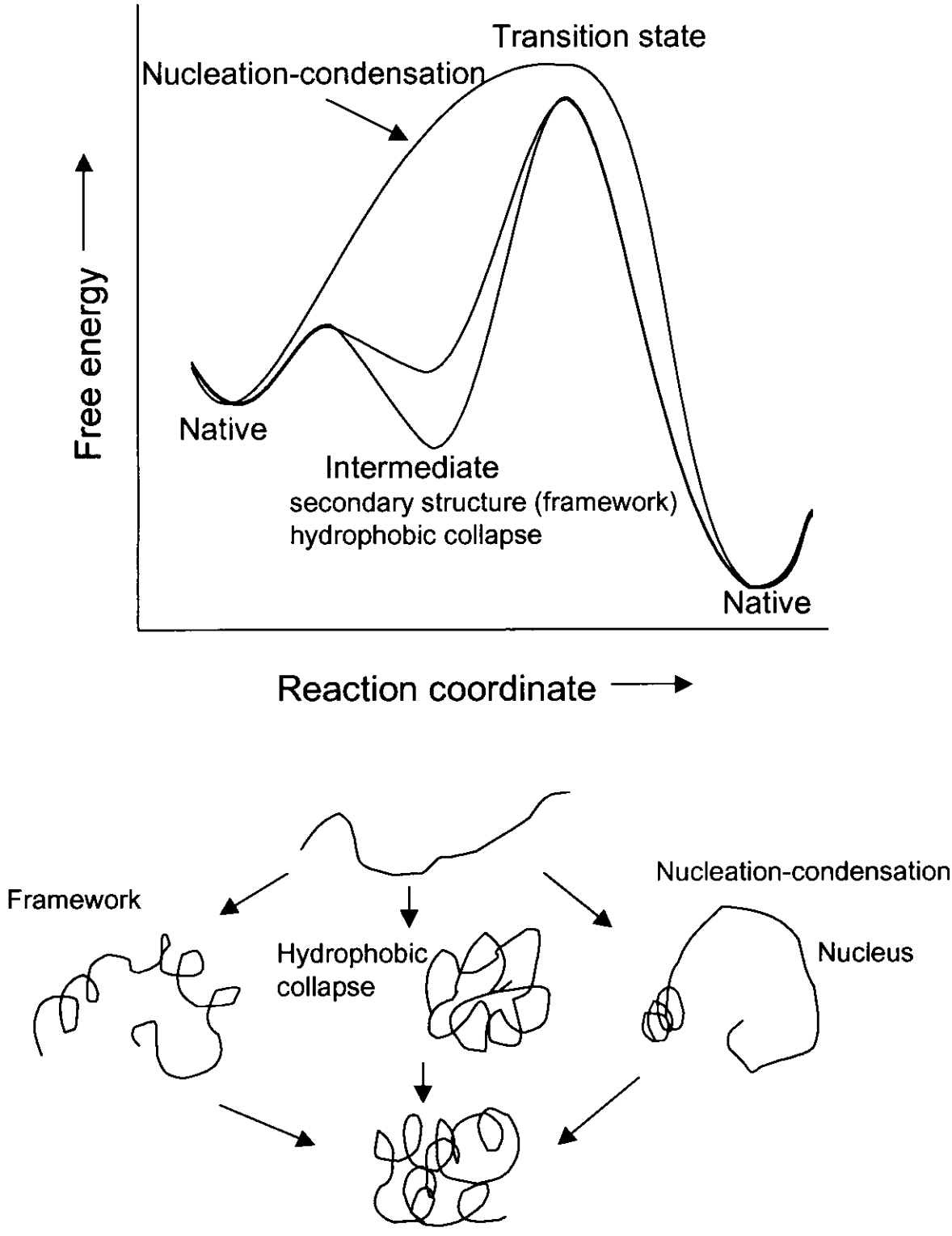


Figure 3-1. 2. Simplified energy diagrams for two-state folding model and multikinetic folding model.

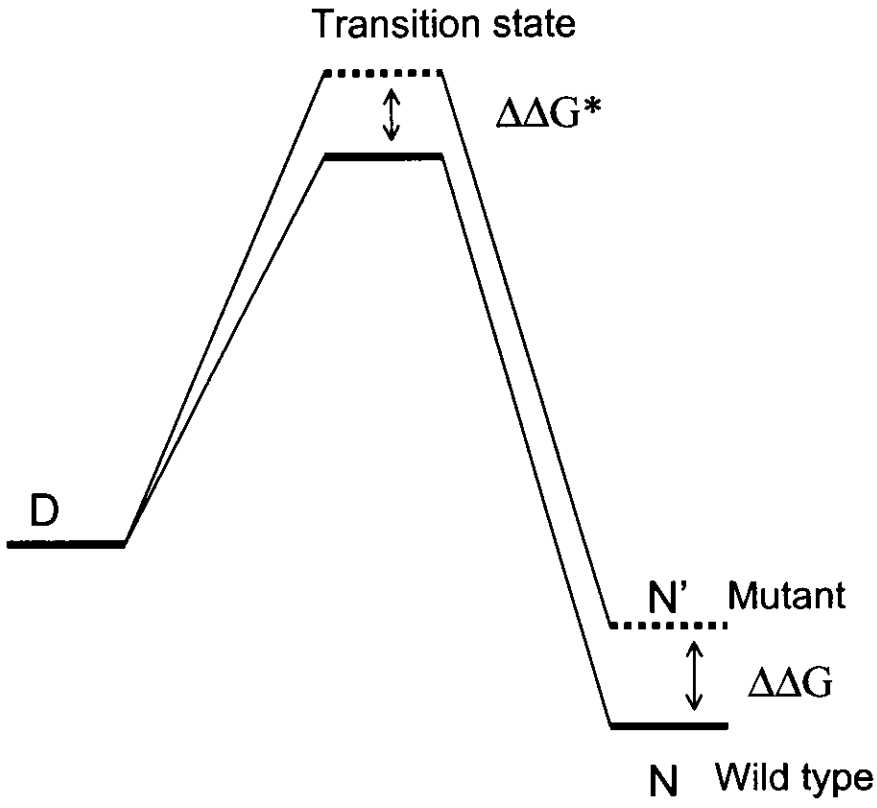


Figure 3-1.3. Free energy diagrams for the two-state folding protein. The $\Delta\Delta G^*$ value was calculated by the equation $\Delta\Delta G^* = RT \ln(k_{wt}/k_{mt})$, where k_{wt} and k_{mt} are folding rate of wild type and mutant, respectively. The ϕ value is determined by $\phi = \Delta\Delta G^* / \Delta\Delta G$.

Table 3-1.1. *Detection techniques for studying of protein folding*

Technique	*Dead time (Trigger)*
Secondary structure (backbone)	
Near UV CD	a, b
NMR (pulsed hydrogen exchange)	Milliseconds
IR	a, b,c
UVRR (below 220 nm)	b,c
Tertiary structure (side chain)	
Fluorescence of intrinsic Trp	a,b, c
Far UV CD	a, b
Real-time NMR	Seconds
UVRR (280-220 nm)	b, (c)
Others	
Small-angle X-ray scattering	a,b
Force spectroscopy using the AFM	Seconds
ESI MS (pulsed hydrogen exchange)	Milliseconds
Fluorescence (ANS binding, FRET)	Milliseconds

a, stopped flow; b,ultrafast mixing; c, laser induction

Table 3-1.2. *Method of triggering for refolding /folding reaction*

Technique	*Dead time*
a. Stopped-flow	milliseconds
b. Ultrafast mixing	50 microseconds
c. Laser induction (pump-probe)	
Temperature jump	Nanoseconds
Photo-induced (redox, photolysis)	<Nanoseconds

3-2. Equilibrium Acid-Unfolding Process of Horse and Sperm Whale Apomyoglobins

(Contents of this chapter has been submitted)

Abstracts

The 244-nm excited UVRR spectroscopy was applied to the folding process of apomyoglobin. One characteristic folding intermediate appeared around pH 4 upon the acid unfolding process of horse apoMb by analyzing of two UVRR bands (W3 and W7) of Trp side chain. In the pH 4 intermediate, Trp14 was buried inside of protein and its hydrophobicity was as same as that of the native state, but the structure of the side chain was different from the native one. On the other hand, Trp7 was already exposed to the solvent without the secondary steric interaction. The pH dependencies of apoMb were different between horse and sperm whale apomyoglobin.

Abbreviations: UVRR, ultraviolet resonance Raman; Mb, myoglobin; apoMb and holoMb, apomyoglobin and holomyoglobin; hMb and swMb, horse Mb and sperm whale Mb; IR, Infrared; NMR, nuclear magnetic resonance; CD, circular dichroism; NOE, nuclear overhauser effects; CW, continuous wave; N, native; I, intermediate; U, unfolded.

Introduction

Structure of folding and unfolding intermediates give us the strong information for understanding the folding mechanism and expression of protein function. The folding intermediate of apomyoglobin (apoMb) is one of the ideal systems to know the interaction and stabilization of the secondary structure and tertiary structure because it forms a stable intermediate at pH 4. Moreover, this equilibrium intermediate is indicated to be same as the kinetically folding intermediate (*1*).

Myoglobin (Mb) has 8 helices, labeled A, B-H, and apoMb has 7 helices that removing the hem melts F helix, which connected to the hem iron with His93 (Figure 3-2.1) (*2, 3*). The stable pH 4 folding intermediate has been characterized by many measurement; NMR(*1, 4, 5*), CD(*6-8*), fluorescence (*9*), small angle X-ray scattering (*10, 11*), and IR techniques(*12*). This intermediate contains A, B, G, and H helices (*1, 4, 5*) with the native-like contacts, called the AGH hydrophobic core (*13, 14*). Since the Trp14 was located on the A-helix and near the G and H helix, it has been using the good probe for the conformational change to detect by the fluorescence (Figure 3-2.1) (*15-17*). While the advantage of the fluorescence spectroscopy such as sensitive and simple without long accumulation of the spectra due to strong intensity suites for the kinetic experiments, it is difficult to interpret the structure of the intermediate in detail because the quenching of fluorescence is complicated with many factors.

We used the ultraviolet resonance Raman spectroscopy (UVRR) to determine the structure of intermediates. Previously we have characterized the UVRR spectra of Trp7 and Trp14 separately using W14F and W7F mutants of Mb, respectively, and pointed out that a structural change upon ligand binding to heme is communicated to the N-terminal of globin through the proximal His (Section 2-2) (*18*). According to this experiment, we also

focused on the two Trp bands, W3 and W7, to chase the folding process with expectation to separate the conformational information from environmental one on Trp residues.

Materials and Methods

Preparation of Samples Apomyoglobin (apoMb) was prepared from horse myoglobin (Sigma) and sperm whale myoglobin (Biozyme) by 2-methylethylketone (MEK, 2-Butanone) extraction of the heme (*19*). The ice-cold holo protein solution was adjusted to pH 2 with HCl, and mixed with equal volume of ice-cold MEK on the ice. Separation takes place into two layers, ketonic supernatant including heme and a lower aliquot layer containing the protein. Remove the supernatant and add equal volume of ice-cold MEK, again. This manipulation was repeated forth or fifth times until lower aliquot layer was colorless and transparent. The solution removed the heme was dialyzed against ice-cold distilled water to remove the dissolved ketone and HCl during over night with change the distilled water fifth times.

UVRR Measurements The protein concentration of sample solution for stationary experiments was adjusted to 60 μM and prepared the each pH by adding the adequate amount of 1M HCl. To avoid the effect of salt, such as aggregation and stabilization of intermediate, we don't use some buffer and checked the pH before and after the UVRR measurement. The 300 μl equilibrium sample was introduced to the spinning cell and the UVRR spectra were measured to reach the total exposure time to less than 120 min at each pH with changing the fresh sample every 10 minute.

The UVRR spectra were excited by intracavity frequency doubling 244 nm CW laser

(Coherent, Innova FreD). The measurement of UVRR spectra was essentially the same as described in section 2-2. The excitation laser power was 200 μW for avoiding the photoreaction and the measurement was performed at 20 °C.

Results

Figure 3-2.2 shows the 244-nm excited UVRR spectra of horse holoMb and apoMb, and difference spectrum subtracting the spectrum of apoMb from holoMb, which normalized the Raman intensity by the band of SO_4^{2-} (980 cm^{-1}). The Trp bands, W1(1620 cm^{-1}), W3 (1559 cm^{-1}), W7 (1358/1338 cm^{-1}), W10 (1238 cm^{-1}), W13 (1129 cm^{-1}), W16 (1011 cm^{-1}), W17 (877 cm^{-1}) and W18 (752 cm^{-1}) assigned are based on Harada and co-workers (*20*), has not large difference between the apoMb and holoMb than Tyr bands. Tyr bands, which assigned Y8a (1620 cm^{-1}), Y7a (1206 cm^{-1}), Y9a (1178 cm^{-1}) and Y1 (852 cm^{-1}) bands indicated the environmental and structural change of Tyr residues were occurred dramatically. The intensity decrease of Y9a and Y1 bands in the apoMb indicated that the Tyr residues exposed to the hydrophilic environment compared to the holoMb. The largest difference between the apoMb and holoMb was that the F helix was melted, where the proximal His (H93) included and connected with the hem iron in the holoMb (*4, 21*). Moreover, the NOEs showed the apoMb structure, which maintain A, B, G and H helices has same packing as horse holoMb (*22*). Tyr146 has hydrogen bond with the backbone of the F helix and Tyr103 interacts with C helix. Therefore, it seems reasonable to suppose that the melting of the F helix cause particularly the perturbation and expose of Tyr146.

The W7 band of Trp residue was the N_1C_8 stretching mode of indole ring and split by

the Fermi resonance with two combination bands (*23*). The I_{1358}/I_{1338} ratio (*R*) of W7 doublet band was sensitive to the environment whether the polar and non-polar. The *R*-value of holoMb showed 1.2 and that of apoMb showed 1.05. The calculation with *R*-value of Trp residue, which completed exposure to solvent (0.48), resulted that apoMb was 20 % less hydrophobicity than holoMb. Since two Trp residues located at the A helix and interact with EF loop and H helix (Trp7) and E, H and G helices (Trp14), the melting F helix induced the fluctuation of E-F loop and the environment of Trp7 residue changed to hydrophilic than holoMb.

The UVRR spectra for the equilibrium acid-unfolding process of horse Mb at 20 °C are shown in Figure 3-2.3, where the trace at bottom denotes the difference spectrum between pH 5.5 and 1.9, each of which is referred as the native and the acid-unfolded (low salt) states, respectively. The difference spectrum means that the bands at 1559, 1358 and 1172 cm^{-1} decrease in intensity upon acid unfolding. Intensity of the 1172 cm^{-1} band of Tyr is already reduced in the spectrum at pH 5.5 compared with that of holoMb at pH 5.6 due to removal of heme, and further decreases at pH 1.9. The intensity reduction means that the environment of Tyr residue becomes more hydrophilic, and thus that Tyr side chain is more exposed to solvent in apoMb than in holoMb at pH 5.5, and so is at pH 1.9 than pH 5.5 (*18*).

The W7 doublet and W3 bands of Trp residue and Y9a band of Tyr exhibited dramatically changes during acid unfolding. The W3 band (C_2-C_3 stretch) (*18*) is shifted to lower frequencies and shows broadening at lower pH. This feature is expanded in Figure 3-2.4, where solid lines indicate the observed spectra. It is known that the W3 frequencies of Trp have a correlation with the torsion angles ($\chi^{2,1}$) of the C_α-C_β-C_3-C_2 linkage (*24*). The W3 frequency at pH 1.9 is identical with that of tryptophan in an

aqueous solution, indicating that the Trp residues of acid-unfolded Mb at pH 1.9 receive no steric hindrance from other parts of protein.

The frequency shift and intensity decrease of W3 band with pH could not be satisfactorily interpreted by weighted sum of the native and unfolded spectra. Deviation of the observed spectra from the calculated ones, which were simulated under the assumption of the presence of two components, was conspicuous around pH 4, suggesting the appearance of an intermediate around pH 4. Therefore, we tried to simulate the observed spectra with three components, that is, the native, an intermediate and the unfolded species, according to the following idea (Figure 3-2.5). We know from the previous mutation study (*18*) that Trp7 and Trp14 in the native state give the W3 bands at 1558 and 1561 cm^{-1}, respectively. Accordingly, the relative intensity and frequencies observed there were used as those of native species in the simulation. Both Trp7 and Trp14 in the unfolded state are considered to have the same band shape and intensity as those of aqueous tryptophan solution. It seems likely that Trp7 is already exposed to the protein surface in the intermediate species, because the studies on pH dependence of far-UV CD spectra indicated that A helix of horse apoMb at pH 4 was partially unfolded, while A helix of sperm whale apoMb at pH 4 was not unfolded (*4*, *8*). Accordingly, it would be reasonable to assume that Trp7 in the intermediate species has W3 band at 1553 cm^{-1} with the same intensity as that of tryptophan in an aqueous solution. Thus, only a spectrum of Trp14 in the intermediate species is unknown. The spectral fitting calculations suggested its presence at 1557 cm^{-1} as depicted by dotted lines in Figure 3-2.4. In this way the W3 spectra of Trp7 and Trp14 in the native, intermediate, and unfolded species were determined, and their sum weighted with populations were calculated as shown by dashed lines in Figure 3-2.4, which well reproduce the observed spectra.

Figure 3-2.6 shows the pH dependence of the population of the three species obtained from deconvolution of W3 band depicted in Figure 3-2.5. The native species decreases upon lowering of pH from 5.5 to 4.0, and conversely, the intermediate species increases with the midpoint pH of transition at pH 4.5. Since this intermediate species was a dominant component around pH 4, this will be called "equilibrium pH 4 intermediate" hereafter. Below pH 4, there are practically two species in equilibrium, that is, the pH 4 intermediate and the unfolded species, with the midpoint pH of transition at ~3.5.

The pH dependence of the *R*-value is illustrated in Figure 3-2.7, where the values for horse (open circles) and sperm whale Mbs (closed circles) are contained and the corresponding hydrophobicity, under the assumption of 100% hydrophobicity for horse holoMb and 0% for acid-denatured state, is graduated at the right axis. The *R*-value of the native state was 1.05 for apoMb, meaning that the Trp residues were buried in the hydrophobic environment of the protein inside. Since the W7 bands at various pH values could be well fitted by the weighted superposition of native and unfolded spectra, there would be no other effects that determine the W7 doublet intensity except for hydrophobicity. The plot in Figure 3-2.7 indicates the presence of two transitions with the midpoint pH at 4.5 and 3.5, which are identical with the transition pH values of the native to the equilibrium pH 4 intermediate states and of the equilibrium pH 4 intermediate to the unfolded states determined with the W3 band (Figure 3-2.6).

The *R*-value of the equilibrium pH 4 intermediate indicates that average hydrophobicity in the environments of Trp is 55 % of that in the native holoMb. In the previous study of the W7F and W14F mutants of Mb, we concluded that the environments of Trp14 was 1.3 times more hydrophobic than that of Trp7 (*18*). Supposed that the

parameter representing buried extent of Trp7 is X, the corresponding parameter of Trp14 would be 1.3X and their average would give the hydrophobicity of Trp in native holoMb. When Trp7 is exposed completely (X=0) at pH 4 while Trp14 remained buried, the average value of the buried parameter would be 1.3X/2 and it would correspond to 57% hydrophobicity of native holoMb (= (1.3X/2)/(2.3X/2)), in agreement with the observed value. Accordingly, it is highly likely that the change from the native to the pH 4 intermediate is accompanied by complete exposure of Trp7 under retention of Trp14 in the hydrophobic environments. It suggests that not only the AGH hydrophobic core but also the E helix are retained in the hydrophobic condensation, although the NMR-hydrogen exchange experiments indicated the amide proton of E helix was not protected from access of solvent water (*4*).

Figure 3-2.8 shows the UVRR spectra of sperm whale apoMb at various pH. The difference spectrum between the pH 5.3 and pH 1.9 is very similar to the difference spectrum of horse apoMb. However, the spectral change around pH 4 is different between two apoMbs, as described about W7 band. We also tried to deconvolute the W3 band of sperm whale apoMb as horse apoMb. However, it is difficult to apply the result of horse apoMb to sperm whale one because the W3 band change of sperm whale apoMb seems to be more complicated. The complication of W3 band change indicated the presence of some intermediates in the process of sperm whale apoMb as Jamin and Baldwin pointed out the presence of two intermediate from the result of fluorescence experiment (*9*).

Discussion

Horse ApoMb and Serm Whale ApoMb The schematic diagram of the structure of horse apoMb at pH 4 was drawn in Figure 3-2.9 on the basis of this study. Since accumulated data on folding of apoMb have been derived from both horse and sperm whale Mbs and they are partly controversial, it is important to compare the two Mbs. The present Raman spectra showed that the Trp UVRR bands were similar between horse and sperm whale Mbs in the native holoMb (*18*), but slightly dissimilar in apoMb; more hydrophobic for sperm whale apoMb than horse apoMb at pH 5.6. Furthermore, pH dependencies of the W7 doublet of apoMb were different as shown in Figure 3-2.7.

The pH dependencies of Raman frequency shift and intensity change of the W3 band were also different between the horse and sperm whale apoMbs. This difference seems to reflect differences in the formation of A helix. The α helix content of whole protein at pH 4.2 was lower than the sum of the contents of A, G and H helices for horse apoMb, whereas it was higher for sperm whale apoMb (*8*). Since the equilibrium intermediate species at pH 4 has the AGH helix core, in which the A helix has contact with stable G and H helices at the C-terminal side of A helix, it is reasonable to assume that the N-terminal side of A helix is more melted in horse Mb than in sperm whale Mb and accordingly, Trp residues might be influenced by it. In the analysis of fluorescence spectra of apoMb, it is assumed that the fluorescence of Trp14 is more quenched by Met131 of H helix when they are closer but that of Trp7 is not at all. However, discussion on Trp 7 requires some caution due to possible differences between horse and sperm whale Mbs. The present results suggest that Trp7 of horse apoMb in the equilibrium pH 4 intermediate is exposed to the solvent in the form of random coil, while Trp7 of sperm whale apoMb is on the α helix and in contact with H helix.

Structural characteristics of the equilibrium pH 4 intermediate Although the pursuit of packing process of the secondary structures is not easy with conventional methods, vibrational spectroscopy of Trp can serve as a direct probe of the structural changes around Trp side chains. The frequency shift of the W3 band between the native and the equilibrium pH 4 intermediate indicates that some steric interaction is present between A and E helices in the native state but is absent in the pH 4 intermediate due to melting of E helix. In holoMb, the presence of heme would help the globin overcome the potential barrier of this final stage in folding toward construction of the active site. In this regard the folding process of Mb seems to consist of two independent steps, formation of the "frame" structure and construction of the "functional" site.

The equilibrium unfolding intermediate of apoMb at pH 4 has been studied extensively using various spectroscopic techniques (*4-8, 11, 25, 26*). Many of them concluded that the equilibrium pH 4 intermediate was the same as the kinetic intermediate appearing in the millisecond time regime for neutral solutions (*1, 9, 10, 16, 27*). In this intermediate, the G, H and A helices are fully or partially folded (*4, 5*), and the native-like AGH hydrophobic core is formed (*14*). On the basis of these studies, the structure of native apoMb seems to be categorized into two structural portions, regions I and II. Region I includes A, G and H helices which remain intact in the equilibrium pH 4 intermediate. Accordingly, region I must be formed in the earlier stage of folding and provide a rigid core to construct region II. Region II includes B, C, D, E and F helices, which are almost frayed in the equilibrium pH 4 intermediate. The tertiary structure formation of region II is the rate-limiting step of folding process. Most residues which control the ligand binding and recognition including distal and proximal His are contained

in this region. Two introns which devide the gene of Mb into three parts, interpose B and G helices (*28*). Regions I and II may correspond to the "frame" and "functional" parts, respectively.

References

1. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892-6.

2. Eliezer, D., and Wright, P. E. (1996) *J Mol Biol 263*, 531-8.

3. Lecomte, J. T., Kao, Y. H., and Cocco, M. J. (1996) *Proteins 25*, 267-85.

4. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) *Science 249*, 1544-8.

5. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) *Nat Struct Biol 5*, 148-55.

6. Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., and Kutyshenko, V. P. (1988) *J Mol Biol 202*, 127-38.

7. Goto, Y., and Fink, A. L. (1990) *J Mol Biol 214*, 803-5.

8. Sabelko, J., Ervin, J., and Gruebele, M. (1998) *J. Rhys. Chem. B 102*, 1806-1819.

9. Jamin, M., and Baldwin, R. L. (1998) *J Mol Biol 276*, 491-504.

10. Eliezer, D., Jennings, P. A., Wright, P. E., Doniach, S., Hodgson, K. O., and Tsuruta, H. (1995) *Science 270*, 487-8.

11. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, Y. (1995) *J Mol Biol 249*, 215-28.

12. Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1997) *Biochemistry 36*, 15006-12.

13. Kay, M. S., and Baldwin, R. L. (1996) *Nat Struct Biol 3*, 439-45.

14. Kay, M. S., Ramos, C. H., and Baldwin, R. L. (1999) *Proc Natl Acad Sci U S A 96*,

2007-12.

15. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *nature structural biology 3*, 923-926.

16. Jamin, M., u, and Baldwin, R. L. (1999) *J Mol Biol 292*, 731-740.

17. Tcherkasskaya, O., Ptitsyn, O. B., and Knutson, J. R. (2000) *Biochemistry 39*, 1879-89.

18. Haruta, N., Aki, M., Ozaki, S., Watanabe, Y., and Kitagawa, T. (2001) *Biocemistry 6956-6963.*

19. Teale, F. W. (1959) *Biochim. Biophys. Acta 35*, 543.

20. Harada, I., and Takeuchi, H. (1986) *Advances in Spectroscopy: Spectroscopy of biological System*, 113-175.

21. Cocco, M. J., Kao, Y. H., Phillips, A. T., and Lecomte, J. T. (1992) *Biochemistry 31*, 6481-91.

22. Lecomte, J. T., Sukits, S. F., Bhattacharya, S., and Falzone, C. J. (1999) *Protein Sci 8*, 1484-91.

23. Harada, I., Miura, T., and Takeuchi, H. (1986) *Spectrochimica Acta 42a*, 307-312.

24. Miura, T., Takeuchi, H., and Harada, I. (1989) *J. Raman Spectrosc. 20*, 667-671.

25. Griko, Y. V., and Privalov, P. L. (1994) *J Mol Biol 235*, 1318-25.

26. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) *Proc Natl Acad Sci U S A 92*, 5446-50.

27. Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, H. J., and Wright, P. E. (1999) *Protein Sci 8*, 45-9.

28. Go, M. (1985) *Adv Biophys 19*, 91-131.

29. Evans, S. V., and Brayer, G. D. (1990) *J Mol Biol 213*, 885-97.

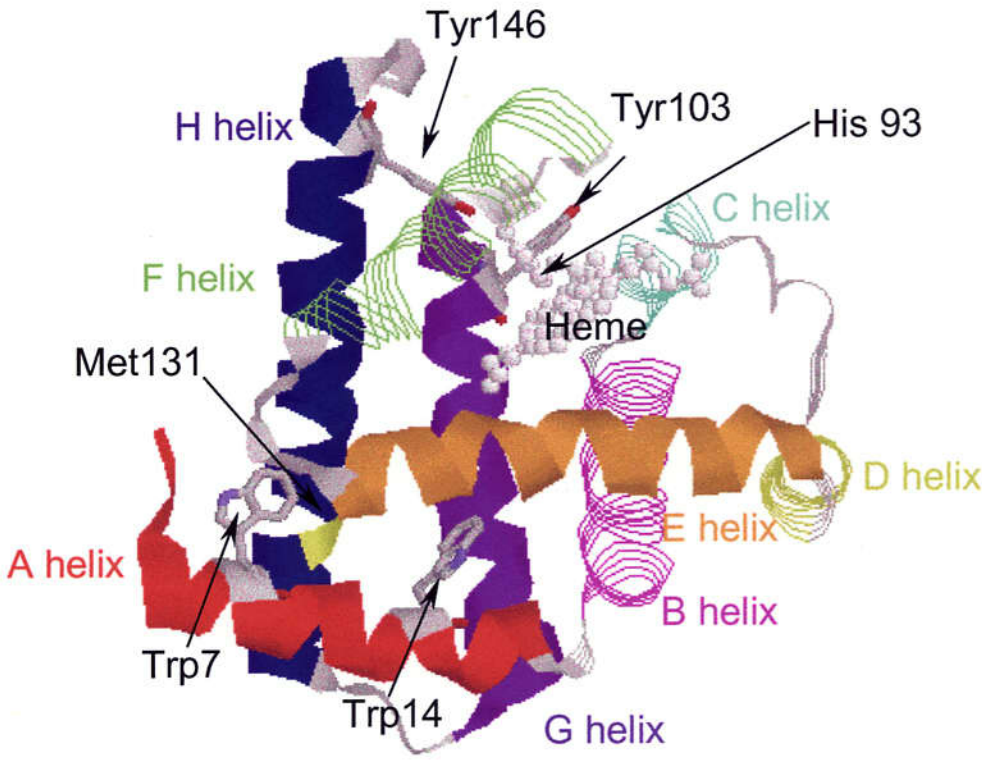


Figure 3-2.1. The ribbon diagram of horse myoglobin, with each helices labeled. Trp and Tyr residue are showed by ball and sticks. The X-ray crystallographic data was reported by Evans et al.*(29)*.

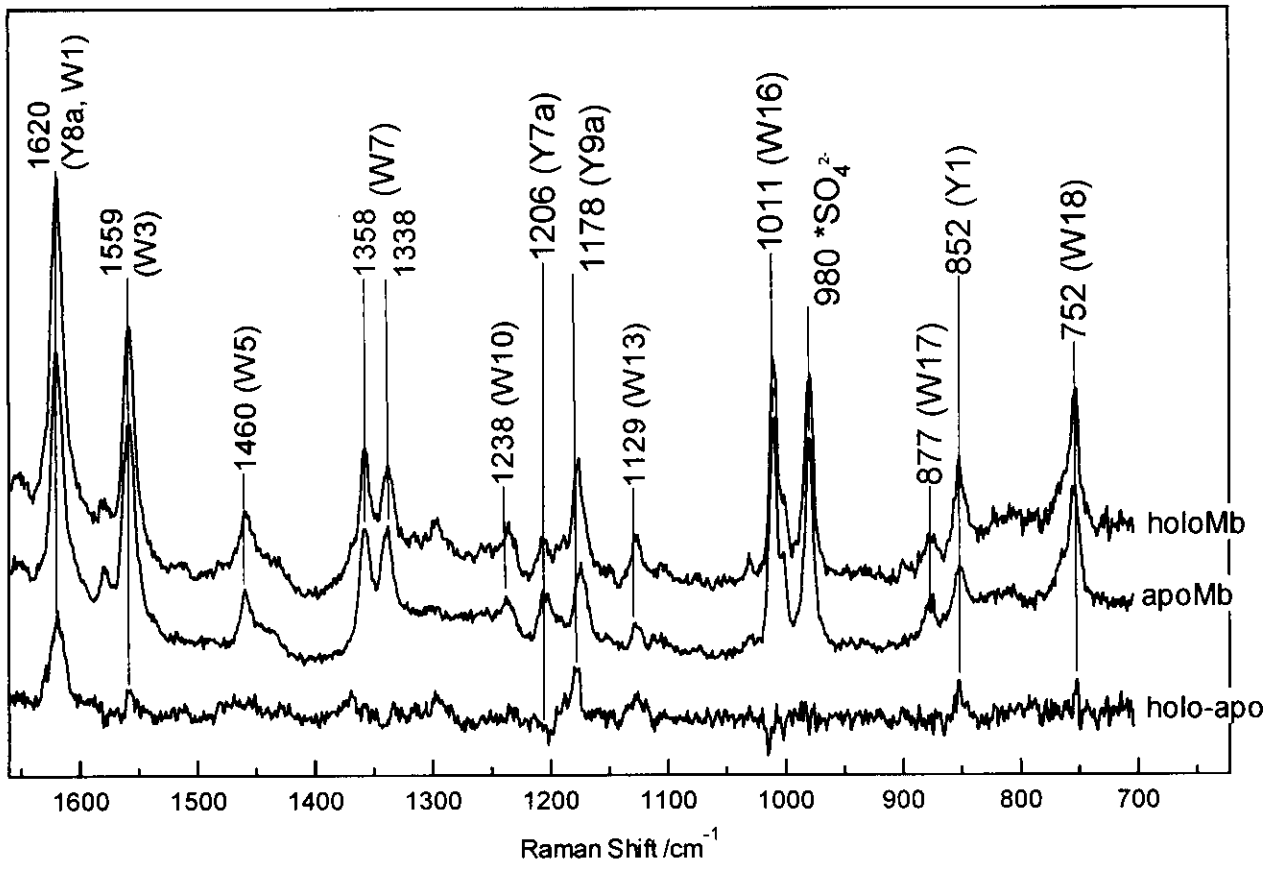


Figure 3-2.2. The 244-nm excited UVRR spectra of horse holoMb and apoMb at 20 °C. The concentration of holoMb and apoMb was prepared to 400 μM at pH 7.5 with 50 mM phosphate buffer. The difference spectrum was calculated to subtract the spectrum of apoMb from that of holoMb by normalizing with SO_4^{2-} band at 980 cm^{-1}.

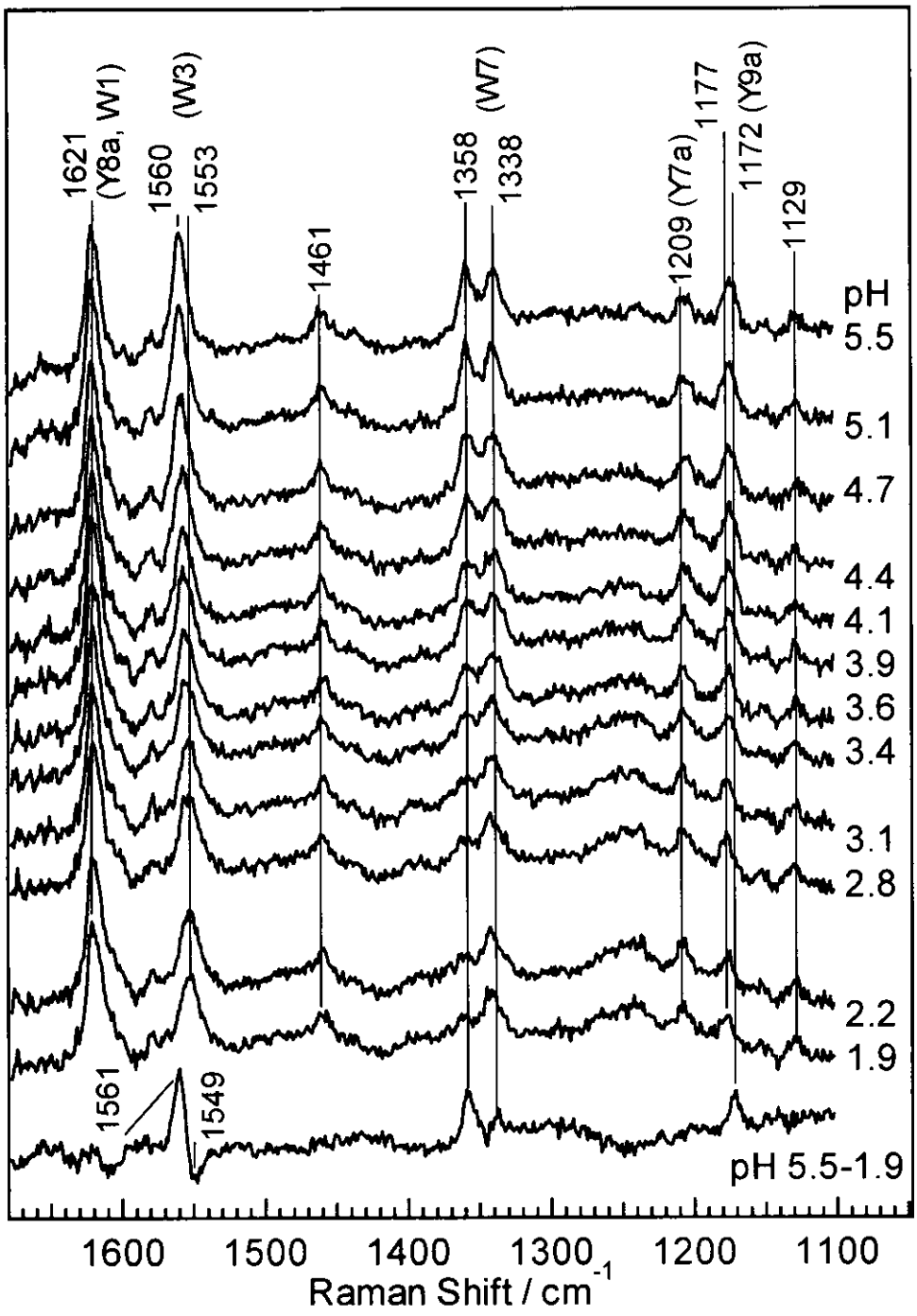


Figure 3-2.3. The stationary state UVRR spectra of 60 μM horse apoMb at various pH at 20 °C. The laser power was 0.2 mW and total exposure time at one spectrum was 2 hours. The apoMb solutions at various pH values were prepared with a suitable amount of 1 M HCl from the same stock solution. The spectra were normalized with the water band at 1650 cm^{-1}, and the water contribution was subtracted from the observed spectra.

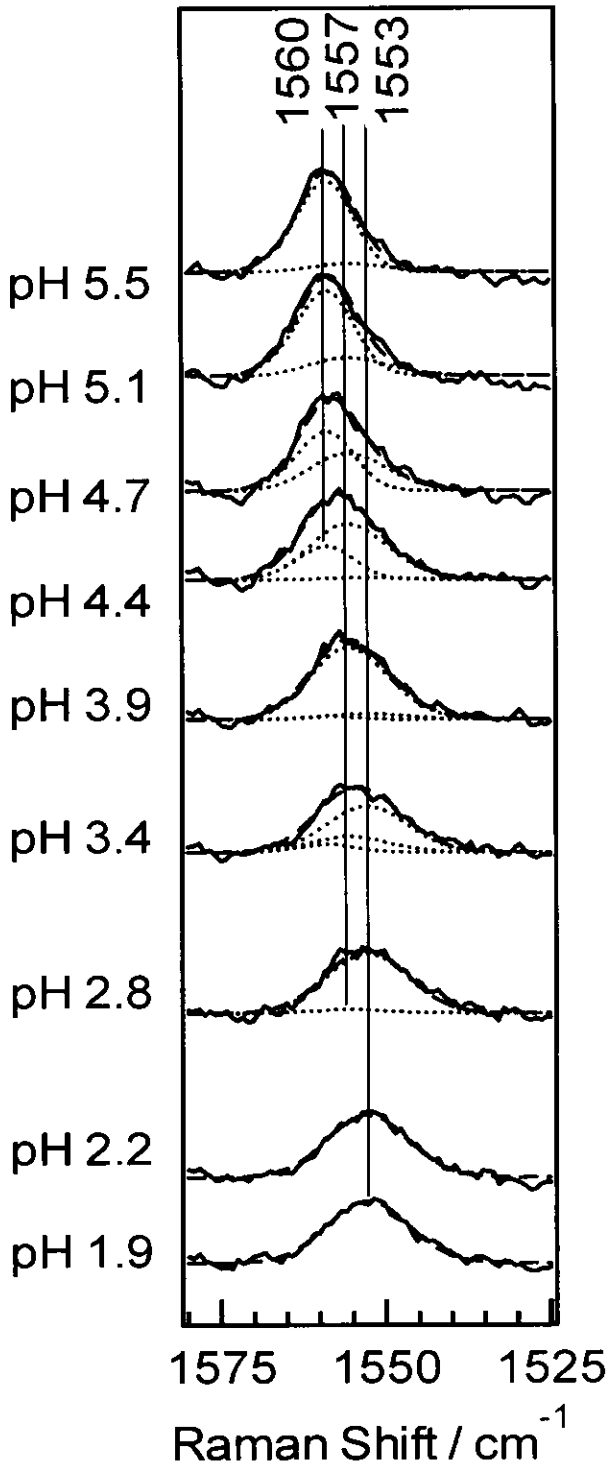


Figure 3-2.4. The expanded W3 spectra of apoMb in equilibrium at various pH. The solid lines denote the observed spectra, and the dotted lines show the assumed spectra for pure native, intermediate and unfolded species, which have a band center at 1559, 1557, and 1553 cm^{-1}, respectively. The dashed lines represent the simulated spectra (weighted sum of the three assumed spectra).

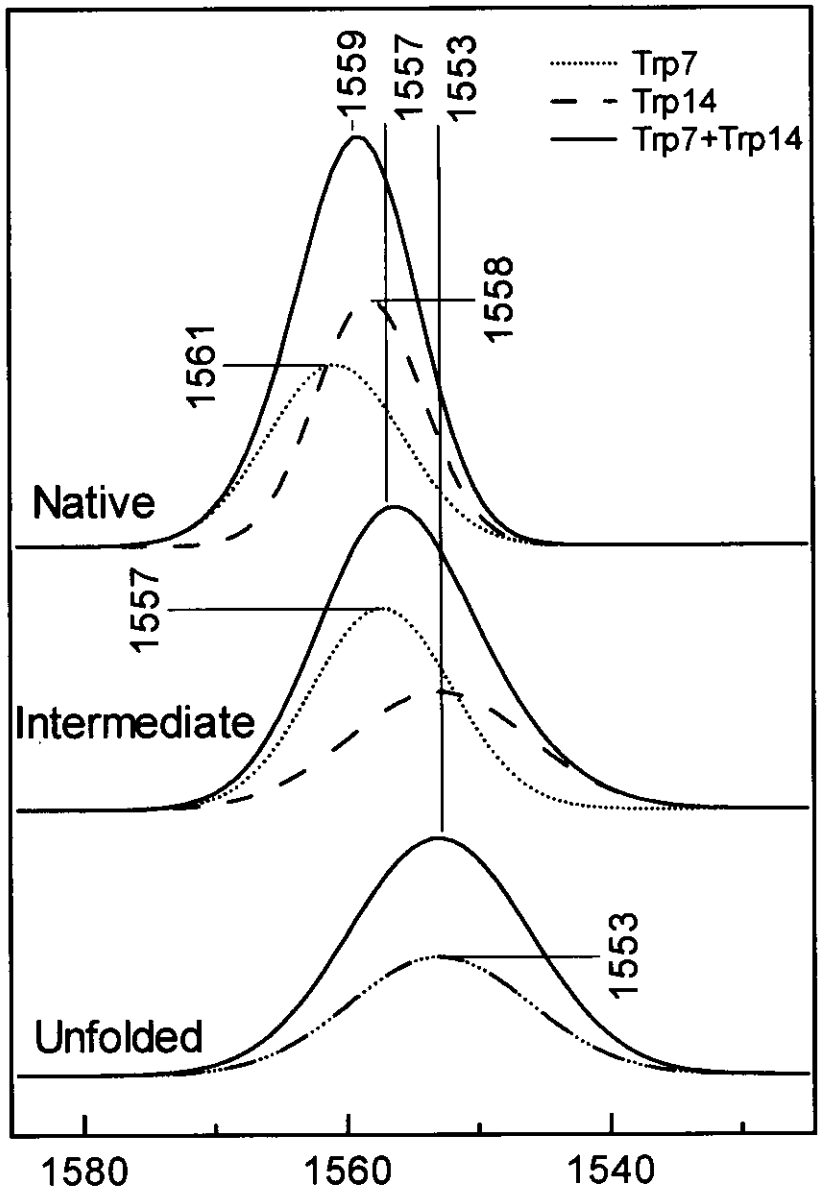


Figure 3-2.5. The decomposition of W3 band. Trp7 and Trp14 in the native state give the W3 bands at 1558 and 1561 cm^{-1}, respectively, from the mutant. Both Trp7 and Trp14 in the unfolded state and Trp7 in the intermediate are considered to have the same band shape and intensity as those of aqueous tryptophan solution. The spectrum of Trp14 in the intermediate species is unknown. The spectral fitting calculations using the Igor Pro software (WaveMetrics) suggested its presence at 1557 cm^{-1}.

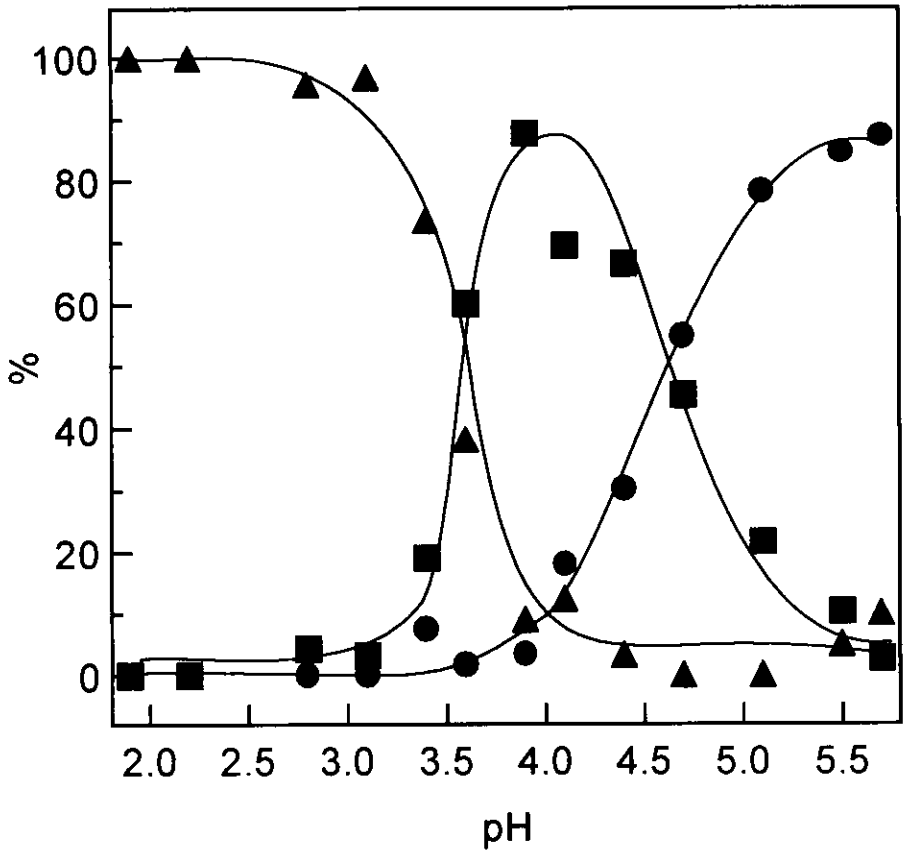


Figure 3-2.6. The pH dependence of the populations of the native (N, ●), intermediate (I, ▲) and unfolded species(U, ■) calculated by the W3 band. The midpoint pH for the N ↔ I and I ↔ U transitions are pH 4.5 and pH 3.5, respectively.

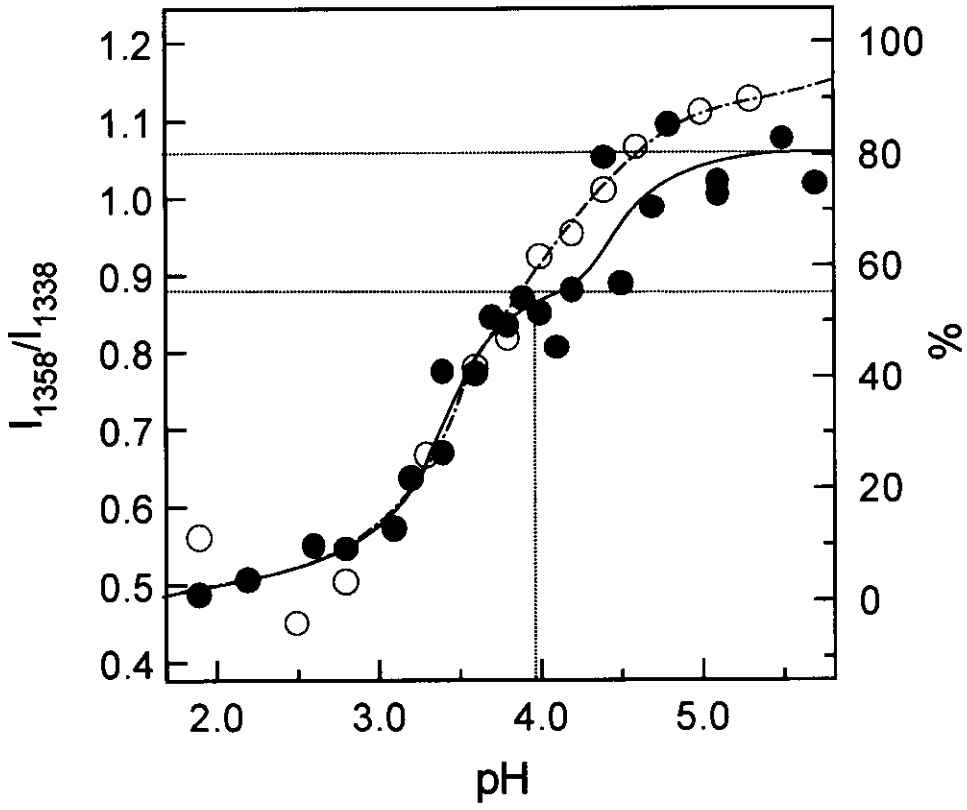


Figure 3-2.7. The pH dependence of the intensity ratio of W7 doublet (R=I_{1358}/I_{1338}) for 60 μM horse (●) and sperm whale (○) apoMbs at 20 °C. The right axis denotes the degree of hydrophobicity, which was scaled by assuming that the R-values of holoMb at pH 7 and of apoMb at pH 1.9 are 100 and 0 %, respectively. The R-values of native apoMb and the equilibrium pH 4 intermediate are 1.05 and 0.85, respectively, meaning that hydrophobicity around Trp for the native apoMb and the pH 4 intermediate is decreased by 20 % and 45 % than that of native holoMb, respectively.

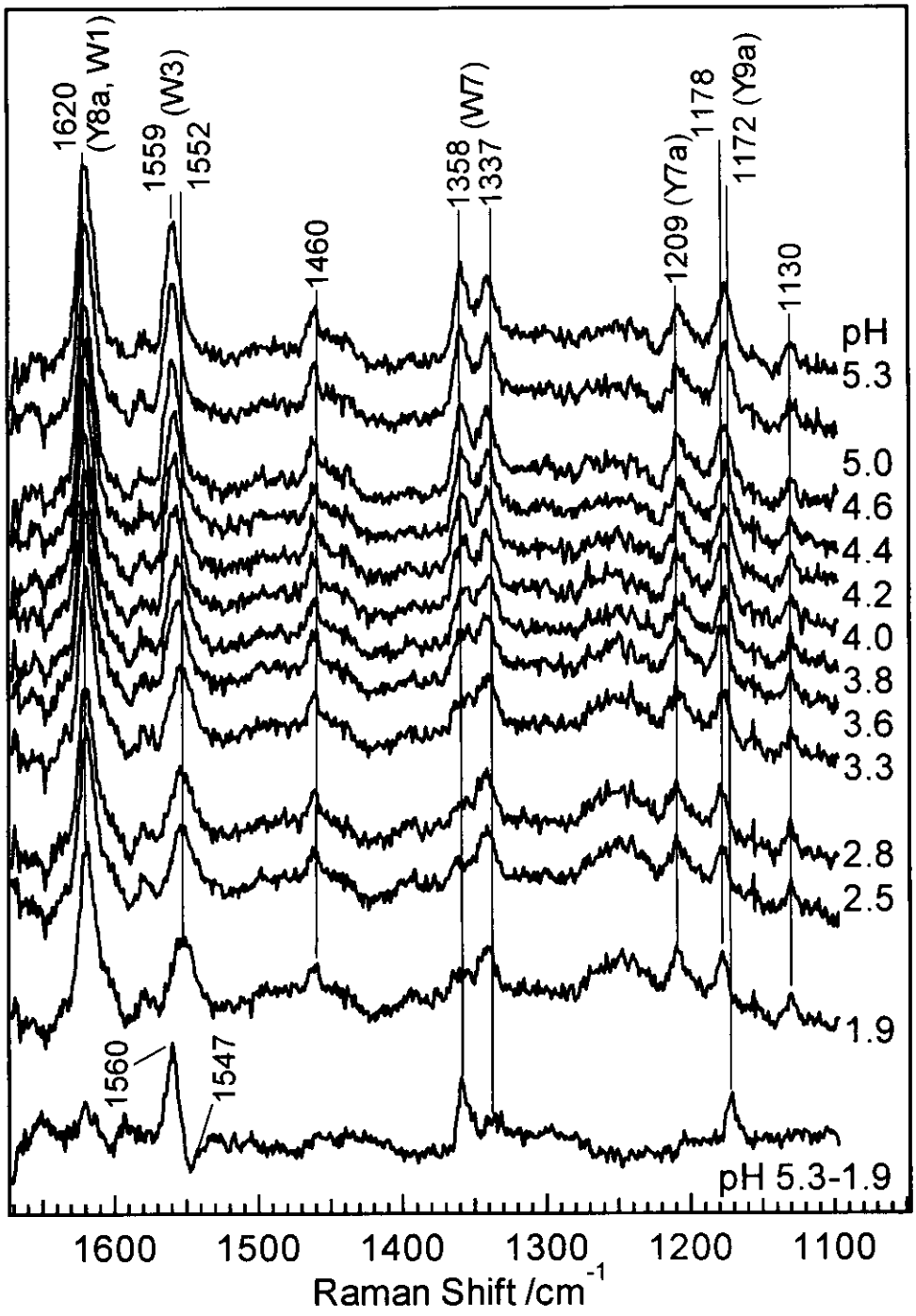


Figure 3-2.8. The stationary state UVRR spectra of 60 μM sperm whale apoMb at various pH at 20 °C.

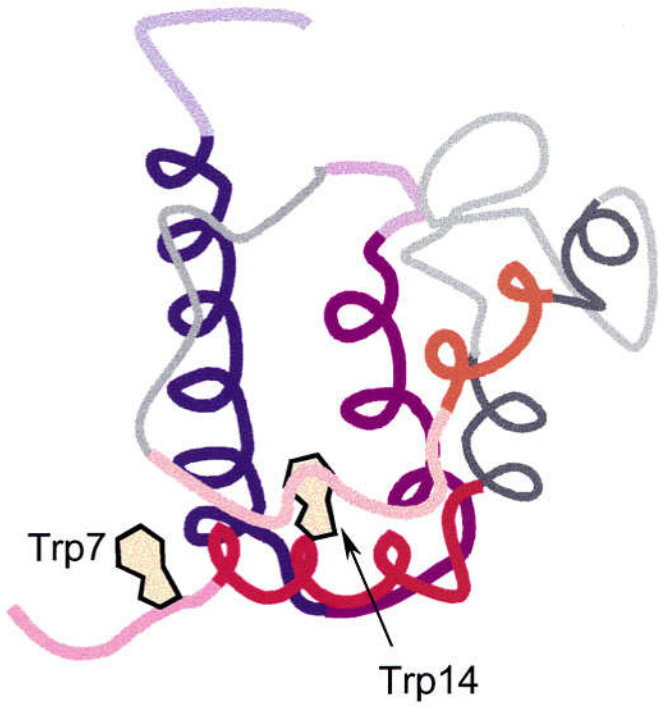


Figure 3-2.9. The schematic diagram of pH 4 folding intermediate of horse apoMb; Trp7 is exposed to the solvent and no secondary steric interaction, whereas Trp14 is buried inside of protein and has some steric hindrance that is different from packing of native state. The regions drawn by light colors indicate the unfolded structures and dark color for fold structures.

3-3. Observation of Submillisecond Protein Folding Events of Horse Apomyoglobin Detected by UVRR Spectroscopy

(Contents of this chapter has been submitted)

Abstract

The kinetic folding intermediates of apomyoglobin were investigated by 244 nm-excited UV resonance Raman spectroscopy combined with a new rapid mixer. The dead time for mixing in the newly constructed flow-mixer was determined to be 150 μs with UV Raman spectral changes of imidazole to imidazolium upon mixing with an acid. The pH-jump experiments of apomyoglobin (apoMb) from 2.2 to 5.6 conducted with this device demonstrated the presence of three folding intermediates. Based on the analysis of W3 and W7 bands of Trp 7 and Trp14, the first intermediate, formed before 250 μs, involved inclusion of Trp14 into an α-helix from a random coil. In the second intermediate, formed around 1 ms after the start of folding, the surroundings of both Trp7 and 14 were significantly hydrophobic, suggesting the formation of hydrophobic core. In the third intermediate appearing around 3 ms, the hydrophobicity was relaxed to the same level as that of the pH 4 equilibrium intermediate, which was investigated in detail with the stationary state technique. The change from the third intermediate to the native state takes time longer than 40 ms, while the appearance of the native spectrum after the mixing of the same solutions was confirmed separately.

Abbreviations: Im and ImH^+, imidazole and imidazolium; N, native; I, intermediate; U, unfolded.

Introduction

Since Anfinsen's finding of the principal dogma on a protein structure (*1*), understanding of the tertiary structure formation on the basis of the primary structure has become an attractive theme, and all the newly developed techniques which may possibly elucidate it, have been applied to promote current understanding of protein folding, particularly in relation with the recently developed funnel model(*2*). The early stage of protein folding has been explained with a few controversial models; the framework model(*3*) postulates that the native secondary structures are independently formed as local elements, the diffusion-collision model(*4, 5*) extends it so that the individual elements diffuse until they collide and are associated into a specific tertiary structure, and the hydrophobic collapse model(*6*) assumes that overall condensation of polypeptide segments occurs through clustering of non-polar side chains first and then, secondary structure formation follows in the restricted space. In either case, the early processes take place rapidly and therefore, time-resolved experiments are desirous to select a proper model. Side chains of individual amino acid residues are arranged as a result of balance among various interactions involving hydrogen bond formation and hydrophobic and electrostatic attraction or repulsion, being accompanied with some dynamics of the main chain. However, with ordinary techniques available, it is generally difficult to get specific information on the side chain arrangements and their environments in the events occurring faster than millisecond time regime.

Apomyoglobin is a one of the most studied proteins about its folding reactions (Section 2-2). The hydrogen exchange pulse labeling method succeeded to determine the order of the formation of α-helix and we could image the folding pathway after few milliseconds (*7*). However, it is already finished the formation of A, G and H helix

and the AGH hydrophobic core is seemed to be appeared in this regime.

The kinetic studies before the formation of the AGH helix core are fewer because the trigger of the reactions and the detection with time-resolved required the special instruments. The fastest trigger, which could be applied the apoMb without any modification, is the temperature jump with the pulse light for the refolding process from the cold denature state (*8*) and for fast unfolding reaction (*9*). However, it remains the disputable and essential questions that the cold denaturation and thermal denaturation are the same as pH denaturation and the unfolding and folding pathway can discuss according to the same aspect with pH perturbation. The pH-jump trigger from acid to neutral pH can be performed only the mixing, the fastest dead time is a few microseconds (*10, 11*). The equipment of the fluorescence spectroscopy with the rapid continuous flow system resulted that two intermediate, which adopt the two forms of the pH 4 intermediates in the equilibrium pH experiments of sperm whale apoMb (*11, 12*). Since the Trp14 was located on the A-helix and near the G and H helix, it has been using the good probe of the conformational change to detect the fluorescence (*8, 11, 13*). While the advantage of the fluorescence spectroscopy, sensitive and simple without long accumulation of the spectra, is suited for the kinetic experiments, it is difficult to interpret the structure of the intermediate.

In section 3-2, the pH 4 folding intermediate of apomyoglobin (apoMb) was distinguished by 244 nm excited ultraviolet resonance Raman spectroscopy (UVRR). Accordingly, we have constructed a rapid mixing device for pH-jump and combined it with UVRR spectroscopy to explore the primary process in folding of apomyoglobin (apoMb).

Materials and Methods

Construction and Check of Rapid Mixing Device Submillisecond kinetic folding experiments were performed using the rapid-continuous flow apparatus constructed in this study which is essentially the same as that reported by Takahashi et al (Figure 3-3.1) (*10*). The mixing effect of this apparatus was checked by the protonation reaction of imidazole (Im) to imidazolium (ImH^+) ion, in which 100 mM Im solution was mixed with 100 mM citric acid under the conditions similar to those used for the measurements of apoMb. In the equilibrium conditions after mixing, Im is expected to change into ImH^+ completely. The Raman bands of Im (1257 and 1327 cm^{-1}) were easily distinguished from those of ImH^+ (1212 and 1457 cm^{-1}) (Figure 3-3.2). On account of resonance enhancement effect of intensity on Im bands upon excitation at 244 nm, the remaining Im as small as 3% could be detected. Complete absence of the Im band at 150 μs after mixing was confirmed in the kinetic pH jump experiment, although the data after 250 μs were used for the analysis of apoMb in this study. The protein concentration of sample solution for the kinetic pH-jump was made to be 120 μM with 10 mM HCl at pH 2.2, and it was mixed with the same volume of 100 mM potassium phosphate buffer at pH 6.2 to achieve pH-jump to pH 5.6.

Results

The time-resolved 244 nm excited UVRR spectra are depicted in Figure 3-3.3, where the stationary state spectra at pH 2.2 and 5.6 are presented at the bottom and top, respectively. The Trp W3 band at 1553 cm^{-1} in the acidic solution is shifted to 1559 cm^{-1} and sharpened in the neutral solution, and the relative intensity of the W7 doublet

($R = I_{1358}/I_{1338}$) is also changed. This feature is well reproduced by the stationary spectrum of the mixture of the unfolded sample with the refolding buffer depicted with a dashed line. As seen in the raw spectra, the relative intensity of W7 doublet at 250 μs after pH jump is similar to that of the stationary spectrum at pH 2.2 but has been changed at 1 ms.

Figure 3-3.4 shows the time profile of the *R*-value. It is noted from this plot that the *R*-value becomes maximum around 1 ms but decreases slowly around 2-5 ms. Fitting of the *R*-value changes with the model of successive two first order reactions ($I_1 \rightarrow I_2 \rightarrow I_3$), which is represented by the solid line in the inset, yielded the rate constants of $k_{1,2} = 2.5 \times 10^3$ s^{-1} and $k_{2,3} = 0.33 \times 10^3$ s^{-1} (correspond to $\tau_1 = 0.4$ ms and $\tau_2 = 3$ ms). Here I_1 is not always equal to a denatured state (U) as will be discussed later. The *R*-value was 1.05 for I_2 and 0.83 for I_3. The value of 1.05 is close to that of the native state, and therefore cannot be interpreted by the contribution of a single Trp14 residue, if Trp7 is left exposed in the hydrophilic environments. The results suggest that the environment of Trp7 also becomes hydrophobic in I_2 and returns somewhat hydrophilic in I_3. It may mean that Trp7 is involved in the hydrophobic condensation on the $I_1 \rightarrow I_2$ process and in the hydrophilic rearrangements on the $I_2 \rightarrow I_3$ process. Note that such features have not been recognized in the absence of time-resolved UVRR experiments and that the spectra for the equilibrium unfolding process reflect the species with longer life times but are apt to pass unnoticed transient species with a short lifetime.

Figure 3-3.5 shows the expanded time-resolved UVRR spectra in the W3 region. The frequency of W3 band is already shifted from 1553 to 1557 cm^{-1} in the earliest delay time after pH jump (0.25 ms in Fig. 3-3.3) at which the W7 doublet are not altered yet. The frequency of 1557 cm^{-1} remains unchanged between 0.4 and 3.2 ms and is the

same as that observed for the equilibrium pH 4 intermediate. This fact means that the structural change on the $\chi_{2,1}$ angle (torsion angle at C_2-C_3-C_β-C_α of Trp side chain) occurs prior to the environmental change and yields the first intermediate (I_1) before 0.25 ms, while Trp residues are still left in the hydrophilic environments.

The frequency change of W3 band from 1557 to 1559 cm^{-1}, which is the value of native state (N), did not occur even 40 ms after the pH jump, at the allowable longest delay time with the present apparatus. This means that the conversion from the last intermediate (I_3) to the native state must be a slow process. Other studies also reported that the conversion of the last intermediate to the native state takes the order of second (*7*). While the side chain of Trp14 is already fixed in a certain structure in the I_1 state, the final arrangement of the $\chi_{2,1}$ angle would take place in the rate-limiting step ($I_3 \rightarrow N$) of the whole folding reaction.

Discussions

What Does Occur in the Earliest Stage of Folding Event of ApoMb ? Protein folds from random conformations to a unique ordered structure without random searches of astronomical numbers of conformations (Levinthal paradox)(*14*). A view from this study is illustrated in Figure 3-3.6. In this experiment, the W3 band had been shifted in I_1 state before hydrophobic environment was produced. The frequency shift from 1553 cm^{-1} to 1557 cm^{-1} ($U \rightarrow I_1$) suggests the secondary structure change from a random coil to an α helix, and this was confirmed by an independent experiment on a model peptide of the A helix. Although the model peptide had some substitution of amino acid residues from the A helix to increase the helicity and solubility, it formed the

α helix in an organic solvent but not in water. The W3 band of Trp was shifted to a higher frequency in the organic solvent compared with in water, where its frequency was the same as that in a free tryptophan solution (see Section 3-4).

The change of *R*-value clearly indicated that the state (I_2) with increased hydrophobicity in the environments of Trp has been formed prior to reach to the equilibrium intermediate (I_3). The side chain conformation of Trp, suggested by the 1557 cm^{-1} band of the intermediates (I_2 and I_3), is a consequence of restriction from other parts of protein, and its release from the restriction allows assembly of the native hydrophobic core ($I_3 \rightarrow N$). The origin of the structural restriction of the Trp side chain might be the local contacts owing to the formation of α helix. This view is compatible with the fluorescence anisotropy decay kinetics of Trp which showed that movement of Trp residues was restricted in the pH 4 intermediate (*13, 15*) and that this restriction was stronger than that in the coil-helix transition of model peptides (*16*). According to X-ray crystallography data (*17*), the indole ring of Trp14 has contacts with Val17 of A helix. The fact that the formation of local structure is faster than the increase in hydrophobicity in the environment around the aromatic residues seems to suggest that the framework model(*3*) is more suitable than the hydrophobic collapse model (*6*). However, since the formation of the local structure is limited and the local structure seems unstable by itself, it is more reasonable to emphasize the diffusion-collision process (*4, 5*) following the formation of the framework.

The temperature-jump experiment from cold-denaturation exhibited very fast fluorescence quenching (2.0×10^5 s^{-1}) and it was considered that the A helix was in the proximity of G and H helices, that is, rapid formation of AGH helices core(*8, 18*). The IR observation for α helix formation indicated that the formation rate of AGH helices

core reached the diffusion limit (1.0×10^4 s^{-1}), although the total helicity was ~10 % (*19*). The time scale of reaction would depend on conditions of the system, particularly in the denatured state, because cold denaturation restricts the space and conformation for related residues. It seems likely that local α helix conformations, which include small hydrophobic patches, diffuse until colliding together and are stabilized by adhesion of three helices referred as the AGH hydrophobic core. Presence of another intermediate at low pH is suggested by fluorescence(*11, 12, 20*) and IR spectra(*19*). The characters of these intermediates are not clear except that A, G and H helices are partially formed and have contacts. This species may correspond to I_2 of the present kinetic intermediates, although we failed to characterize these intermediates in the equilibrium acid-unfolding process.

Formation of Large Hydrophobic Core The second kinetic intermediate (I_2), which appeared with the formation rate of 2.5×10^3 s^{-1}, has two Trp residues in the hydrophobic environments. Although Trp7 is exposed to surface in the equilibrium pH 4 intermediate, the environment of Trp7 in this intermediate was hydrophobic. To interpret this observation, presumably a large scale hydrophobic collapse involving Trp7 would have occurred on the random coils, because the AGH helices core only cannot keep many hydrophobic side chains without exposing to solvent. It is possible that the hydrophobic side chain on the frayed E and F helices and the N-terminal residues in the frayed A helix gathered around the AGH helices core, playing a role as a hydrophobic nucleus. In this intermediate, the helicity would not increase and the backbone could move flexibly. The formation rate of A helix and flexibility around Trp7 might be different between horse and sperm whale Mbs in this time regime.

To form the third intermediate (I_3), which is close to the equilibrium pH 4 intermediate, A and B helices would be extended and the melted E helix would be located in the topologically correct place, while Trp7 is made exposed to the solvent again.

The Non-specific Packing of the Intermediate around Trp14 Many experiments using mutants tried to determine whether the factor stabilizing the equilibrium pH 4 intermediate is the native-like specific hydrophobic contacts or non-specific hydrophobic interaction (*21-23*). The Trp14 → Phe mutation had a small effect on the stability of the equilibrium pH 4 intermediate, indicating that Trp14 contributes to the native specific interaction as well as to non-specific hydrophobic interaction (*23*). The independence of the frequency shift of W3 band from the increase of hydrophobicity in the earlier stage of folding ($I_1 \rightarrow I_3$) supports the idea that the Trp14 has no specific interaction with the hydrophobic core present in the equilibrium pH 4 intermediate. On the other hand, packing process from the intermediate to the native structure ($I_3 \rightarrow$ N) caused the frequency shift of W3 band from 1557 to 1559 cm^{-1} (change of the C_2-C_3-C_β-C_α torsion angle) under the same hydrophobicity.

Conclusion

Taking advantage of the sensitive dependence of W3 frequency on the torsion angle of Trp indole ring, we succeeded in detecting the structural change of Trp side chain during the formation ($U \rightarrow I_1$) and association ($I_3 \rightarrow N$) of helices of apoMb from UVRR experiments. Mobility of Trp side chain was somewhat restricted in the very

early stage (I_1), though the restriction was different from the interaction in the native state. The change in hydrophobicity around Trp revealed by intensity ratio of the W7 doublet suggested that the local secondary structure was formed in the I_1 state prior to the hydrophobic collapse. Since formation rates of α helices are different between regions I and II and collisions among α helices occurred in the first step ($U \rightarrow I_1$) before the formation of A helix was completed, the folding process of apoMb might be approximately explained by the diffusion-collision model. This is the first direct observation for the structural restriction and environmental changes of Trp side chain of apoMb in the early stage of folding from the almost unfolded structure.

References

1. Anfinsen, C. B. (1973) *Science 181*, 223-30.
2. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) *Proteins 21*, 167-95.
3. Kim, P. S., and Baldwin, R. L. (1990) *Annu Rev Biochem 59*, 631-60.
4. Bashford, D., Cohen, F. E., Karplus, M., Kuntz, I. D., and Weaver, D. L. (1988) *Proteins 4*, 211-27.
5. Karplus, M., and Weaver, D. L. (1994) *Protein Sci 3*, 650-68.
6. Dill, K. A. (1985) *Biochemistry 24*, 1501-9.
7. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892-6.
8. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *nature structural biology 3*, 923-926.
9. Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., and Dyer, R. B.

(1997) *Biochemistry 36*, 15006-12.

10. Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) *Nat Struct Biol 4*, 44-50.
11. Jamin, M., u, and Baldwin, R. L. (1999) *J Mol Biol 292*, 731-740.
12. Jamin, M., and Baldwin, R. L. (1998) *J Mol Biol 276*, 491-504.
13. Tcherkasskaya, O., Ptitsyn, O. B., and Knutson, J. R. (2000) *Biochemistry 39*, 1879-89.
14. Levinthal, C. (1968) *J. Chim. Phys. 65*, 44-45.
15. Kemple, M. D., Buckley, P., Yuan, P., and Prendergast, F. G. (1997) *Biochemistry 36*, 1678-88.
16. Semisotnov, G. V., Zikherman, K. K., Kasatkin, S. B., Ptitsyn, O. B., and Anufrieva, E. V. (1981) *Biopolymers 20*, 2287-2309.
17. Evans, S. V., and Brayer, G. D. (1988) *J Biol Chem 263*, 4263-8.
18. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *Proc Natl Acad Sci U S A 93*, 5759-64.
19. Gilmanshin, R., Callender, R. H., and Dyer, R. B. (1998) *Nat Struct Biol 5*, 363-5.
20. Gilmanshin, R., Dyer, R. B., and Callender, R. H. (1997) *Protein Sci 6*, 2134-42.
21. Kay, M. S., Ramos, C. H., and Baldwin, R. L. (1999) *Proc Natl Acad Sci U S A 96*, 2007-12.
22. Hughson, F. M., Barrick, D., and Baldwin, R. L. (1991) *Biochemistry 30*, 4113-8.
23. Kay, M. S., and Baldwin, R. L. (1996) *Nat Struct Biol 3*, 439-45.

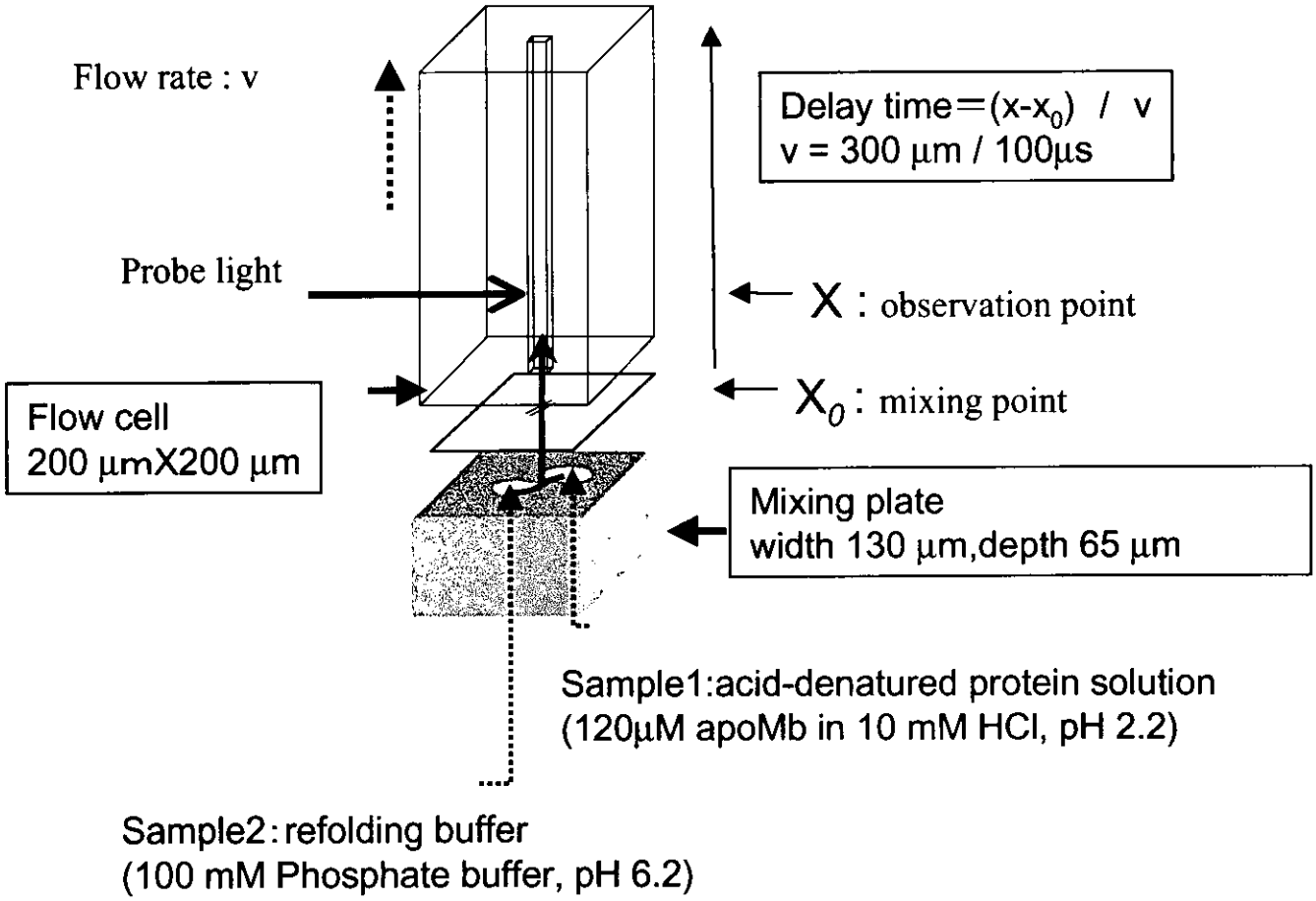


Figure 3-3.1. Continuous-flow rapid mixing apparatus. Two solutions to be mixed, protein sample and reaction buffer, enter the T shaped nozzle with high velocity from the opposite ends of 130 μm wide and 65 μm deep channel on the top of metal plate and meet at a point directly below the entrance of an observation cell. The quartz observation cell is tightly pushed against the metal plate with thin teflon spacer. The distance between the mixing point and laser excited point is correlated to the delay time of folding reaction. The Raman scattering light was collected at right angle.

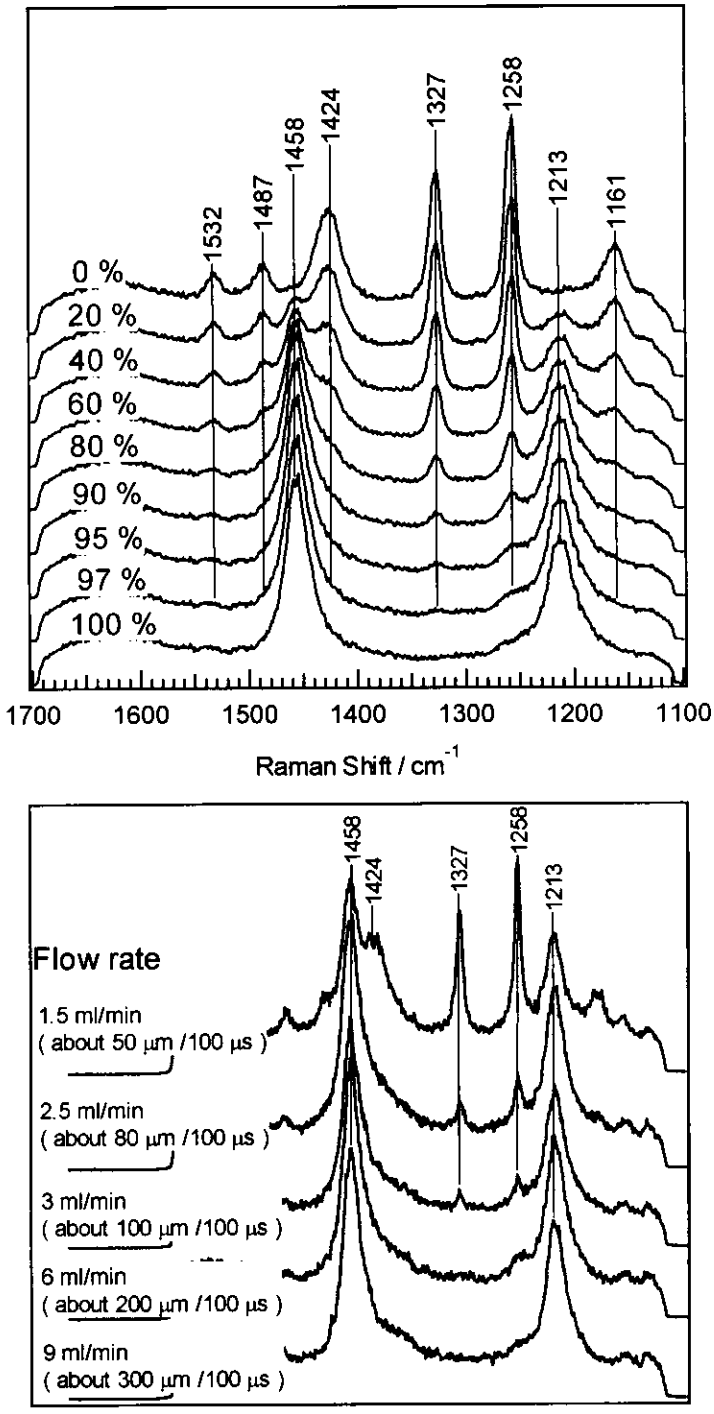


Figure 3-3.2. The UVRR spectra of imidazole and imidazolium ion. Above spectra are the sum of the spectra of the imidazole (0 %) and imidazolium ion (100%) with various ratio. Below spectra were the results of pH jump of imidazole to low pH and detected at 450 μm apart from the mixing plate with various flow rate; 9 ml/min (150 μs), 6 ml/min (225 μs), 3 ml/min (450 μs), 2.5 ml/min (540 μs) and 1.5 ml/min (900 μs). The mixing efficiency and dead time significantly depends on the flow rate.

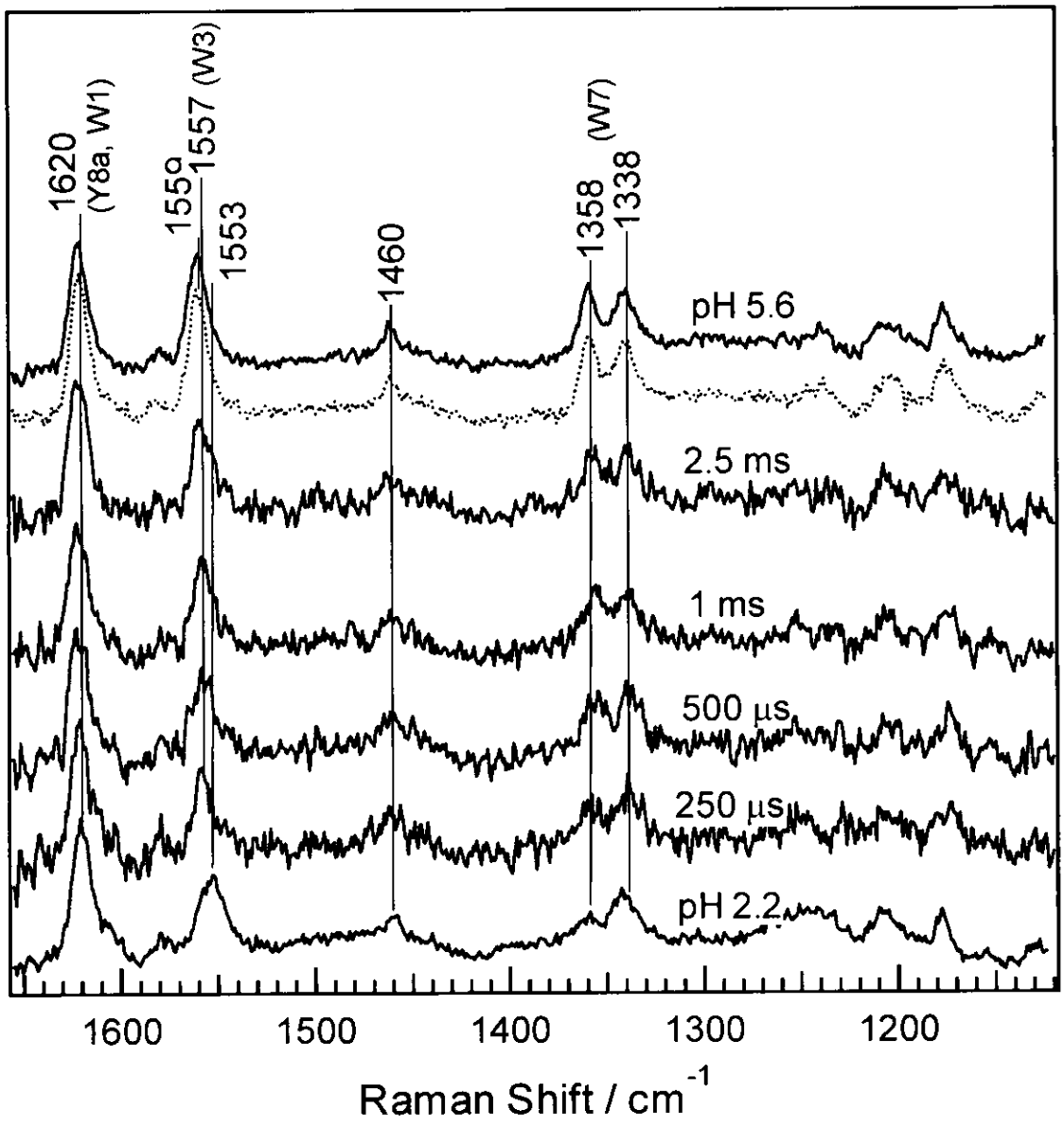


Figure 3-3.3. The time-resolved UVRR spectra excited by the 244 nm CW laser following pH jump from 2.2 to 5.6 at 20 °C. The top and bottom traces represent the stationary state spectra obtained at pH 5.6 and 2.2, respectively. The spectrum represented by a dotted line shows the spectrum of the stationary mixture of the unfolded sample with the refolding buffer. Delay times following mixing of two solutions by the rapid mixing device is specified on each spectrum.

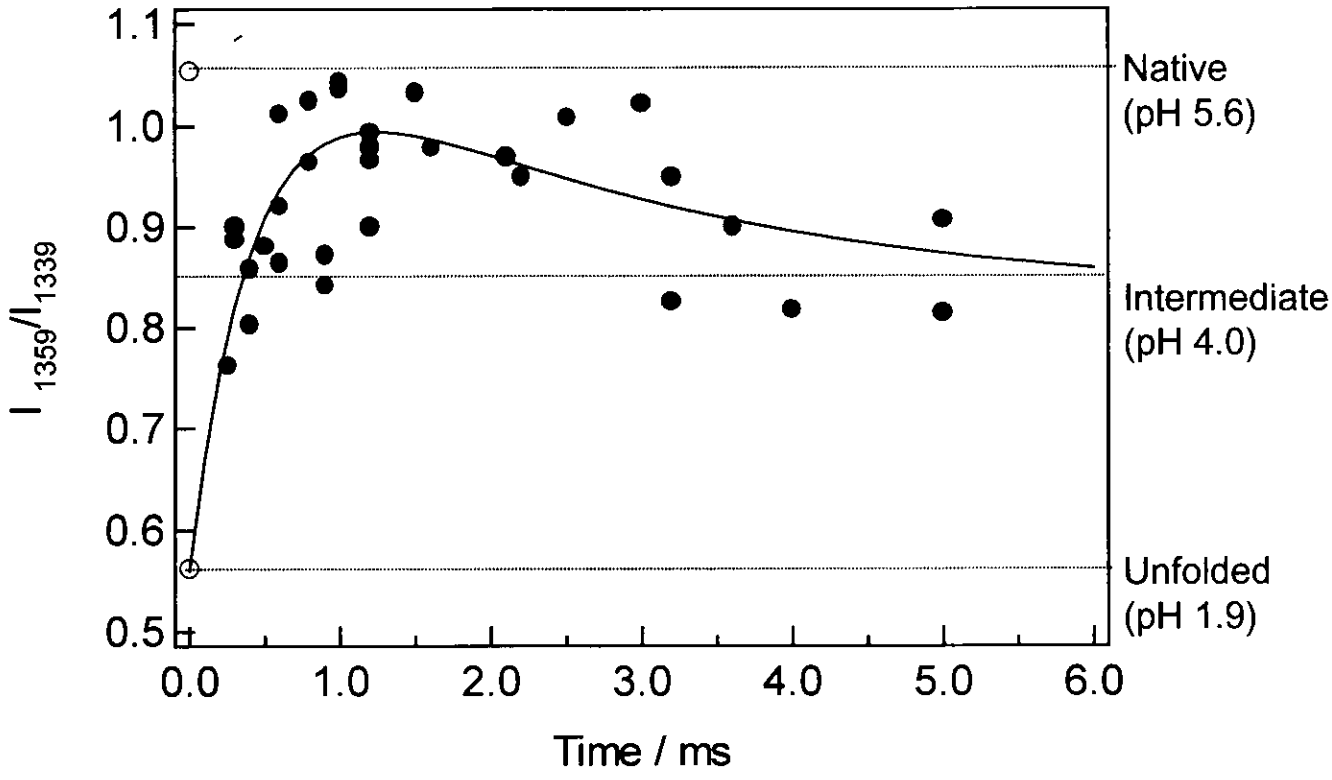


Figure 3-3.4. The time dependence of observed *R*-values (closed circles). The *R*-values of the pH 4 intermediate (R=0.56) and the native state (R=1.05) are also included (open circles). The solid line denotes the theoretical results calculated under the assumption that two successive first order reactions ($I_1 \rightarrow I_2 \rightarrow I_3$) occur to change the *R*-value; $\tau_1 = 0.4$ ms and $\tau_2 = 3$ ms.

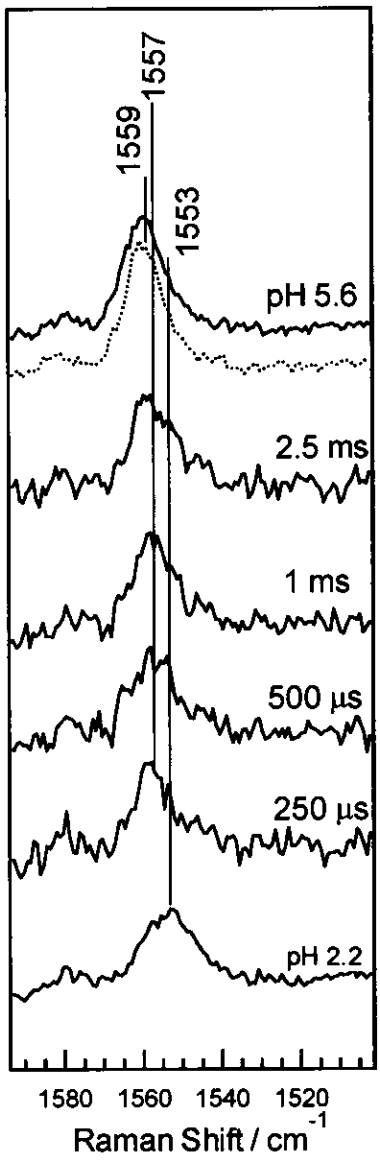


Figure 3-3.5. The expanded W3 bands of the time-resolved UVRR spectra between 0.25 ms and 3.2 ms following pH-jump from 2.2 to 5.6. The frequency of W3 band, which is sensitive to the torsion angle of Trp indole plane, is 1553 cm^{-1} for (U), 1557 cm^{-1} for I_1, I_2 and I_3, and 1559 cm^{-1} for N. The spectrum represented by a dotted line shows the spectrum of the stationary mixture of the unfolded sample with the refolding buffer.

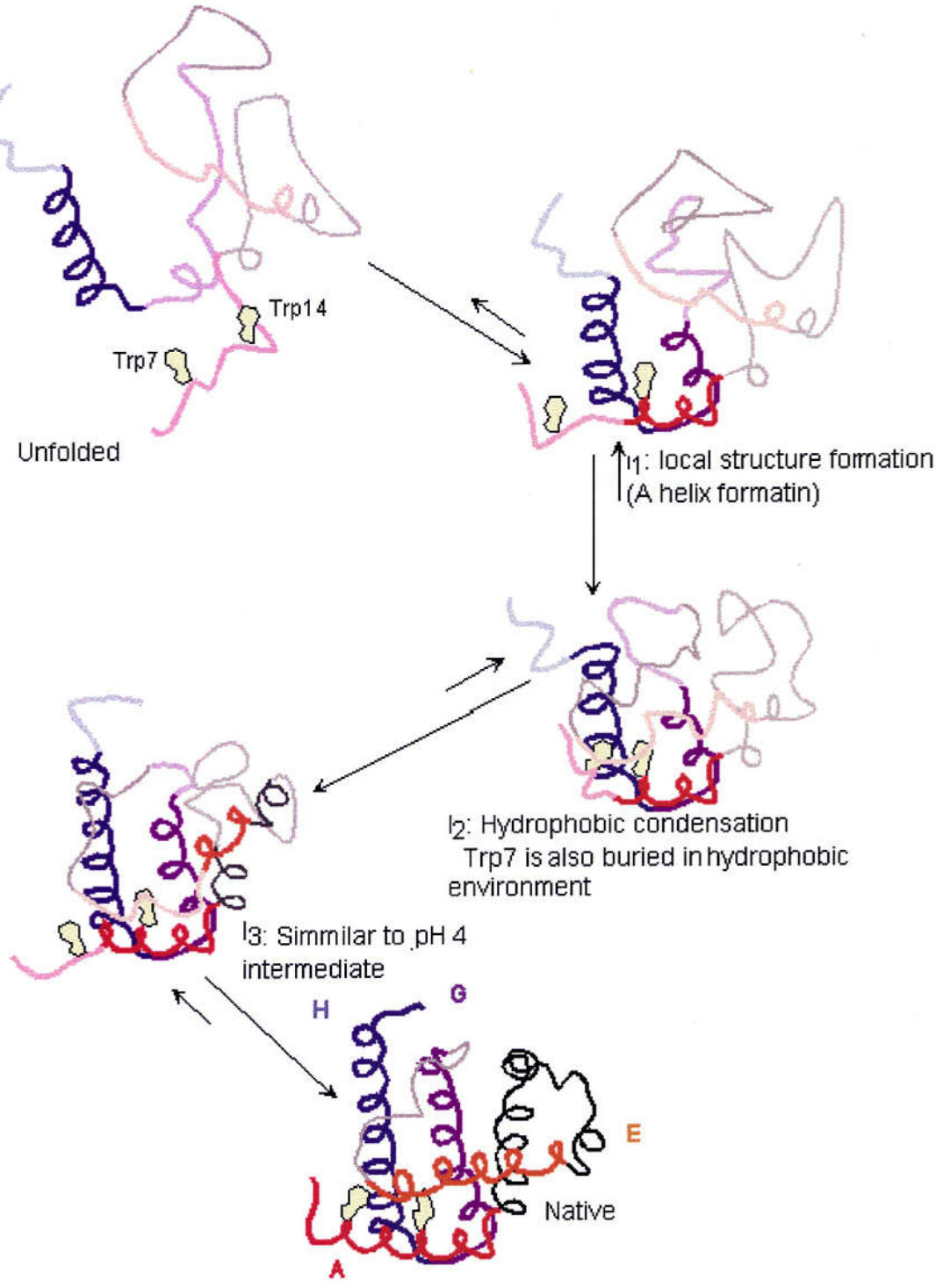


Figure 3-3.6. Schematic diagram for folding of apoMb derived from the present kinetic UVRR experiments. Trp7 and Trp14 are attached on red segment. I_1 has a local structure for which the frequency of W3 band is shifted to 1557 cm^{-1} in the dead time of the pH jump. I_2 contains a large hydrophobic core around the AGH helices core. The character of I_3 seems to be similar to the equilibrium pH 4 intermediate, in which the torsion angle of Trp indole plane is different from that of native state.

3-4. The Effect of the Helix-Coil Transition on the Conformation of Trp Side Chain in a Model Peptide of A Helix of Myoglobin

Abstract

Model peptide of A helix of myoglobin was constructed to analyze the conformational change of Trp side chain by the formation of α helix. The conformation of Trp side chain was investigated with the 244 nm excited ultraviolet resonance Raman (UVRR) spectroscopy. It was described in previous study (section 3-2, 3) that, the W3 band of Trp residue was sensitive to the conformational change of Trp side chain. This band was observed at 1561, 1557 and 1553 cm^{-1} for Trp 14 in the native, the pH 4 intermediate and unfolded states, respectively. The proposal of this experiment is to explain the structural implication of the frequency change. It became clear from the experiment on A peptide that the W3 band shift from 1553 to 1557 cm^{-1} came from the local interaction, that is α helix formation, rather than tertiary contacts. The result is strongly supported to the interpretation of the kinetic studies of apoMb folding reaction.

Abbreviations: CD, circular dichroism; NOE, nuclear overhauser effects; A peptide, model peptide of the A helix region of Mb; HPLC, high performance liquid chromatography; Tris, tris(hydroxymethyl)aminomethane; TFE, 2,2,2-trifluoroethanol; MeOH, methanol.

Introduction

α helix is one of the most important structural unit of protein structures and it has been well studied with model peptides and actual proteins. Although some studies reported that the formation rate of isolated α helix, such as Ala-rich peptide, seemed to be very fast, form nanoseconds to microseconds (*1*), the formation of α helix in actual protein occurs in various time regions and in various steps of folding process (*2, 3*). Some proteins were reported to perform the temporary formation of non-native α helix during the folding reaction (*4*). The formation of α helix is dependent not only on its intrinsic helicity, which is encoded in the sequence, but also on the surrounding conditions, especially influencing to its stabilization. In this time range, the noncovalent interactions among side chains, including hydrogen bond, electrostatic interaction and hydrophobic interaction, contribute to this stabilization. In many cases, a polar residue is localized on one side of α helix, which is exposed to solvent, and nonpolar residues are located on the other side of α helix and interact with the internal structure is localized by nonpolar residues. Thus, it is said that the hydrophobic interaction especially contributes to stabilize the α helix.

The folding process of protein is the process that reduces the number of conformation, namely conformational entropy (*5*). Not only the overall conformational entropy of the backbone but also that of the side chain reduces during the protein folding from unfolded state to a unique native state with restriction of the rotation. Although final folded protein is stabilized by the free energy only in the range from 5 to 10 kcal/mol compared to the unfolded state as a result of counteracting the large opposite forces, the cost of restricting side chain motion ($T\Delta S_{conf}$) is near −1 kcal/mol per residue in protein folding or −0.5 kcal/mol per rotamer (*6*). Thus, the restriction of

side chain motion is considerably important in protein stability. However, compared to the backbone motion, it is difficult to probe the dynamics of the side chain in an early stage of the protein folding. In the stationary state, fluorescence anisotropy decay (*7, 8*) method and nuclear Overhauser effects (NOE) (*9*) are applied to measure the mobility of Trp side chains.

The ultraviolet resonance Raman (UVRR) spectroscopy is one of the strong tools to yield the structural and environmental information of the aromatic residues. We showed that the W3 band of Trp excited at 244 nm exhibited sensitive frequency change in the folding and unfolding processes of apomyoglobin (apoMb) in previous sections (3-2,3). The 244 nm excited UVRR spectra of W3 band of Trp14 could be located at three bands at 1561, 1557 and 1553 cm^{-1} in the native, intermediate and unfolded states, respectively. The equilibrium and kinetic studies indicated that the band at 1561 cm^{-1} means the tight and native packing and 1557 cm^{-1} means the presence of Trp in a local structure formed in the dead time. To reveal the conformational effect on the W3 band shift, the model peptide including the 19 residues is constructed with the sequence based on A helix of myoglobin, in which Trp 14 is located. The results provide not only the evidence that frequency shift of W3 band indicates the formation of α helix in apoMb, but also proposal that W3 band frequency is a good maker to the conformational interaction.

Materials and Methods

Preparation of Samples All reagents without any specific description were purchased from Wako pure chemical cooperation. The model peptide of A helix (A peptide) was

purchased from the Sawady Technology Company. A peptide and a tryptophan for reference were dissolved in the 50 mM Tris buffer at pH 8 and adjusted to 2.4 mM as a stock solution. The concentrations of them were calculated by the molar absorbance coefficient of Trp residue at 280 nm, $\varepsilon_{280} = 5690$ (M^{-1} cm^{-1})(*10*).

Measurements of CD Spectra CD spectra were measured at 20 °C by use of a Jasco J-720WI spectropolarimeter. The concentration of A peptide was adjusted to 70 μM at various solvent conditions. The helix content was calculated by the mean residue ellipticity at 222 nm using the method of Chen et al.(*11*).

Measurements of UVRR Spectra UVRR measurement were performed by using the instrumentation described in detail in section 2-2 (*12, 13*). The 244 nm excitation light was obtained by an intracavity frequency-doubling CW laser (Coherent, Innova FreD). The concentration of samples was adjusted to 0.6 mM and placed in a spinning cell. Raman scattered light was dispersed with a 1.26 m single spectrometer (Spex, 1269) equipped with a 3600 grooves/mm holographic grating, and detected with an intensified charge coupled detector (Princeton Instruments, ICCD-1024MG-E/1). The laser power at the sample point was 0.3 mW and total exposure time of one spectra was 20 min at 20 °C.

Results

Table 3-4.1 shows the amino acids sequences of A helix region in sperm whale myoglobin (swMb) and horse myoglobin (hMb) and the sequence of A peptide that we

prepared. The sequence of the A peptide is based on that of swMb, since the helical propensity of A helix of swMb is higher than that of hMb (*14*). We changed some residues for increasing the helical propensity and furthermore for increasing the solubility, since a peptide with the native sequence of A helix is hard to dissolve in water (*15*). Thus, Trp7 and other hydrophobic residues, which are for from Trp14 and are considered to have no interaction with Trp on the basis of X-ray structures of Mb, were changed to Ala or Glu. As a results, this model peptide could dissolve in water easily around pH 8.

Table 3-4.1. *Amino acid sequencse of sperm whale and horse Mbs in A helix region and that of model peptide*

	3	8	13	18
Sperm whale	-SEAEW	QQVLH	V**W**AKV	EADV-
Horse	-SDAEW	QQVLN	V**W**GKV	EADV-
Model peptide	N-SEAEA	QQAEH	A**W**AKV	EADG-C

To determine the structure of this model peptide, we measured the CD spectra at various solvent conditions at 20 °C (Figure 3-4.1). In the water with 50 mM Tris buffer at pH 8, A peptide has a random coil structure. The addition of alcohols, such as TFE and methanol, increased the negative ellipticity that indicated the formation of α helix. The helix contents were calculated by the mean residue ellipticity at 222 nm. Figure 3-4.2 is a plot of the helix contents of A peptide versus the alcohol concentration. In 95 % alcohol solvent, the model peptide had about 50 % helix content and the helix contents were almost proportional to the alcohol concentration.

Figure 3-4.3 shows the 244 nm excited UV resonance Raman spectra of free Trp solution and A peptide in various solvents. The Raman bands derived from solvents

were too strong to measure the Raman spectroscopy of Trp and A peptide in 100 % TFE and MeOH. Thus the concentration of TFE was adjusted to 70 % v/v and that of MeOH was to 75 % v/v. The contribution of the solvent bands were subtracted from the observed spectra and the resulted spectra were roughly normalized with the band intensity of the W5 band at 1460 cm^{-1}.

The spectra of W3 region are expanded above right of Figure 3-4.3. The W3 bands of free Trp solution and A peptide appeared at 1552 cm^{-1}. The W3 band of free Trp solution was hardly changed in other solvent conditions, except a slight change in 70 % TFE solvent that is considered to be due to the change of polarity or hydrogen bond. On the other hands, the W3 bands of A peptide in alcohol-added solutions were shifted to 1555 cm^{-1}. These differences between free Trp solution and Trp residue on A peptide obviously supported that Trp residue was influenced the conformational change of backbone, that is, the formation of α helix.

The intensity ratio (I_{1359}/I_{1340}) of W7 doublet band of Trp residue, *R*-value, is sensitive to the polar and nonpolar environment of Trp side chain (*16*). As Trp residue is located in more nonpolar environment, the band at high frequency side (1359 cm^{-1}) is increasing in intensity and *R*-value becomes higher. *R* values of free Trp solution and A peptide depend on the solvent, if they are lined up from the highest, 75 % MeOH, mixture of MeOH and TFE, 70 % TFE, water. In all solvent conditions, the *R* value of A peptide is higher than that of free Trp, when they are compared to the same solvent.

Discussion

In the previous kinetic studies of apoMb, we found that the W3 band shift from

1553 cm^{-1} in unfolded state to 1557 cm^{-1} occurred not later than 250 μs, although the environment of Trp side chain remained hydrophilic. Since the estimations of the α helical contents by the graph of figure 3-4.2 indicated that A peptide has helical contents about 30 % in the 70-75 % MeOH or TFE, the frequency of W3 band in A peptide at 1555 cm^{-1} is consistent with the estimation from that the W3 band of Trp on α helix is at 1557 cm^{-1}. This result indicates that the shift of W3 band came from the local interaction rather than tertiary contacts. The result in this section, that W3 band shift of model peptide during the coli-helix transition, strongly supports the interpretation of the kinetic studies of apoMb.

Mobility of Trp side chain were measured by time resolved fluorescence anisotropy (*7, 8, 17*) and NOE in NMR spectroscopy(*9*). The studies of the peptide suggested that some restriction of the mobility arise from the coil-helix transitions (*9*). However, Tcherkasskaya and coworkers suggested that the restriction of mobility of Trp side chain in the apoMb intermediate is more severe than that in the peptide (*9*). Otherwise, the comparison between the peptides with difference sequences was resulted in that the hydrophobic residues that located adjacent to Trp residue is more affected to the mobility of Trp side chain, that the amplitude of the tortional vibrations are smaller and the rotational isomerization times larger than in the coil (*7*). In our UVRR studies, it is difficult to distinguish between the peptide studies and apoMb intermediate. There are two possible reasons. One is positive interpretation that the W3 band of Trp on the peptide and intermediate are same because of the strong interaction with Val17. Figure 3-4.4 shows the X-ray crystallographic structure (*18*) of the swMb at the region of A helix. In the native form, Trp14 has strong hydrogen bonding with Glu18 and hydrophobic interactions with Val17. Although the structure of Trp14 in the

intermediate is changed and its W3 frequency is different from native form, it seems that Trp 14 is interacting with Val17 in the intermediate. Actually, the intensity ratio of W7 band that is sensitive to the polar/nonpolar environment suggested that the environment of Trp side chain is slightly hydrophobic in the peptide than in the free Trp solutions.

Another reason for the difference between UVRR spectroscopy and fluorescence anisotropy study of apoMb intermediate is that the static vibrational spectroscopy reflects the total fluctuations and distributions of conformations as a band in terms of inhomogeneous and homogeneous contributions (*19*). The fact that W3 band of the unfolded state or free Trp solutions has a broad band than native band indicated that the amplitude of fluctuations and conformational numbers are increased than native one. The band shift of W3 band indicated the structural change of Trp by the interaction with surrounding residues but it is difficult to separate the contributions of the amplitude of fluctuations from that of static conformational distributions. Furthermore, since model peptide could not be folded 100 %, it is difficult to determine the bandwidth of W3 in the α helix formation of the model peptide to compare with apoMb intermediate determined in section 3-2.

It is clear, however, that the changes of side chain interaction with other residues or structures are reflected sensitively by the band shape and frequency of W3 band, and it will be advantageous to use this band for analysis the fast dynamics of proteins.

Reference

1. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H.,

Woodruff, W. H., and Dyer, R. B. (1996) *Biochemistry 35*, 691-7.

2. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892-6.
3. Akiyama, S., Takahashi, S., Ishimori, K., and Morishima, I. (2000) *Nat Struct Biol 7*, 514-20.
4. Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y., and Roder, H. (2001) *Nat Struct Biol 8*, 151-5.
5. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., and Socci, N. D. (1995) *Proc Natl Acad Sci U S A 92*, 3626-30.
6. Privalov, P. L., and Makhatadze, G. I. (1993) *J Mol Biol 232*, 660-79.
7. Semisotnov, G. V., Zikherman, K. K., Kasatkin, S. B., Ptitsyn, O. B., and Anufrieva, E. V. (1981) *Biopolymers 20*, 2287-2309.
8. Tcherkasskaya, O., Ptitsyn, O. B., and Knutson, J. R. (2000) *Biochemistry 39*, 1879-89.
9. Kemple, M. D., Buckley, P., Yuan, P., and Prendergast, F. G. (1997) *Biochemistry 36*, 1678-88.
10. Edelhoch, H. (1967) *Biochemistry 6*, 1948-54.
11. Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972) *Biochemistry 11*, 4120-31.
12. Aki, M., Ogura, T., Shinzawa-Itoh, K., Yoshikawa, S., and Kitagawa, T. (2000) *J. Phys. Chem. B 104*, 10765-10774.
13. Haruta, N., Aki, M., Ozaki, S., Watanabe, Y., and Kitagawa, T. (2001) *Biocemistry* 6956-6963.
14. Sabelko, J., Ervin, J., and Gruebele, M. (1998) *J. Rhys. Chem. B 102*, 1806-1819.
15. Reymond, M. T., Merutka, G., Dyson, H. J., and Wright, P. E. (1997) *Protein Sci 6*, 706-16.

16. Harada, I., Miura, T., and Takeuchi, H. (1986) *Spectrochimica Acta 42a*, 307-312.

17. Nishimoto, E., Yamashita, S., Szabo, A. G., and Imoto, T. (1998) *Biochemistry 37*, 5599-5607.

18. Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R. M., and Schlichting, I. (1999) *Biophys J 77*, 2153-74.

19. Mukamel, S. (1995) *Principle of nonlinear optical spectroscopy*, Oxford University Press, New York.

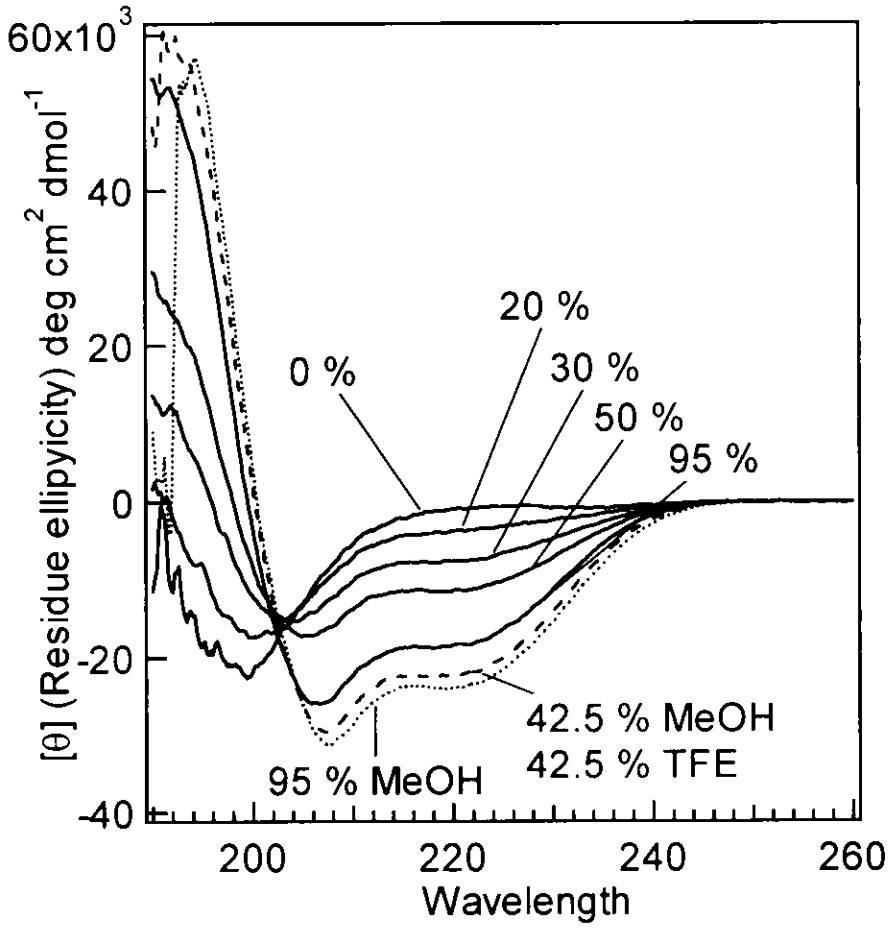


Figure 3-4.1. CD spectra of A peptide were measured in the various concentrations of TFE and Methanol.

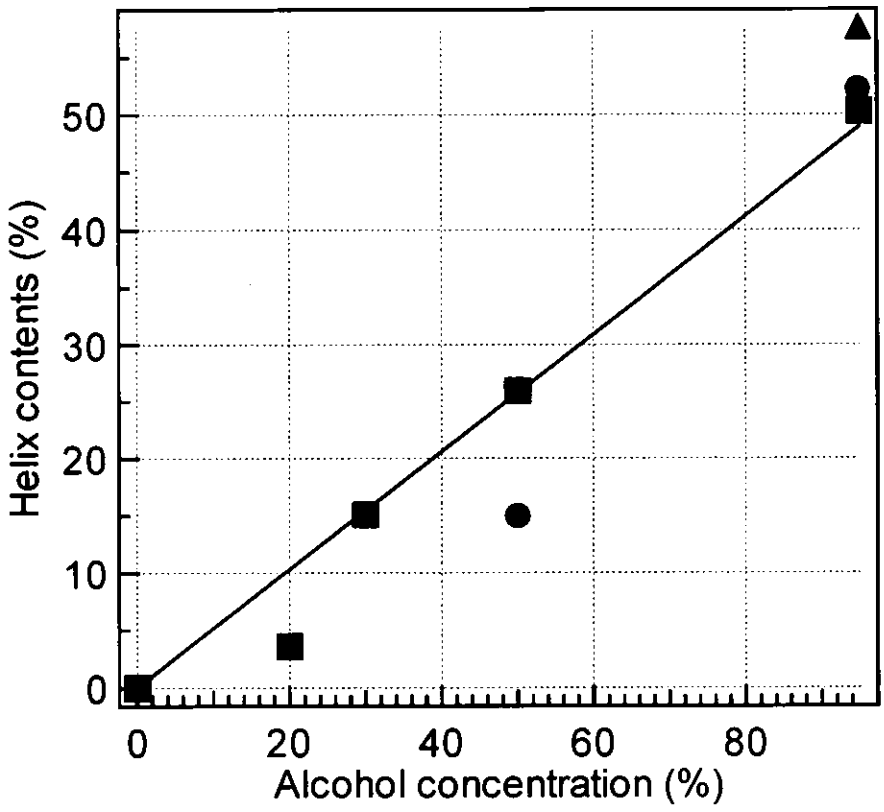


Figure 3-4.2. The α helix contents depend on the concentration of alcohol calculated by the ellipticity at 222 nm; TFE (■), MeOH (●) and mixture of TFE and MeOH (▲).

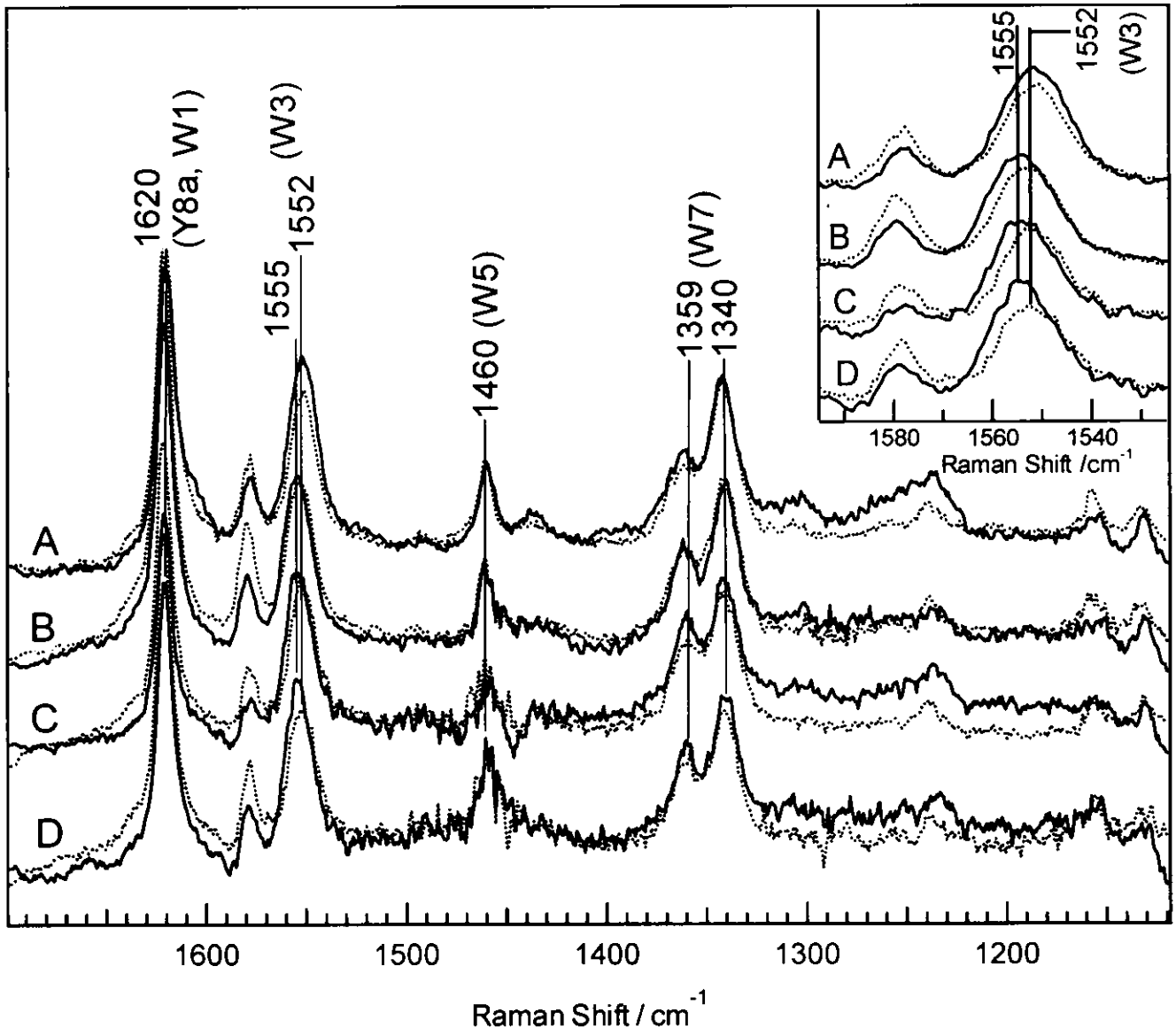


Figure 3-4.3. 244-nm excited UVRR spectra of free Trp solution (broken line) and A peptide (solid line) in various solvents: A, water; B, 70 % v/v TFE; C, 75 % v/v MeOH; D, 32.5 % TFE and 32.5 % MeOH. The inset shows the expanded spectra around W3 band.

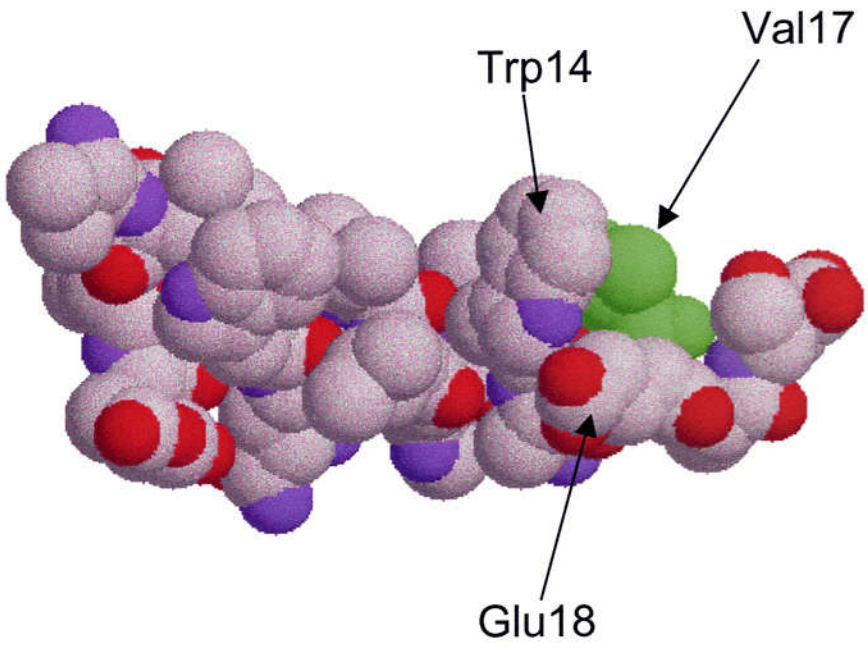


Figure 3-4.4. The structure in A helix region of sperm whale Mb *(14)* drawn by the spacefill styles.

LIST OF PUBLICATION

1. Nami Haruta, Michihiko Aki, Shin-ichi Ozaki, Yoshihito Watanabe and Teizo Kitagawa, "Protein conformation change of myoglobin upon ligand binding probed by ultraviolet resonance Raman spectroscopy" *Biochemistry* **40**, 6956-6963 (2001)

2. Takeshi Tomita, Nami Haruta, Michihiko Aki, Teizo Kitagawa and Masao Ikeda-Saito, "UV resonance Raman detection of a ligand vibration on ferric nitosyl heme proteins" *J. Am. Chem. Soc.* **123**, 2666-2667 (2001)

3. Nami Haruta and Teizo Kitagawa, "UV time-resolved resonance Raman investigation of protein folding: characterization of kinetic and equilibrium intermediate of apomyoglobin" to be submitted