Two-Dimensional Infrared Correlation Spectroscopy Study of Sequential Events in the Heat-Induced Unfolding and Aggregation Process of Myoglobin

Yong-Bin Yan,*† Qi Wang,‡ Hua-Wei He,* Xin-Yao Hu,‡ Ri-Qing Zhang,*† and Hai-Meng Zhou§
*NMR Laboratory, Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China; †State Key Laboratory of Biomembrane and Membrane Biotechnology, Tsinghua University, Beijing 100084, China; ‡Department of Chemistry, Tsinghua University, Beijing 100084, China; and §Protein Science Laboratory of the Ministry of Education, Tsinghua University, Beijing 100084, China

ABSTRACT Unfolding and aggregation are basic problems in protein science with serious biotechnological and medical implications. Probing the sequential events occurring during the unfolding and aggregation process and the relationship between unfolding and aggregation is of particular interest. In this study, two-dimensional infrared (2D IR) correlation spectroscopy was used to study the sequential events and starting temperature dependence of Myoglobin (Mb) thermal transitions. Though a two-state model could be obtained from traditional 1D IR spectra, subtle noncooperative conformational changes were observed at low temperatures. Formation of aggregation was observed at a temperature (50–58°C) that protein was dominated by native structures and accompanied with unfolding of native helical structures when a traditional thermal denaturation condition was used. The time course NMR study of Mb incubated at 55°C for 45 h confirmed that an irreversible aggregation process existed. Aggregation was also observed before fully unfolding of the Mb native structure when a relative high starting temperature was used. These findings demonstrated that 2D IR correlation spectroscopy is a powerful tool to study protein aggregation and the protein aggregation process observed depends on the different environmental conditions used.

INTRODUCTION

In the study of protein folding, the occurrence of aggregation has traditionally been considered an unproductive off-pathway reaction. However, the aggregation of proteins has important technical implications in disease treatment and biotechnology. Studying protein aggregation in vivo is important to a range of diseases that arise from protein misfolding or improper folding including Alzheimer's disease, Parkinson's disease, Huntington's disease, prion encephalopathies, and sickle-cell anemia disease (Johnson, 2000; Hoffner and Djian, 2002; Kopito, 2000; Noguchi and Schechter, 1985), whereas studying aggregation in vitro is crucial to protein refolding procedures in protein engineering (Kopito, 2000). Much effort has been devoted in recent years to elucidating the mechanism of protein aggregation and factors that affected the progress of aggregation. Electrostatic interactions and the exposure of hydrophobic sections are two basic mechanisms important for protein aggregation (Bruinsma and Pincus, 1996), although new observations and new mechanism have been proposed by many groups (Ren et al., 2001; Srisailam et al., 2002). In general, aggregation must involve formation of specific intermolecular contacts. Both theoretical and experimental studies have addressed the aggregation process directly. The conclusion is that either the fully unfolded proteins or the partially unfolded intermediates were supposed to be responsible for aggregation. For example, De Young et al. found the primary aggregation species were denatured protein molecules when measuring the solubility of apomyoglobin (apoMb) in aqueous urea solution (De Young et al., 1993). Tsai et al. (1998) found that the occurrence of Ribonuclease A thermal aggregation in physiological pH was detected when the protein was present predominantly in the folded configuration. The relationship between aggregation and the folding/unfolding state of proteins is not clear yet. This relationship is important to elucidate the mechanism and prevention of protein aggregation, although sequential events involved are at the heart of the protein aggregation process. However, probes used are not always sensitive enough to characterize the related sequential events in protein unfolding and the aggregation process. It is important to select a sensitive tool that can detect the minor noncooperative events and closely related changes.

For the thermal aggregation model, differential scanning calorimetry (DSC) is a useful tool to study the thermal denaturation, tertiary structure, and thermodynamic properties of proteins (Pace et al., 1989). However, DSC only brings out the simple two-state mechanism of heat-induced transitions (Lyubarev et al., 1999) and yields no detailed information about the molecular structure and dynamics of proteins. The protein aggregation process is also difficult to be followed by NMR or standard optical techniques including circular dichroism (CD) and fluorescence because of inhomogeneous or light scattering induced by the aggregates. In contrast, IR spectroscopy is insensitive to light scattering and thus provides a valuable method for studying protein aggregation (Dong et al., 2000). Fourier transform infrared (FT-IR) spectroscopy has been widely

Submitted December 20, 2002, and accepted for publication May 19, 2003.

Yong-Bin Yan and Qi Wang contributed equally to this work.

Address reprint requests to Dr. Yong-Bin Yan, NMR Laboratory, Dept. of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China. Tel.: +86-10-6278-3477; Fax: +86-10-6277-1597; E-mail: ybyan@mail.tsinghua.edu.cn.

0006-3495/03/09/1959/09 $2.00

used to study the secondary structure and dynamics of polypeptides and proteins in solution (Gerwert, 2000). However, a shortcoming of general IR spectroscopy is the highly overlapping of different secondary component bands. Combined with computerized Fourier transform instrumentation and powerful mathematical resolution-enhancing techniques including Fourier deconvolution, second derivative analysis, and two-dimensional correlation spectroscopy, different secondary structures components of proteins can be distinguished and quantitatively determined by FT-IR (Dong et al., 2000). In recent years, two-dimensional infrared (2D IR) correlation spectroscopy proposed by Noda (1989, 1990) has been applied extensively to analyze IR spectra of proteins (for example, Fabian et al., 1999; Murayama et al., 2001; Paquet et al., 2001). The basic idea of 2D IR has been thoroughly described by many authors (Fabian et al., 1999; Noda, 1989, 1990; Murayama et al., 2001; Paquet et al., 2001). Briefly, time-dependent spectral variations, which are called dynamic spectra, can be induced by certain external perturbation including mechanical, thermal, chemical, electrical, or acoustic stimulations. With the application of a correlation analysis to spectral fluctuations, new types of spectra defined by two independent spectral variable axes are obtained. By spreading peaks along the second dimension, the spectral resolution is enhanced, and the order of the actual sequence of processes induced by the perturbation can be revealed. Protein unfolding and aggregation were usually induced by specific environmental changes, thus 2D IR provides a useful tool to investigate the sequential events occurring in the unfolding and aggregation processes of proteins. In this article, sequential events occurring in the heat-induced denaturation and aggregation of the all α-helix protein, Myoglobin (Mb), were investigated by 2D IR. Different thermal conditions were chosen to characterize the relationship between the folding/unfolding states and initiating structures formed during aggregation.

MATERIALS AND METHODS

Sample preparation

Horse heart Myoglobin was purchased from Sigma Chemical (St. Louis, MO) and used without further purification. The protein was dissolved in D_2O and stored overnight at 35°C to ensure complete H/^{2}H exchange of all solvent accessible protons. This solution was then lyophilized. The lyophilized protein was redissolved into D_2O at a concentration of 50 mg/ml (p^{2}H 5.4) to form homogeneous solution. The p^{2}H is the uncorrected (for D_2O) pH-meter reading at room temperature.

Infrared measurements

Infrared spectra were measured at a spectral resolution of 4 cm^{-1} in a single-beam mode with a Perkin-Elmer Spectrum 2000 spectrometer equipped with a dTGS detector. Protein samples were prepared in a temperature cell with CaF_2 windows separated by a 50-μm Teflon spacer, which was placed in a heating jacket controlled by a digital temperature controller. The spectra of protein solutions were collected in the temperature range of 30–80°C, 60–70°C, or 70–80°C at intervals of 2°C every 20 min. The temperature stability was better than ±0.2°C. For each measurement, 256 scans were recorded to ensure a good signal/noise ratio. Reference spectra were recorded under identical scan conditions with only the corresponding solvent (D_2O) in the cell. The protein spectra were acquired by subtracting the reference spectra of D_2O from the spectra of protein solutions at the same temperature. To ensure the reliability of the experiments, all experiments were repeated 2–3 times.

The thermal denaturation curve was obtained by measuring the intensity at 1651 cm^{-1} and analyzed with a nonlinear least-squares algorithm. The data analysis was also carried out by the spectral image analysis method established recently (Yan et al., 2001). The spectrum parameters (Shannon entropy, H, and Correlation coefficient, C) that describe the nature of each spectrum and the correlation between different spectra were calculated using MATLAB software (MATLAB 5.2, The MathWorks, Natick, MA) by programs developed in-house. The spectral region used for the analysis was 1600–1700 cm^{-1}. The thermal denaturation curves were represented directly by these parameters. Curve fit was obtained by the standard method using the Marqurdt-Levenberg routine (Jackson and Fersht, 1991) as provided in the Origin 6.0 software (Microcal, Northampton, MA).

2D correlation analysis

Synchronous and asynchronous correlation intensities were computed for the IR spectra at different temperatures by applying the generalized 2D correlation algorithm of Noda (1989, 1990). Before 2D calculation, each spectrum was smoothed and baseline corrected. Software used for calculation was developed in-house according to the generalized 2D correlation algorithm based on Hilbert transform. Rules used for the analysis of the sign of the peaks in the 2D IR plots were followed by those proposed by Noda (1990).

NMR measurements

The NMR sample was prepared by dissolving 25 mg protein in 500 μl 90% H_2O/10% D_2O (v/v) at a pH of 5.2. No insoluble aggregates were found after a 10-min 6000-g centrifugation. The NMR sample was stabilized at room temperature for 4 h and was inserted into the magnet preequilibrated to a temperature of 30 or 55°C. The first 1D ^{1}H spectrum was obtained after 15 min equilibrium. For thermal transition studies, spectra were recorded every 10 min at 2.5°C temperature intervals. For time course studies, spectra were continuously collected every hour during 45 h incubation at 55°C. All NMR experiments were performed on a Varian Unity Inova 500NB NMR spectrometer at Tsinghua University. A total of 16 transients were collected with a 20 ppm spectra width. The improved WATERGATE pulse sequence (Liu et al., 1998) was used for water suppression in all experiments. All data were processed and analyzed by using the VNMR software provided by Varian (Palo Alto, CA).

RESULTS AND DISCUSSION

Myoglobin is one of the most thoroughly studied proteins with respect to structure, function, and folding. The secondary structure of Mb is mainly composed of α-helices (~77%, including 3_{10} helices), random coil (~13%), and β-turns (~10%) (Evans and Brayer, 1990). The IR spectra of native holomyoglobin (hMb), native apoMb, and various acid-destabilized forms of apoMb have been thoroughly investigated by many groups (for example, Dong et al., 2000; Gilmanshin et al., 1997, 2001). Assignment of the different secondary structure components used in this re-

search was referenced to prepublished works (Byler and Susi, 1986; Dong et al., 1990, 2000; Gilmanshin et al., 1997, 2001; Haris and Chapman, 1995).

Mb unfolding from traditional 1D IR spectroscopy

Fig. 1 shows the changes in the amide I band upon heating. At physiological temperatures, the amide I band is dominated by the α-helix band at 1651 cm^{-1} (Fig. 1 *A*). As temperature increased, the intensity of this component gradually decreased, and the formation of two new bands at 1681 and 1618 cm^{-1} is observed. According to the classical idea, the heat-induced unfolding of most proteins at high concentrations can be described with a two-state model, followed by the loss of native secondary structure and accompanied aggregation. The protein aggregation is reflected by the IR spectroscopy in the formation and growth

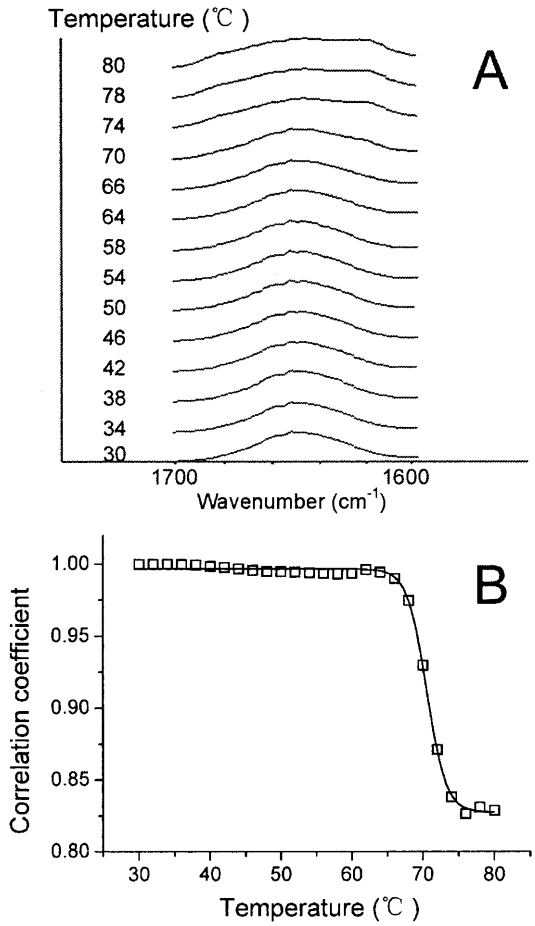

FIGURE 1 The original infrared spectra (*A*) and the thermodynamic curve (*B*) of Mb in D_2O measured at a temperature range from 30 to 80°C. At each temperature, the spectrum of the solvent has been subtracted from the corresponding spectrum of the Mb solution. The thermodynamic curve was represented by correlation coefficient from image analysis. (Yan et al., 2001, 2002a).

of two well-defined bands at ~1618 and ~1681 cm^{-1}. These two bands have been previously assigned to hydrogen-bonded extended intermolecular β-sheet structures formed upon aggregation of thermally denatured proteins (Damaschun et al., 2000).

The thermal denaturation curve could be obtained by measuring the intensity at 1651 cm^{-1}. But error was introduced for a baseline shift found at high temperatures. Thus the data analysis was carried out by the spectral image analysis method established recently (Yan et al., 2001). The result from spectral image analysis can reflect all changes of the spectra at every wavenumber instead of one wavenumber from traditional methods, and thus can be used to detect whether the change at different wavenumber is a cooperative process. The spectrum parameters (*H* and *C*) that describe the nature of each spectrum and the correlation between different spectra were calculated. In agreement with the theory established before (Yan et al., 2002a), a peak function was found when the values of *H* are plotted versus temperature with a midpoint temperature at ~71°C (data not shown). The result from image parameter *C* was shown in Fig. 1 *B* and was quite consistent with those previous observations upon aggregation. A sigmoid curve is found with a midpoint temperature, $T_{1/2}$, at ~70.6 ± 0.1°C, which is quite consistent with previous work (Meersman et al., 2002). All the results from 1D IR spectra show a standard two-state thermal transition model and no intermediate was found.

As mentioned above, thermal unfolding of proteins is always assumed to be a fully cooperative event and shows a two-state process without any intermediate observed (Fig. 1 *B*). If so, different probes used to characterize the unfolding process should give identical unfolding curves (Jackson and Fersht, 1991). However, folding intermediate, pretransition state or inhomogeneous folding was observed at least for some proteins during thermal transitions (Carrotta et al., 2001; Dong et al., 2000; Fabian et al., 1999; Paquet et al., 2001; Gilmanshin et al., 1997; Yan et al., 2002b). Thus, what is needed to study protein thermal transition is to probe the closely related events and noncooperative events. To address this issue, 2D correlation spectra were calculated at intervals of 10°C in this study to investigate details of sequential events occurring in Mb thermal unfolding.

Thermal transition of Mb at a different temperature range from 2D IR correlation plots

Fig. 2 shows the synchronous (*A*) and asynchronous (*B*) 2D IR correlation plots of Mb thermal transitions covering the temperature interval 30–38°C with a calculated spectral region of 1700–1600 cm^{-1}. The synchronous spectrum was dominated by a prominent autopeak at 1646 cm^{-1}, which is mainly due to the random coil structure in the amide I band. Since the synchronous 2D IR correlation plot recognizes the similarity between the variations of spectral intensities to a perturbation (in our case, temperature), the appearance of

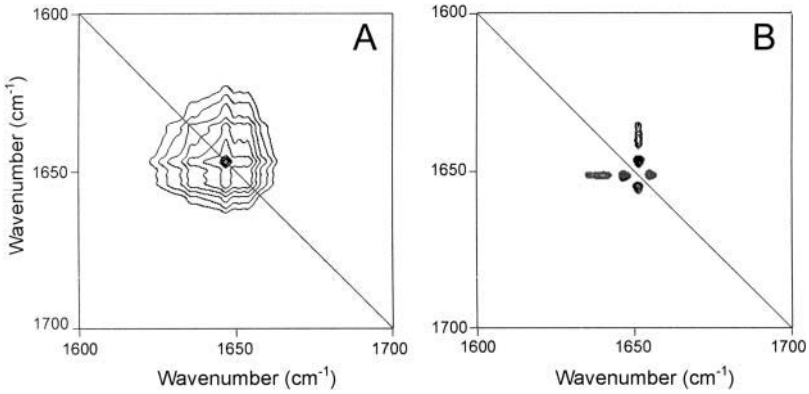

FIGURE 2 Synchronous (*A*) and asynchronous (*B*) 2D IR correlation maps constructed from 1D IR spectra recorded at 30, 32, 34, 36, and 38°C during Mb thermal transition from 30 to 80°C. Clear and dark peaks are positive and negative, respectively.

this autopeak indicates that random coil is more sensitive to heating, and thermal variations of secondary structures of Myoglobin might start from them. New information can be obtained from the analysis of the asynchronous correlation map shown in Fig. 2 *B*. Generally, the asynchronous 2D IR correlation plot detects the difference between the variations of spectral intensities and thus is of particular interest because it can be used to distinguish different components and changes that occurred out of phase. In Fig. 2 *B*, three characteristics asynchronous crosspeaks could be identified, located at 1651–1646 cm^{-1}, 1655–1651 cm^{-1}, and 1651–1636 cm^{-1}, respectively. The appearance of asynchronous crosspeaks always indicates changes at a different rate as a function of applied perturbation (in our case, temperature) between the two corresponding bands during a specific process. In Fig. 2 *B*, a band at 1636 cm^{-1}, which should be assigned to the hydrogen bonded extended chain (Byler and Susi, 1986; Dong et al., 1990; Haris and Chapman, 1995), was asynchronously correlated with the band of α-helix at 1651 cm^{-1}. Additionally, important asynchronous crosspeak was observed between 1651 cm^{-1} and 1655 cm^{-1}, which were both due to the helical structure. The appearance of this crosspeak suggested that, when heated, the eight α-helices of Mb might have different vibration patterns and exhibit dissimilar characters upon perturbation. However, the absence of an autopeak related to α-helices indicated that no significant changes occurred in the helical structure of Mb.

During the temperature interval from 40 to 48°C, no obvious spectral changes were observed in the traditional 1D IR spectra (Fig. 1). Fig. 3 shows the synchronous (*A*) and asynchronous (*B*) 2D IR plots of the Mb thermal transitions during this temperature range. Besides the peak at 1646 cm^{-1} representing random coil, a clear autopeak was observed at 1651 cm^{-1} along the diagonal line, corresponding to the α-helices conformation. The intensity of autopeaks always represents the overall extent of dynamic fluctuations of spectral intensity observed at a specific wavenumber. Thus, the appearance of this peak suggested that the unfolding of α-helix started during this temperature range. Additionally, synchronous crosspeaks have developed between 1651 and 1646 cm^{-1}, 1651 and 1643 cm^{-1}, and 1646 and 1643 cm^{-1}. A new band identified by the synchronous map at 1643 cm^{-1} is also due to random coil. The corresponding asynchronous spectrum in Fig. 3 *B* shows a number of asynchronous crosspeaks, located at 1652–1624 cm^{-1}, 1652–1630 cm^{-1}, 1652–1638 cm^{-1}, 1652–1647 cm^{-1}, 1657–1652 cm^{-1}, and 1663–1652 cm^{-1}. In addition to the band at 1636 cm^{-1}, the asynchronous map reveals two new bands located at 1624 and 1630 cm^{-1}, which are generally assigned to β-sheet. However, the x-ray analysis suggested that the Mb crystal structure has no β-sheet (Evans and Brayer, 1990); these three bands should all be assigned to the hydrogen bonded extended chains that connect the helical cylinders (Byler and Susi, 1986). Comparison between the spectra in Figs. 2 and 3

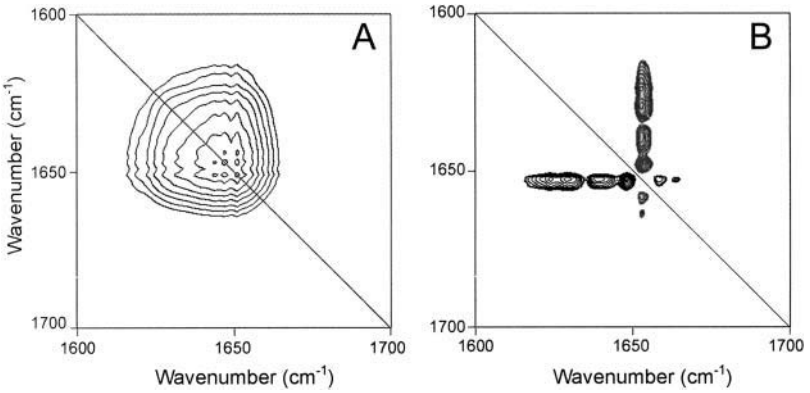

FIGURE 3 Synchronous (*A*) and asynchronous (*B*) 2D IR correlation maps constructed from 1D IR spectra recorded at 40, 42, 44, 46, and 48°C during Mb thermal transition from 30 to 80°C. Clear and dark peaks are positive and negative, respectively.

indicated that more absorption bands related to extended chain are revealed. The signs of the crosspeaks in Fig. 3 indicated that the change of 1652 cm^{-1} was after all other events. Combined analysis of the appearance of the crosspeaks and the sign of these peaks in Fig. 3 suggested that the whole protein might be much looser than the native globular state, whereas the helical structures were changed after the exposure to solvents.

From the asynchronous plot constructed from 40 to 48°C, different vibration patterns of the eight α-helix, at least three bands located at 1652 cm^{-1}, 1656 cm^{-1}, and 1663 cm^{-1}, could be clearly distinguished. The former two bands were also assigned to α-helix as identified in Fig. 2, whereas the band at 1663 cm^{-1} should be assigned to 3_{10} helix (Byler and Susi, 1986; Dong et al., 1990, 2000; Haris and Chapman, 1995). Thus according to different responses to certain perturbation, the broad absorption band of α-helices in amide band I can be decomposed into at least three components in this study. This suggested that even the different changes of the same secondary structure could also be distinguished by 2D IR correlation spectroscopy. Dong et al. (2000) has found that an intermediate rich in 3_{10} helix was observed during Mb thermal unfolding in the presence of 1.0 M guanidine hydrochloride (GdnHCl), although no such intermediate was found during Mb thermal unfolding in the absence of GdnHCl. It is a question whether the band appearing at 1663 cm^{-1} is from the native Mb structure or from the 3_{10}-helix-rich intermediate. In our case, no thermal stable intermediate could be identified observed from both 1D and 2D IR spectra in this study, and the formation of the 3_{10}-helix-rich intermediate of Mb in GdnHCl was at a relatively higher temperature (Dong et al., 2000). Thus the band at 1663 cm^{-1} in Fig. 3 *B* should be assigned to the native Mb structure.

It had been suggested that no changes of Mb thermal transition could be observed under 70°C by traditional 1D IR spectra (Meersman et al., 2002; and in this article). In this study, it is quite interesting that the unfolding of α-helices is detectable in the temperature range of 40–50°C by 2D IR spectroscopy, although this temperature is far before the midpoint of the intensity-temperature curve. This suggested that although the overall intensity of the band at 1651 cm^{-1} had not decreased dramatically, different components of the helical conformation might change asynchronously due to different molecular environments and different responses to the perturbation. Thus subtle changes of secondary structure transitions can be successfully revealed in the asynchronous spectrum, which shows the potential of wide applications of 2D IR correlation analysis.

The synchronous and asynchronous plots of Mb thermal transition between 50 and 58°C are shown in Fig. 4. Four main auto peaks located at 1632 cm^{-1}, 1645 cm^{-1}, 1651 cm^{-1}, and 1656 cm^{-1} were observed in the synchronous map (Fig. 4 *A*) along the main diagonal line. These amide I components could be assigned to hydrogen bonded extended chains that connect the helical cylinders (1632 cm^{-1}), random coil (1645 cm^{-1}), and α-helix (1651 cm^{-1} and 1656 cm^{-1}), respectively. At the same time, synchronous crosspeaks develop between these absorption bands. The asynchronous map (Fig. 4 *B*) was characterized by several prominent crosspeaks located at 1632–1617 cm^{-1}, 1644–1617 cm^{-1}, 1651–1617 cm^{-1}, and 1656–1617 cm^{-1}. Weak crosspeaks were also observed at 1684–1647 cm^{-1} and 1684–1656 cm^{-1}. Significant new bands appeared at 1617 and 1684 cm^{-1}, which are characteristic for intermolecular antiparallel β-sheet and have been identified as a sign of protein aggregation (Damaschun et al., 2000). The sign of the crosspeaks in the asynchronous plot reveals that the intensity decrease of absorption bands corresponding to α-helices and random coil is somewhat ahead of the intensity increase of β-sheet, which suggested that the formation of β-sheet is after the unfolding of helical structures (Noda, 1990). Although the intensity increasement at 1617 and 1681 cm^{-1} is not so remarkable when observed in the traditional 1D spectra (Fig. 1), the occurrence of Mb aggregation could be clearly distinguished in the 2D IR asynchronous map.

Sequential events occurred in the Mb unfolding process

Remarkable inhomogeneous unfolding of Mb could be observed when comparing the 2D IR correlation plots constructed upon different temperature ranges as shown in

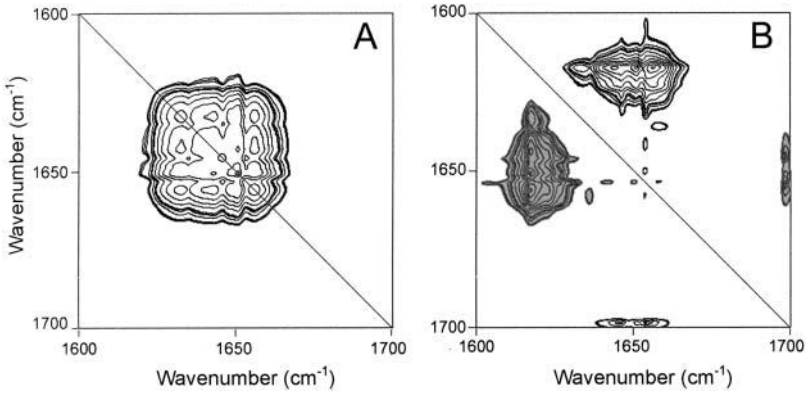

FIGURE 4 Synchronous (*A*) and asynchronous (*B*) 2D IR correlation maps constructed from IR spectra recorded at 50, 52, 54, 56, and 58°C during Mb thermal transition from 30 to 80°C. Clear and dark peaks are positive and negative, respectively.

Figs. 2–4. The sequential events occurred in the inhomogeneous unfolding could be characterized as: 1), Change of random coil was firstly observed upon perturbation at low temperatures (30–38°C; Fig. 2 *A*), which suggested that the irregular structure in Mb was quite sensitive to temperature perturbation and might be the starting point of Mb unfolding. The appearance of bands at 1630 and 1624 cm^{-1} (assigned to extended chains that connect the helical cylinders; Fig. 3 *B*) also suggested that the starting of Mb unfolding was accompanied with looser globule structures. 2), Different bands of α-helix could be clearly distinguished by 2D IR correlation plots (Figs. 2 *B*, 3, and 4), which suggested that the eight α-helix of Mb had different vibration patterns and exhibited dissimilar characteristics upon perturbation. This result also suggested that different vibration patterns that existed might cause the possibility of intermediate(s) observed in chemical or thermal-chemical unfolding of Mb (for example, Dong et al., 2000; and the well-known acid unfolding of Mb). 3), The band at 1663 cm^{-1} (assigned to 3_{10} helix) observed in Fig. 3 *B* disappeared in Fig. 4 *B*, which suggested that the unfolding of 3_{10} helix might start at the temperature range of 40–48°C and finish at the temperature range of 50–58°C. Mb is mainly composed of α-helix, whereas the 3_{10}-helix structures exist in A-, C-, E-, and G-helix. No other evidence of thermal stable unfolding intermediate(s) could be obtained during Mb thermal transition in physiological conditions (Dong et al., 2000; Meersman et al., 2002; and in this article). However, the result in this article indicated that at least pretransition state might exist. This pretransition might include the unfolding of 3_{10} helix and some other α-helix structures. 4), The appearance of the intermolecular β-sheet was at a quite low temperature (between 50 and 58°C). A much higher aggregated temperature (~70°C) was observed by Meersman et al. (2002). This difference might be caused by insensitivity of traditional 1D IR spectroscopy to subtle changes at the initial stages of protein thermal unfolding. The fourth conclusion was quite consisted to some previous studies using different techniques (Dong et al., 2000; Gilmanshin et al., 1997; Tsai et al., 1998), whereas the former three results were first characterized by the 2D IR method. Additionally, usually 2D IR correlation plots are constructed upon a relative wide perturbation range and only the remarkable transitions are observable. Thus these four results also suggested that details, especially the subtle changes, of protein unfolding could be clearly characterized by 2D IR correlation plots constructed at small-step perturbations.

Temperature dependence of the Mb aggregation process

Though the thermal transition of Mb has been thoroughly characterized, a question remained concerning the relationship of protein unfolded state and aggregation. It has been observed by some authors and in this study that thermal aggregation is accompanied with unfolding at high temperatures above 55°C (De Young et al., 1993; Dong et al., 2000; Meersman et al., 2002; Paquet et al., 2001). But several authors also suggested that aggregation could occur at nearly native or partially folded state of proteins (Tsai et al., 1998; Carrotta et al., 2001). The conclusion that assembly must involve formation of specific intermolecular contacts and some nonnative conformation is required to prime the process is intuitive (Carrotta et al., 2001), but how aggregation primed and extended is not clear yet. A basic question is whether a fully unfolded state is necessary for protein aggregation. Thermal transition studies of proteins were always carried out using step-by-step temperature increase. This method helped us demonstrate sequential events occurred, but also raised a problem that a pretransitioned state was introduced into latter aggregation. When aggregation occurs, partially unfolded or even fully unfolded proteins already existed in the solutions. To avoid this problem, starting temperature (60 or 70°C) dependence of the Mb aggregation process was studied in this article.

Fig. 5 shows the synchronous (*A*) and asynchronous (*B*) plots calculated from the spectra started at 60°C and ended at 70°C at intervals of 2°C every 20 min. The plots were dominated by crosspeaks located at 1651–1615 cm^{-1}, which indicated that changes concerning the α-helix and intermolecular β-sheet structure were the main process. In Fig. 5 *A*, the negative sign of the crosspeak at 1651–1615

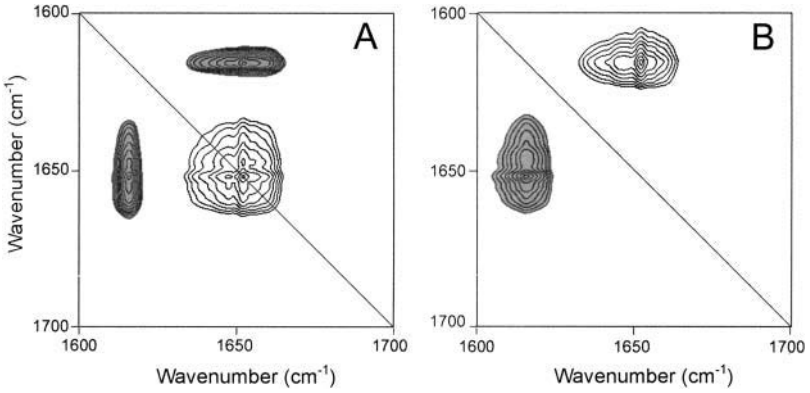

FIGURE 5 Synchronous (*A*) and asynchronous (*B*) 2D IR correlation maps constructed from 1D IR spectra recorded at 60, 62, 64, 66, and 68°C during Mb thermal transition from 60 to 68°C. Clear and dark peaks are positive and negative, respectively.

cm^{-1} indicated that the intensity of these two absorption bands was changing in the opposite direction, which means the unfolding of α-helix and formation of intermolecular β-sheet. At the same time, a positive asynchronous crosspeak develops between 1651 and 1615 cm^{-1}. Judging from the sign of the crosspeaks appearing in synchronous and asynchronous maps using rules proposed by Noda (1990), the formation of β-sheet is somewhat ahead of the unfolding of α-helices. Similar result could be obtained from the analysis of 2D IR correlation plots constructed from a starting temperature of 70°C (Fig. 6). The synchronous map (Fig. 6 *A*) was dominated by a negative crosspeak at 1651–1687 cm^{-1}, whereas a weak positive crosspeak (data not shown) at the same wavenumbers was observed in the asynchronous map. The main crosspeak developed between 1651 and 1615 cm^{-1} in the asynchronous map was quite weak (positive; data not shown) in the synchronous map. The observations in Fig. 6 also suggested that the formation of β-sheet was ahead of the unfolding of α-helices. It is interesting that the results in Figs. 5 and 6 indicated that the formation of Mb aggregation was faster than the unfolding of the native helical structure. In other words, aggregation could form between native or at least partially folded structures (Fig. 1) under certain conditions.

Time course study of Mb aggregation by NMR at 55°C

The sequential events of Mb thermal unfolding and aggregation were determined thoroughly in the current study. The most remarkable feature was that the formation of aggregates occurred at a relative low temperature when the protein is dominated by native structure. The 1D IR spectra (Fig. 1) and 1D NMR spectra (data not shown) clearly show that only a minimal amount of aggregates formed at this low temperature range and could hardly be observed by normal techniques. To obtain additional detailed information regarding the time evolution of aggregation at a certain temperature, one-dimensional NMR measurements were taken at 55°C continuously for 45 h (Fig. 7). A well-known constraint on solution NMR is the line width, $\nu_{1/2}$, of a certain peak is inversely proportional to T_2, which is the time constant for spin-spin relaxation (Cavanagh et al., 1996). T_2 decreases with increasing molecular size and increases with increasing flexibility. Thus the line width can reflect the corresponded molecular size changes. In our case, when aggregation occurred, the large aggregates will have a smaller T_2 and a wider line width. Line width has been successfully used to monitor the aggregation process by many groups (for example, Alexandrescu and Rathgeb-Szabo, 1999; Yan et al., 2002b). Fig. 7 shows the stack plot of the typical part of the ^{1}H NMR spectra for Mb heated at 55°C continuously for 2 days. Significant line broadening could be observed over time, indicating that protein association occurred.

Besides the line width changes, it is interesting that the disappearing of some native peaks and the appearance of nonnative peaks, which was indicated by arrows in Fig. 7, could also be observed. However, in agreement with the 1D IR experiments, denatured peaks could not be observed in the continuously heated NMR experiments at 55 ~ 60°C (data not shown). Since aggregation is a competing process with the unfolding or intermediate formation process, the appearance of denatured peaks in the time course study at this low temperature suggested that an irreversible aggregation process existed and more and more molecules in the sample became part of the aggregate. Thus, these NMR data provide direct evidence that Mb aggregates could form at 55°C and these aggregates were derived from denatured proteins.

CONCLUSIONS

The results in this article indicate that when Mb suffered thermal perturbation step by step from native conditions to unfolding/aggregation conditions, aggregation was observed accompanied by unfolding of Mb helical structures. It is interesting that aggregation could form at a relative low temperature between 50 and 58°C, at which temperature the protein was dominated by native structure. The time course NMR studies further confirmed that irreversible aggregation could form at 55°C. When a starting temperature higher than the aggregation occurrence temperature but below the melting temperature, T_m, (in our case, 60 and 70°C) was

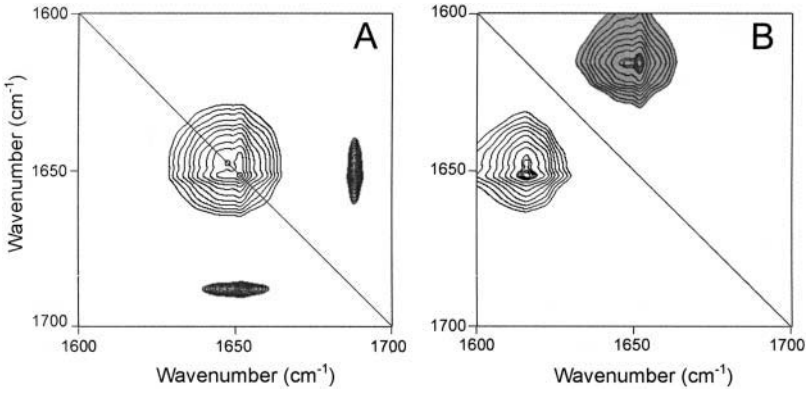

FIGURE 6 Synchronous (*A*) and asynchronous (*B*) 2D IR correlation maps constructed from 1D IR spectra recorded at 70, 72, 74, 76, 78, and 80°C during Mb thermal transition from 70 to 80°C. Clear and dark peaks are positive and negative, respectively. Weak positive crosspeaks located at 1651–1615 cm^{-1} in the synchronous map and weak positive crosspeak located at 1651–1687 cm^{-1} in the asynchronous map were not shown.

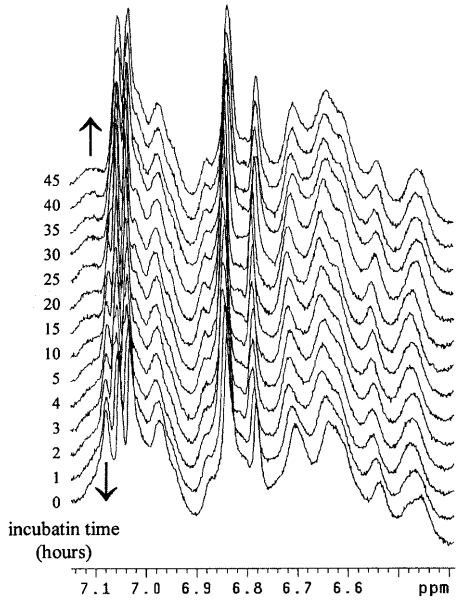

FIGURE 7 Stack plot of ^{1}H NMR spectra recorded continuously at different time points for 50 mg/mL Mb incubated at 55°C for 45 h. Only the typical part (part of the aromatic resonance region) of the whole spectra was shown for clarity. The up and down arrows indicated the appearance of nonnative peaks and the disappearance of native peaks, respectively.

used, aggregation was formed before the fully unfolding of the protein. The results herein suggested that the aggregation process could be affected by environmental conditions. Different processes might be observed when different conditions were used. In other words, aggregation is a competing process with the unfolding or intermediate formation process. When an aggregation-favored condition is used, protein aggregation is the main process observed (Tsai et al., 1998; Damaschun et al., 2000; Meersman et al., 2002; and in this article), whereas when an unfolding-favored condition is used, protein unfolding (including the formation of unfolding intermediate) will dominate the transitions (Carrotta et al., 2001; Damaschun et al., 2000; De Young et al., 1993; Dong et al., 2000; Paquet et al., 2001).

The authors thank Dr. Qiong Wu and Dr. Fan-Guo Meng (Department of Biological Sciences and Biotechnology, Tsinghua University) for fruitful discussion. The authors also thank the anonymous reviewers for helpful suggestions.

This investigation was partially supported by the Basic Research Funds of Tsinghua University, P. R. China (JC2002047), the 985 Funds of Tsinghua University, P. R. China, and funds from State Key Laboratory of Biomembrane and Membrane Biotechnology, P. R. China.

REFERENCES

Alexandrescu, A. T., and K. Rathgeb-Szabo. 1999. An NMR investigation of solution aggregation reactions preceding the misassembly of acid-denatured cold shock protein A into fibrils. *J. Mol. Biol.* 291:1191–1206.

Bruinsma, R., and P. Pincus. 1996. Protein aggregation in membranes. *Current Opinion in Solid State & Materials Science*. 1:401–406.

Byler, D. M., and H. Susi. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers*. 25:469–487.

Carrotta, R., R. Bauer, R. Waninge, and C. Rischel. 2001. Conformational characterization of oligomeric intermediates and aggregates in β-lactoglobulin heat aggregation. *Protein Sci.* 10:1312–1318.

Cavanagh, J., W. J. Fairbrother, A. G. Palmer, and N. J. Skelton. 1996. Protein NMR Spectroscopy. Academic Press, San Diego.

Damaschun, G., H. Damaschun, H. Fabian, K. Gast, R. Kröber, M. Wieske, and D. Zirwer. 2000. Conversion of yeast phosphoglycerate kinase into amyloid-like structure. *Proteins*. 39:204–211.

De Young, L. R., K. Dill, and A. L. Fink. 1993. Aggregation and denaturation of apomyoglobin in aqueous urea solutions. *Biochemistry*. 32:3877–3886.

Dong, A., P. Huang, and W. S. Caughey. 1990. Protein secondary structure in water from second-derivative amide I infrared spectra. *Biochemistry*. 29:3303–3308.

Dong, A., T. W. Randolph, and J. F. Carpenter. 2000. Entrapping intermediates of thermal aggregation in α-helical proteins with low concentration of guanidine hydrochloride. *J. Biol. Chem.* 275:27689–27693.

Evans, S. V., and G. D. Brayer. 1990. High-resolution study of the three-dimensional structure of horse heart metmyoglobin. *J. Mol. Biol.* 213:885–897.

Fabian, H., H. H. Mantsch, and C. P. Schultz. 1999. Two dimensional IR correlation spectroscopy: sequential events in the unfolding process of the λ Cro-V55C repressor protein. *Proc. Natl. Acad. Sci. USA*. 96: 13153–13158.

Gerwert, K. 2000. Infrared and Raman Spectroscopy of Biological Materials. Marcel Dekker, New York.

Gilmanshin, R., M. Gulotta, R. B. Dyer, and R. H. Callender. 2001. Structures of apomyoglobin's various acid-destabilized forms. *Biochemistry*. 40:5127–5136.

Gilmanshin, R. S., R. H. Williams, R. H. Callender, W. H. Woodruff, and R. B. Dyer. 1997. Fast events in protein folding: relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl. Acad. Sci. USA*. 94:3709–3713.

Haris, P. I., and D. Chapman. 1995. The conformational analysis of peptides using Fourier transform IR spectroscopy. *Biopolymers*. 37:251–263.

Hoffner, G., and P. Djian. 2002. Protein aggregation in Huntington's disease. *Biochimie*. 84:273–278.

Jackson, S. E., and A. R. Fersht. 1991. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry*. 30:10428–10435.

Johnson, W. G. 2000. Late-onset neurodegenerative diseases—the role of protein insolubility. *J. Anat.* 196:609–616.

Kopito, R. R. 2000. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* 10:524–530.

Liu, M., X. Mao, C. Ye, H. Huang, J. K. Nicholson, and J. C. Lindon. 1998. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *J. Magn. Reson.* 132:125–129.

Lyubarev, A. E., B. I. Kurganov, V. N. Orlov, and H.-M. Zhou. 1999. Two-state irreversible thermal denaturation of muscle creatine kinase. *Biophys. Chem.* 79:199–204.

Meersman, F., L. Smeller, and K. Heremans. 2002. Comparative Fourier transform infrared spectroscopy study of cold-, pressure-, and heat-induced unfolding and aggregation of Myoglobin. *Biophys. J.* 82:2635–2644.

Murayama, K., Y. Wu, B. Czarnik-Matusewicz, and Y. Ozaki. 2001. Two-dimensional/attenuated total reflection infrared correlation spectroscopy studies on secondary structural changes in human serum albumin in aqueous solutions: pH-dependent structural changes in the secondary structures and in the hydrogen bonding of side chains. *J. Phys. Chem. B*. 105:4763–4769.

Noda, I. 1989. Two-dimensional infrared spectroscopy. *J. Am. Chem. Soc.* 111:8116–8118.

Noda, I. 1990. Two-dimensional infrared (2D-IR) spectroscopy: theory and applications. *Appl. Spectrosc.* 44:550–561.

Noguchi, C. T., and A. N. Schechter. 1985. Sickle hemoglobin polymerization in solution and in cells. *Annu. Rev. Biophys. Chem.* 14:239–263.

Pace, N. C., B. A. Shirley, and J. A. Thomson. 1989. Measuring the conformational stability of a protein. *In* Protein Structure, a Practical Approach. T. E. Creighton, editor. IRL Press, New York. 311–330

Paquet, M.-J., M. Laviolette, M. Pezolet, and M. Auger. 2001. Two-dimensional infrared correlation spectroscopy study of the aggregation of cytochrome c in the presence of dimyristoylphosphatidylglycerol. *Biophys. J.* 81:305–312.

Ren, H., Y. Nagai, T. Tucker, W. J. Strittmatter, and J. R. Burker. 2001. Amino acid sequence requirements of peptides that inhibit polyglutamine-protein aggregation and cell death. *Biochem. Biophys. Res. Comm.* 288:703–710.

Srisailam, S., T. K. S. Kumar, T. Srimathi, and C. Yu. 2002. Influence of backbone conformation on protein aggregation. *J. Am. Chem. Soc.* 124:1884–1888.

Tsai, A. M., J. H. van Zanten, and M. J. Betenbaugh. 1998. Study of protein aggregation due to heat denaturation: a structural approach using circular dichroism spectroscopy, nuclear magnetic resonance, and static light scattering. *Biotechnol. Bioeng.* 59:273–280.

Yan, Y.-B., B. Jiang, R.-Q. Zhang, and H.-M. Zhou. 2001. Two-phase unfolding pathway of ribonuclease A during denaturation induced by dithiothreitol. *Protein Sci.* 10:321–328.

Yan, Y.-B., X.-C. Luo, H.-M. Zhou, and R.-Q. Zhang. 2002a. Two-state kinetics characterized by image analysis of nuclear magnetic resonance spectra. *Chinese Sci. Bull.* 47:389–393.

Yan, Y.-B., R.-Q. Zhang, and H.-M. Zhou. 2002b. Biphasic reductive unfolding of ribonuclease A is temperature dependent. *Eur. J. Biochem.* 269:5314–5322.