Nativelike Secondary Structure in Interleukin-1β Inclusion Bodies by Attenuated Total Reflectance FTIR†

Keith Oberg,‡ Boris A. Chrunyk,§ Ronald Wetzel,§ and Anthony L. Fink*,‡

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, and Macromolecular Sciences Department, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Received September 9, 1993; Revised Manuscript Received December 2, 1993®

ABSTRACT: Attenuated total reflectance FTIR has been used to study the structure of human interleukin-1β in inclusion bodies (IBs) and other aggregated forms. The secondary structure composition of native wild-type IL-1β determined by FTIR is in excellent agreement with that previously determined by crystallography and NMR: 52% β-sheet, 25% loop/irregular structure, and 23% turn. Remarkably, IL-1β inclusion bodies exhibit secondary structural composition very similar to that of the native protein. The results indicate that the IBs form from a folding intermediate that has nativelike secondary structure. The secondary structure content of aggregated IL-1β, formed either in refolding or by thermal denaturation, was identical within experimental error to that of the IB, indicating that these aggregates were formed from intermediates with structures similar to that of the inclusion body.

Inclusion bodies (IBs)[1] are densely packed particles of protein frequently found when recombinant genes are expressed in bacteria (Marston, 1986; Schein, 1989; Wetzel, 1992a,b). Strong denaturants or detergents are required for their solubilization; in some cases, reduction of disulfide bonds is also necessary. In many cases they are highly enriched in a single protein. Their prevalence, and the difficulty in renaturing protein from them, has made inclusion bodies a major problem in biotechnology. Despite the fundamental and practical importance of inclusion bodies, their structures and mechanisms of formation are little understood. These structures are seen as refractile bodies in phase-contrast microscopy and as dense, amorphous particles in electron micrographs (Williams et al., 1982; Wetzel & Goeddel, 1983). No evidence has previously been published for any regular structure in inclusion bodies, and the prevailing view is that IB formation is an aggregation process mediated by nonspecific interactions. However, the observation of dramatic mutational effects on IB formation in a number of systems (King et al., 1990; Krueger et al., 1990; Fierke et al., 1991; Truong et al., 1991; Wetzel et al., 1991; Rinas et al., 1991; Wetzel & Chrunyk, 1992) is more reminiscent of the high degree of sequence specificity observed in the folding stability of proteins. Three possible mechanisms for their formation have been proposed: aggregation of native protein of limited solubility; aggregation of the unfolded state; and aggregation of partially folded intermediate states. Unfortunately, detailed structural studies of IBs have not been performed due to the scarcity of techniques that can be applied to nontransparent, solid samples.

When expressed in *E. coli* at about 5–10% total cellular protein, wild-type human interleukin-1β accumulates almost exclusively (>90%) in the soluble fraction of the native cell lysate. However, point mutants of the protein have been identified which, although produced at about the same level as wild type, are deposited >50% into inclusion bodies (Chrunyk et al., 1993). Stability studies on a number of these purified mutants suggest that the mutations do not influence IB formation *via* effects on the thermal or thermodynamic stabilities of the native molecules and, hence, might operate *in vivo* at the level of folding intermediates. Particularly strong evidence for this is found for the mutant Lys97Val (K97V), which is produced mostly in the form of IBs, even though it is more thermodynamically stable than the wild-type protein (Chrunyk et al., 1993; Chrunyk & Wetzel, 1992). In addition, K97V displays folding-related aggregation *in vitro* under conditions in which the wild type remains soluble (Chrunyk & Wetzel, 1993a,b). The availability of inclusion body mutants, detailed information on their *in vitro* stabilities and aggregation properties, and high-resolution X-ray (Finzel et al., 1989; Priestle et al., 1989) and NMR structural studies on the wild-type protein (Driscoll et al., 1990; Clore et al., 1991) make IL-1β a particularly attractive subject for structural studies on its IBs and other aggregated states.

Knowledge of the secondary structure of IL-1β in inclusion bodies provides an important step toward understanding the mechanisms of inclusion body formation. To collect FTIR spectra of the highly scattering aggregates, ATR-FTIR was used because of its broad range of sampling capabilities. With the ATR method, a spectrum can be taken of nearly any sample in contact with the surface of the internal reflection element (IRE), regardless of transparency. With this approach, we have examined the structure of three different aggregated forms: inclusion bodies, aggregates formed during thermal denaturation, and aggregates formed during refolding from denaturant. H_2O was chosen as the solvent for this study in order to eliminate potential problems caused by slow exchange of D for H in the inclusion bodies, allowing valid comparison of the spectra of IB (formed in an aqueous environment) and native structure by avoiding isotopic shifts in the IR bands.

† Supported in part by a grant from the National Science Foundation (A.L.F.).

* Author to whom correspondence should be addressed.
‡ University of California, Santa Cruz.
§ SmithKline Beecham Pharmaceuticals.

® Abstract published in *Advance ACS Abstracts,* February 1, 1994.

[1] Abbreviations: IL-1β, interleukin-1β; I/L, irregular/loop structure; ATR, attenuated total reflectance; IRE, internal reflection element used as an ATR sample holder; IB, inclusion body; WT, wild type; FSD, Fourier self-deconvolution.

0006-2960/94/0433-2628$04.50/0

MATERIALS AND METHODS

Production of IL-1β. The plasmid containing the wild-type IL-1β cDNA engineered for expression in *Escherichia coli* has been described previously, along with procedures for cell transformation, growth, and IL-1β induction (Myers et al., 1987). Expression under control of the λ P_L promoter was induced by incubating transformed *E. coli* strain AR58 cells at 42 °C. Cells were recovered by centrifugation, frozen, and processed as will be described.

Inclusion Body Isolation. Frozen cells (45 g) were suspended in 150 mL of cell suspension buffer (20 mM Tris and 0.25 mM EDTA, pH 8.0) containing 1 mM PMSF, and then 0.5 mg of T4 lysozyme (Perry & Wetzel, 1986) was added and the suspension sonicated at 0 °C for a total of 10 1-min pulses (50% duty cycle). The lysed cells were centrifuged for 30 min at 10 000 rpm (Sorval GS3). The insoluble fractions were resuspended in another 150 mL of cell suspension buffer containing 1 mM PMSF and 0.25 mg of T4 lysozyme and then sonicated and centrifuged as before. The pellet was washed in 250 mL of cell suspension buffer containing 0.1% Triton X-100 and centrifuged again to remove nonspecifically adsorbed proteins. Freeze, thaw, lysis, and washing steps were repeated until cell lysis was complete, as judged by the disappearance of a yellow film on top of the Triton X-100 pellet. The pellet was then resuspended in 15 mL of the cell suspension buffer containing 0.1% Triton X-100 and centrifuged at 5500 rpm for 30 min. The supernatant was carefully poured off and the pellet resuspended in 15 mL of fresh buffer, and the centrifugation was repeated. The suspension of inclusion bodies was stored at 4 °C in the cell suspension buffer containing Triton X-100. SDS–PAGE indicated that >90% of the Coomassie blue stained protein in the inclusion bodies is IL-1β.

Sample Preparation for FTIR. Protein was prepared for ATR-FTIR spectroscopy as follows: ammonium sulfate precipitates of native IL-1β were dissolved in 50 mM sodium phosphate buffer (pH 6.5) and dialyzed against the same buffer. Initial solutions were 10–15 mg/mL protein. It was found that both mutant and WT IL-1β undergo a slight, time-dependent conformational change when stored in buffer. After 2 weeks of storage at 4 °C, amide I bands for the native protein were broader than those seen from solutions scanned within 48 h after resuspension of the ammonium sulfate precipitates.[2] All data reported herein, with the exception of those of inclusion bodies, are for freshly prepared samples.

Inclusion bodies and folding and thermal aggregates were washed several times with water or buffer in an Eppendorf microcentrifuge before use. Insoluble thermal aggregates were formed by incubating 10 mg/mL protein solution in sodium phosphate buffer at 52 °C for 60 min. Insoluble refolding aggregates were generated by incubating 100 μL of 6 mg/mL native K97V in a solution containing 3 M Gdn·HCl, 10 mM MES, 2 mM EDTA, and 10 mM DTT (pH 6.5) overnight at room temperature. The resulting solution of unfolded protein was diluted into 5 vol of 10 mM MES, 90 mM NaCl, 2 mM EDTA, and 10 mM DTT (pH 6.5) to give a final guanidine concentration of 0.5 M, and incubated for 1 h at 42 °C to induce refolding and aggregation. The pelleted folding aggregate was dried on the IRE as soon as possible after washing (3× with 0.6 mL of pure water) to minimize time-dependent changes that began when the denaturant solution was removed (to be described). In the spectra and analysis results presented here, 5–6 min elapsed from the first wash until drying was complete. Approximately 25% of the precipitate was lost in the washes.

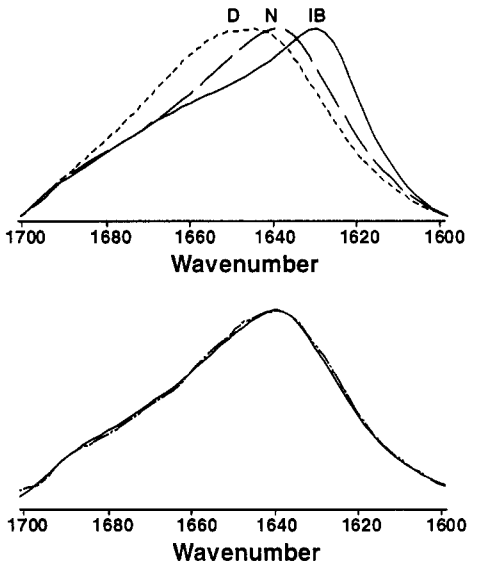

FIGURE 1: Comparison of the amide I region of the FTIR spectrum for interleukin-1β K97V in the native, inclusion body, and acid-denatured forms, using the thin-film ATR mode (see text) (top). D represents the denatured protein (pH 2), N the native (pH 7), and IB the inclusion body. Comparison of the interleukin-1β K97V amide I spectra for native protein in a hydrated thin film (dried) on a germanium IRE (solid line) and in solution (broken line) (bottom). Buffer and water vapor contributions have been subtracted as described in Materials and Methods.

FTIR Spectroscopy. Spectra were collected with a Nicolet 800 FTIR spectrometer equipped with an MCT detector, and purged with dry nitrogen. Spectra at 4-cm^{-1} resolution were obtained by coadding 3000 interferograms for each sample. All samples were scanned in an out-of-compartment horizontal ATR accessory (SPECAC) with a high-throughput 73 × 10 × 6 mm, 45° trapezoidal germanium crystal (IRE). The crystal holder was configured as a flow cell for solution spectra. Dry samples were collected with the IRE uncovered. To collect spectra of dried proteins, the IRE was cleaned with soap and ethanol before use. Samples were applied to the IRE as 15–40 μL of 1–10 mg/mL solution (25–400 μg of protein) and dried while being spread constantly with a spatula. The drying process took 3–10 min at ambient temperature. Spectra of buffer alone were taken for the purpose of subtracting the residual water signal from dry samples.

It is known that protein in solution can bind to the surface of the IRE. To remove any signal due to adsorbed protein, a "flush" sample was prepared after each solution sample by gently rinsing the ATR flow cell with buffer. Flush spectra were used in place of pure buffer spectra for the purpose of liquid water subtraction. This subtraction method provided only the signal from protein in the bulk phase. Spectra of thin films were used for analysis because of their larger S/N. To verify that the drying process did not disrupt IL-1β conformation, solution spectra of the native protein (12 mg/mL) and the IB (as a paste) were collected. In both cases, solution and dry (thin-film) spectra were the same within experimental error (Figure 1) and gave identical analysis results.

FTIR Spectral Analysis. All data processing was done with Lab Calc (Galactic Industries). For solution samples, flush and buffer spectra were ratioed against the spectrum of

[2] Native preparations of 2-week-old solutions gave curve-fitting results as follows: WT, 56% β-sheet, 23% I/L, and 21% turn; and K97V, 49% β-sheet, 29% I/L, and 22% turn. Thus, both wild-type and mutant protein showed an increase of about 4% β-sheet as a result of the aging.

a clean IRE or assembled flow cell. Flush spectra were subtracted from protein spectra using a program based on the algorithm described by Powell et al. (1986) (base-line correction was unnecessary). For thin films, subtraction of the residual water spectrum was necessary for accurate analysis. Scaling factors for water subtraction were typically between 0.2 and 0.7, indicating a substantial contribution from water. The water vapor signal was carefully subtracted as described by Prestrelski et al. (1991).

Three to five independent spectra of each sample type were collected and analyzed individually; errors reported are the standard deviations of all analyses of each sample type. Band positions were determined using second derivatives and Fourier self-deconvolution (FSD). Savitsky–Golay second derivatives were taken with 5–11 convolution points. Excellent agreement between second derivatives and FSD peak positions was observed at FSD parameters $\gamma = 3$ and $f = 0.40$–0.45, where γ is the band narrowing factor, and f is a high-pass filter (Griffiths & Pariente, 1986).

Curve fitting was performed with a modified version of Lab Calc's CURVEFIT.AB program, which allowed curve-fit parameters to be monitored during, and altered between, the stages of the fitting process. Initially, the FSD of a spectrum was fit to generate the best possible starting parameters for the original data. Next, peak widths were increased, and the parameters used for fitting the original spectrum with peak centers fixed and height, width, and % Lorenzian were allowed to vary. Finally, all parameters were freed for 200 iterations. Peak-center drift was always minimal in the final stage of fitting. Band assignments were made on the basis of the studies of Byler and Susi (1986), Surewicz and Mantsch (1988), Dong and Caughey (1990), Wilder et al. (1992), and Prestrelski et al. (1991). In all cases, excellent agreement between the sum of the curve-fit spectra and the experimental spectrum was observed (Figure 3).

RESULTS AND DISCUSSION

Structure of Native Wild-Type IL-1β. X-ray and NMR analyses show that IL-1β consists of 12 strands of β-sheet linked by eight turns and three large loops that connect three similar subdomains (βββLβ), each consisting of two pairs of double-stranded sheets. Six of the β-strands (one pair from each subdomain) form a β-barrel in the center of the protein. The native molecule thus has pseudo-3-fold symmetry (Clore et al., 1991; Finzel et al., 1989). The IR spectrum of native IL-1β is dominated by bands corresponding to β-sheet/extended structure at 1638 and 1624 cm^{-1} (Figures 1–3). The band at 1657 cm^{-1} was assigned to immobile irregular/loop structure: several published studies of proteins with large immobile irregular loops and no α-helix[3] have also reported bands near 1656 cm^{-1} (Dukor et al., 1992; Wilder et al., 1992; Prestrelski et al., 1991). On the basis of curve-fitting analysis of ATR spectra, we find the secondary structure of native wild-type IL-1β to be 52% β-sheet, 23% turn, and 25% irregular/loop (Table 1). This excellent agreement in the secondary structural composition of native human IL-1β determined by both solution-phase and thin-film ATR-FTIR with that determined by X-ray crystallography and NMR (Clore & Gronenborn, 1991) validates the ATR procedures used in the present study. ATR thin films have previously been shown to provide accurate data regarding native protein structure (Goormaghtigh et al., 1990; Swedberg et al., 1990). In addition, the secondary structure of human IL-1β determined here by ATR agrees well with that of murine IL-1β determined by transmission FTIR in D_2O (Wilder et al., 1992) (Table 1).

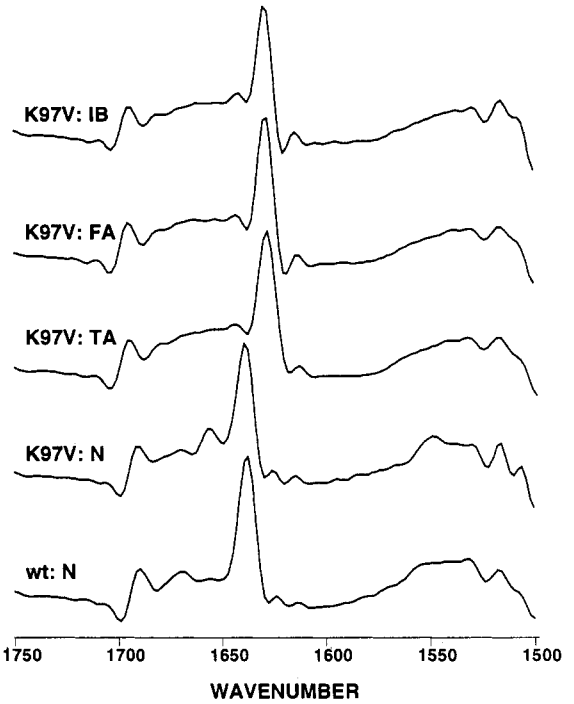

FIGURE 2: Comparison of the interleukin-1β amide I (1600–1700 cm^{-1}) and II (1500–1600 cm^{-1}) bands from ATR-FTIR after Fourier self-deconvolution (FSD) resolution enhancement. IB is the inclusion body, FA is folding aggregate, TA is thermal aggregate, K97V:N is native K97V, and wt:N is native wild type. The major peak (1625 or 1637 cm^{-1}) corresponds to β-sheet or extended structure.

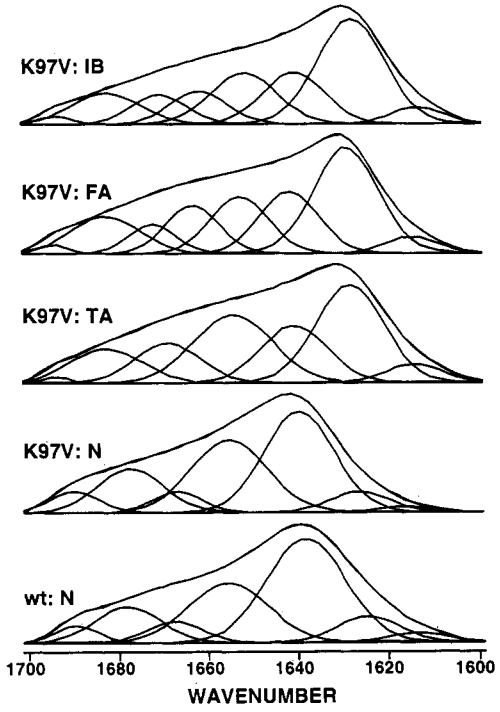

FIGURE 3: Interleukin-1β amide I curve-fit spectra from ATR-FTIR spectra showing the individual amide I components. IB is inclusion body, FA is folding aggregate, TA is thermal aggregate, and N is native. Reconstructed spectra are superimposed on original data; the fits were such that it is difficult to distinguish the two.

[3] In the native structure of IL-1β, there is one turn of 3_{10}-helix in the loop between β-strands 3 and 4, around residue 36.

Table 1: Secondary Structural Composition of Interleukin-1β[a]

	inclusion body K97V	folding aggregate K97V	thermal aggregate K97V	native human K97V	native human WT	murine WT (D_2O) (Wilder et al.)	X-ray human WT
I/L	25.7 ± 1.4	27.9 ± 0.9	26.2 ± 2.7	31.1 ± 1.2	25.4 ± 2.0	20.6	24
extended	51.8 ± 0.4	51.2 ± 1.6	49.9 ± 1.4	45.3 ± 1.8	51.9 ± 1.6	59.2	56
turn	22.6 ± 1.2	20.9 ± 0.8	23.9 ± 2.1	23.7 ± 1.8	22.8 ± 0.9	19.7	20

[a] Based on analysis of amide I spectra (see text). I/L refers to irregular/loop structure (see text).

Structure of Native Human K97V IL-1β. A comparison of the FTIR spectral properties of the native states of wild-type and the K97V mutant is summarized in Tables 1 and 2. Although the FTIR spectra are similar, there are some significant differences: in the mutant, 6–7% of the β-sheet/extended structure is replaced with irregular/loop structure. This suggests that the mutation, at the junction of a loop and a β-strand, leads to some unraveling of the adjacent β-sheet. The overall structure of the protein is not adversely affected by the mutation, however, as indicated by the overall similarity of band positions, the areas of bands in the turn region, and the similar thermodynamic stabilities of the mutant and wild-type IL-1β. The amide II and III regions of the native wild-type and K97V proteins are highly similar, also indicative of similar overall structures (Figures 2 and 4). We therefore conclude that the substitution of valine for lysine at position 97 most likely leads to unraveling of β-strand 8, the strand from residues 100 to 106, which would increase the flexibility of the loop linking topological units B and C (Clore et al., 1991; Finzel et al., 1989).

Structure of the Inclusion Bodies. The curve-fit spectrum for the K97V inclusion bodies is shown in Figure 3, and the secondary structural content is given in Table 1. The amounts of β-sheet, turn, and irregular/loop for the IB are identical to those in the native wild-type protein! This recovery of WT nativelike secondary structure upon IB formation indicates that the protein in the IB is in a highly organized state, with nativelike secondary structure. Because of this, we have chosen to interpret our data from the viewpoint that the overall secondary structure of protein in the inclusion bodies and other aggregates (see below) is nativelike and that variations in FTIR signals reflect subtle changes in secondary structure. The main differences then, between native and aggregated IL-1β, must result from the nature of the interactions and packing of the secondary structural units.

It is clear that there are some significant differences in the actual three-dimensional structure of the native proteins and aggregates, as manifested by changes in amide I component band areas and positions (Figure 3 and Table 2). Although the *amount* of irregular/loop structure is the same in both native and IB states, the single, wide band at 1656 cm^{-1} in the native state is replaced by two bands at 1651 and 1660 cm^{-1} in the IB. We have tentatively assigned both bands to irregular/loop structure. This split suggests that at least one of the loop regions is different in the IB and the native protein (presumably that involving residue 97). In addition, although the total area of the two β-bands is the same in the WT and the mutant, the major β-band in the IB is at 1628 cm^{-1}, corresponding to 32% of the amide I area, whereas this band is the minor one (6% or 9%) in the native state. Since a decrease in frequency of amide I components arises from shorter hydrogen bonds (Susi, 1969), the β-sheet is packed more tightly (favorably) in the IB than in the native protein. Evidence for slight differences in the overall three-dimensional arrangement of the molecule in the IB compared to the native state is also seen in the amide II and III regions (see below). The amide I region of the FTIR spectrum for the acid-denatured form of IL-1β is shown in Figure 1, along with those of the native and IB. Clearly, the spectrum of the denatured state is significantly different from that of the inclusion body. This experiment serves as a control to eliminate the possibility that the IRE affects the conformational state or spectrum of the protein.

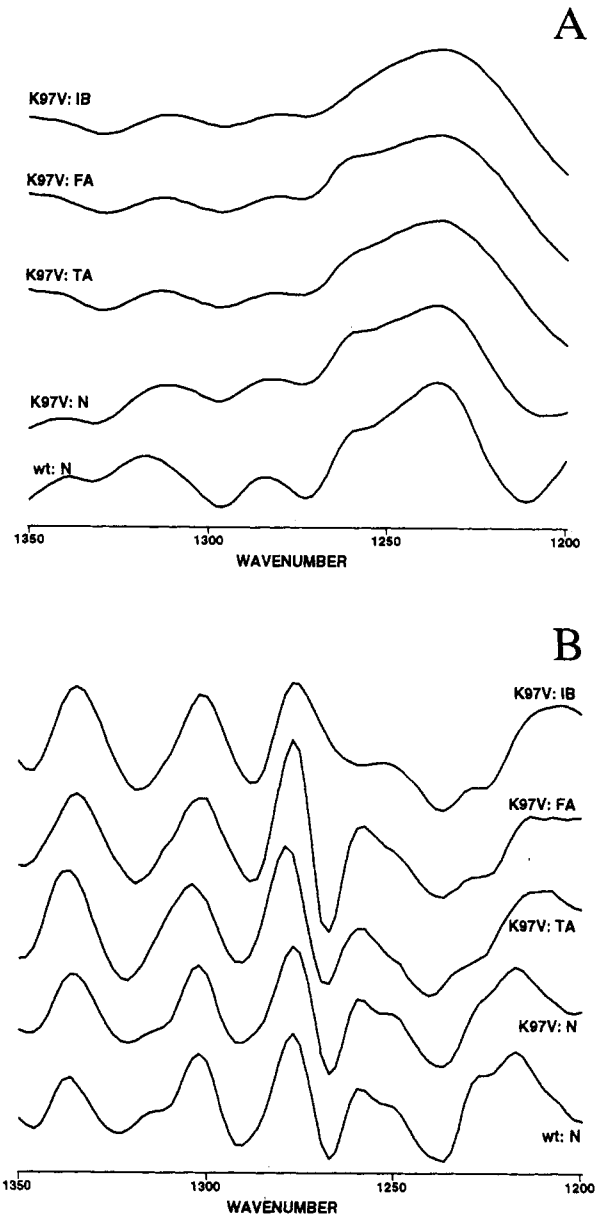

FIGURE 4: Interleukin-1β amide III spectra (a) and second derivative spectra (B) from ATR. IB is inclusion body, FA is folding aggregate, TA is thermal aggregate, and N is native. The band in the vicinity of 1262 cm^{-1} is due to a buffer component.

Structure of the Refolding Aggregate. When the K97V mutant is refolded from 3 M Gdn·HCl at 42 °C, it aggregates (Chrunyk & Wetzel, 1993a; B. A. Chrunyle and R. Wetzel, unpublished results); this behavior is not observed with the wild-type protein. Therefore, it may be concluded that the

Table 2: Amide I Band Positions and Relative Areas for Interleukin-1β[a]

secondary structure type	inclusion body	folding aggregate	thermal aggregate	native K97V	native WT
turn	1693.3 ± 0.2	1694.1 ± 0.1	1693.5 ± 0.1	1689.8 ± 0.2	1689.7 ± 0.1
	1.4 ± 0.2	1.1 ± 0.1	1.1 ± 0.2	4.7 ± 0.6	3.9 ± 0.1
turn	1681.1 ± 0.6	1682.7 ± 0.3	1681.9 ± 0.4	1678.2 ± 0.9	1678.1 ± 0.4
	13.1 ± 0.5	12.5 ± 0.2	11.4 ± 0.2	13.2 ± 1.6	12.4 ± 1.2
turn	1669.5 ± 0.7	1671.8 ± 0.3	1668.5 ± 0.4	1667.5 ± 1.0	1667.7 ± 0.1
	8.1 ± 1.1	7.3 ± 0.6	10.4 ± 1.6	5.8 ± 0.2	6.5 ± 1.9
I/L	1660.5 ± 0.7	1662.7 ± 0.2			
	11.1 ± 0.9	12.9 ± 1.4			
I/L	1651.4 ± 0.8	1652.4 ± 0.1	1654.0 ± 0.1	1655.8 ± 0.5	1655.6 ± 0.4
	14.5 ± 1.4	15.0 ± 2.1	25.1 ± 2.6	31.1 ± 1.0	25.4 ± 2.0
extended	1640.9 ± 0.8	1641.5 ± 0.3	1640.9 ± 0.5	1639.3 ± 0.2	1638.2 ± 0.3
	19.3 ± 0.9	18.4 ± 1.1	14.1 ± 2.6	39.3 ± 1.4	43.2 ± 3.1
extended	1628.5 ± 0.7	1628.6 ± 0.2	1628.0 ± 0.1	1626.0 ± 0.7	1624.3 ± 0.5
	32.4 ± 0.7	32.9 ± 0.7	33.6 ± 2.5	6.0 ± 0.6	8.7 ± 1.7

[a] Peak positions were first determined using FSD resolution enhancement, and areas were determined from the curve-fit procedure, as described in the text. The upper number is the band center in cm^{-1}, and the lower number is the relative area (%) of that component.

mutation interferes in some way with the folding process of IL-1β. The spectra of the IB and the folding aggregate are identical, within experimental error (Figures 2–4, Table 2). Thus, folding of the mutant results in an aggregate structurally similar to the IB, suggesting that both IB and folding aggregate are formed *via* the same pathway (i.e., the same folding intermediate).

Although the spectra of the IB and the refolding aggregate are strikingly similar, a time-dependent change was observed with the latter if the precipitated refolding aggregate was left in contact with the water used to wash it free of denaturant. This change was manifested by the loss of the well-resolved 1652-cm^{-1} band within 15–30 min, followed by a substantial increase in the intensity of the region centered around 1660 cm^{-1} between 30 and 150 min. No further changes were observed after 150 min; the final structure contained approximately 25% β-sheet, 21% I/L, and 54% turn. The cause of this phenomenon is unclear.

Structure of the Thermal Aggregate. Partial thermal denaturation of both WT and K97V IL-1β results in aggregation (Chrunyk et al., 1993; Chrunyk & Wetzel 1993a,b). Both WT and K97V thermal aggregates had very similar spectra; only data for the mutant will be considered here. The data (Figures 2–4 and Table 2) indicate that unfolding is not extensive and that these aggregates have a high degree of nativelike secondary structure (Table 1). Thermal aggregate spectral band positions and areas closely match those of the refolding aggregate, although only a single loop band at 1654 cm^{-1} was observed. Thus, the secondary structure of the thermal aggregate is almost identical to that of the inclusion body and refolding aggregate. Careful inspection of the amide I, II, and III bands for the thermal aggregate indicates that it is *slightly* more nativelike than the other two aggregates. However, the very close resemblance to the other two aggregates in the amide I, II, and III regions suggests that the thermal aggregate is formed from a similar intermediate state.

Amide II and III Spectral Analysis. The amide II (1500–1600 cm^{-1}) and amide III (1200–1400 cm^{-1}) regions were also examined. The spectra of aggregated and native proteins are similar in both of these regions. The resolution-enhanced FSD spectrum of the amide II region is shown in Figure 2. The amide II band corresponds predominantly to N–H bend modes and side-chain modes; in addition, several side-chain contributions are found in this region. As such, it contains some potential tertiary structural information, although interpretation can only be qualitative due to its complexity. As in the amide I region, only a slight structural difference between the mutant and WT native states is evident. The aggregates all share a similar, moderately resolved envelope indicative of some heterogeneity in the IL-1β in the aggregate.

The amide III region, which arises from several amide vibrational modes and also includes some side-chain contributions, is presented in Figure 4. The resolution-enhanced amide III bands reveal that the structures of the three aggregates are extremely similar. The exceptional band in the vicinity of 1262 cm^{-1} is due to buffer components. The second derivative spectra of the native states are also highly similar to each other. The most apparent differences between native and aggregate spectra are in bands near 1310–1325 (turn) and 1220–1225 cm^{-1} (β-sheet); there are also minor differences in the position of the 1290-cm^{-1} band. The major amide III bands corroborate the analysis of the amide I region, in that they show major bands corresponding to β-structure (1230–1250 and 1338 cm^{-1}) and disordered structure (1260 and 1341–45 cm^{-1}) (Anderle & Mendelsohn, 1987). The similarities in the amide II and III regions of the different states, which involve substantial side-chain mode contributions, suggest that much of the tertiary structure is similar in the different aggregates and is not substantially different from that of the native state.

Structural Basis for Aggregation and Inclusion Body Formation. Our results support the hypothesis that a nativelike folding intermediate is responsible for the aggregation of K97V IL-1β. Mutations that increase IB formation by altering the folding pathway might act either by increasing the aggregation rate of the intermediate or by decreasing the rate at which it matures to native protein (Wetzel, 1992a,b), although the exact mechanism by which aggregation occurs remains unclear.

FTIR analysis reveals that the K97V mutant in its native state contains more irregular/loop structure than the WT native state. However, a structural difference between native WT and native K97V cannot be directly involved in IB/aggregate formation: once folded, K97V has the same solubility and higher thermodynamic stability than the WT (B. A. Chrunyk and R. Wetzel, unpublished observations). No difference has been observed between the folding kinetics of WT and those of K97V (Chrunyk & Wetzel, 1993a). In addition, there is a critical step in the thermal *unfolding* of K97V, which leads to aggregation at a temperature significantly lower than expected on the basis of its thermodynamic stability (Chrunyk & Wetzel, 1993a,b).

Evidence that the mutation at position 97 alters the folding pathway of IL-1β lies in the fact that K97V aggregates when folding, while the wild type does not (Chrunyk & Wetzel,

1993a; B. A. Chrunyk and R. Wetzel, unpublished results). The congruence between FTIR data for the inclusion body and folding aggregate (Figure 3 and Table 1) also implies that aggregation occurs by similar pathways in both environments. We consider the generation of insoluble aggregates during refolding from Gdn·HCl to be an *in vitro* model system for IB formation that mimics the *in vivo* situation that produces IBs. That this model system can distinguish between the K97V mutant and wild-type IL-1β indicates that cellular factors (or their absence) are not responsible for the aggregation.

In their analysis of folding-related aggregation *in vitro* of *E. coli* tryptophanase, London et al. (1974) suggested that the aggregation might be mediated by (incorrect) *inter*molecular interactions between structural units normally involved in *intra*molecular interactions in the folded protein. This model proposes rapid, correct folding of subdomains, which then incorrectly associate with the folded regions of other molecules. Although this model for folding-related aggregation is often invoked to account for inclusion body formation (Mitraki & King, 1989), it has not been tested due to the difficulties involved in the structural examination of protein aggregates. Two aspects of our results are consistent with this model. First, the overall nativelike structure for both *in vivo* (IB) and *in vitro* (folding) aggregates of IL-1β shows that the majority of the folding process has occurred before the onset of aggregation. Rapid formation of native subdomains during folding *should* produce the observed nativelike secondary structure. If aggregation involved extensive *inter*molecular H-bonding early in folding, the FTIR spectrum would be anticipated to show a conformation quite different from that of the native state. Further, intermolecular aggregation, such as in heat-gelled proteins or thermally denatured protein aggregates, is associated with a band at 1615–1620 cm^{-1}, reflecting intermolecular β-sheet (Clark et al., 1981). Such a band was not observed in the IB or in other aggregates.[4] Second, the observed increase in β-sheet H-bonding strength between native and IB IL-1β suggests an altered topology between subdomains in the aggregates, presumably involving the manner in which supersecondary structural subdomains are arranged with respect to one another. At least part of the spectral differences must reflect the *inter*molecular interactions in the IB and other aggregates. During aggregation, *inter*molecular hydrophobic interactions will shield regions of the protein from solvent and reduce the driving force to minimize the exposed hydrophobic surface area that would normally be present during folding of the monomer. This effect will increase the available conformations of the folded protein and, thus, allow more optimal H-bonding in both *inter*- and *intra*molecular β-sheets in the aggregates.

The FTIR results for the inclusion bodies (and other aggregates) suggest that the three $\beta\beta\beta$Lβ subdomains of the native protein are retained relatively intact in the IB, but are organized into a different topology. We suggest two models to account for aggregation in this context: the first is one in which the β-barrel structure is "unrolled" by breaking contacts between one set of strands within the barrel core, possibly β-strands 8 and 9. These strands could then form nativelike interactions with the complementary strands of another monomer to form a polymeric antiparallel β-sheet composed entirely of barrel core strands making nativelike hydrogen bonds. Such a model is supported by recently reported folding kinetics of IL-1β indicating relatively slow formation of H-bonds between β-strands 1 and 12 [see Figures 3 and 4 in Varley et al. (1993)].

An alternative possibility is based on the structure in the vicinity of residue 97, which suggests intriguing possibilities for the intermediates. Residue 97 in wild-type IL-1β is part of a long, irregular loop that has a high degree of internal structure due to a number of H-bonded turns and intraloop hydrophobic interactions (Driscoll et al., 1990). The loop also contains a *cis* peptidyl–proline bond at Pro-91 which may establish folding/unfolding kinetics, allowing mutations elsewhere in the loop to be expressed in aggregation phenomena. NMR results on wild-type IL-1β also suggest that the 86–99 loop can take part in a conformational interchange in the native state (Clore & Gronenborn, 1991). Thus, it is possible that the loop region involving residue 97, which in the K97V mutant probably extends from residues 86 to 106, is responsible for the aggregation. In this second model, the remainder of the molecule is essentially nativelike in structure, but interactions between this loop (86–106) and adjacent molecules lead to a three-dimensional network and aggregation. These interactions must occur in a folding intermediate, rather than in the native state, for the reasons enumerated above.

A recent time-resolved amide protection NMR study by Varley et al. (1993) demonstrated that Lys-97 forms a stable H-bond with Val-100 within 0.7–1.5 s after initiation of folding. A mutation that interferes with the formation of this bond may disrupt folding sufficiently for the intermediate (which has exposed hydrophobic surfaces, as revealed by its ability to bind ANS) to aggregate at this stage. Refolding kinetics studies (Craig et al., 1987; Varley et al., 1993) indicate that subdomain A (β-strands 1–4) is formed considerably more slowly than the other two subdomains in the WT protein and that the initially formed β-sheets are only transiently stable. In addition, the slowest step ($t_{1/2} \approx 20$ min) in folding may not be due to proline isomerization, but may involve the final packing of the barrel structure. In this framework, it seems likely that aggregation, in terms of the first model proposed above, may arise from interactions of the A subdomain of one molecule with the B (β-strands 5–8) and C (β-strands 9–12) subdomains of another molecule. Our results are most consistent with aggregation occurring from either intermediate X_3 or X_4 in the kinetic scheme of Varley et al. (1993).

In conclusion, this study has found a remarkable similarity in the overall secondary structure content of human IL-1β inclusion bodies, refolding aggregates, and thermal aggregates and that in the native state. Our data show that the aggregating species, in all cases, is one with nativelike secondary structure. Thus, models for inclusion body formation that involve aggregation of the denatured state, or intermediates formed early in the folding process, can be eliminated.

[4] The small absorbance in the vicinity of 1610 cm^{-1} in all spectra is due to aromatic side-chain absorption. The relative area of this band was the same, regardless of the conformational state of the protein, and was not used in secondary structure calculations.

REFERENCES

Anderle, G., Mendelsohn, R. (1987) *Biophys. J. 52,* 69–74.

Byler, M., & Susi, H. (1986) *Biopolymers 25,* 469–487.

Chrunyk, B., & Wetzel, R. (1993a) In *Protein Folding In Vivo and In Vitro.* (Cleland, J. L. Ed.) pp 46–58, American Chemical Society Books. Washington, DC.

Chrunyk, B. A., & Wetzel, R. (1993b) *Protein Eng. 6,* 733–738.

Chrunyk, B. A., Evans, J., Lillquist, J., Young, P., & Wetzel, R. (1993) *J. Biol. Chem. 268,* 18053–18061.

Clark, A. H., Saunderson, D. H. P., & Suggett, A. (1981) *Int. J. Pept. Protein Res. 17,* 353–364.

Clore, G. M., & Gronenborn, A. M. (1991) *J. Mol. Biol. 221,* 47–53.

Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29,* 8172–8184.
Clore, G. M., Wingfield, P. T., & Gronenborn, A. M. (1991) *Biochemistry 30,* 2315–23.
Craig, S., Schmeissner, U., Wingfield, P., & Pain, R. (1987) *Biochemistry 26,* 3570–3576.
Dong, A., Huang, P., & Caughey, W. S. (1990) *Biochemistry 29,* 3303–3308.
Driscoll, P. C., Gronenborn, A. M., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29,* 4668–4682.
Dukor, R. K., Pancoska, P., Keiderling, T. A., Prestrelski, S. J., & Arakawa, T. (1992) *Arch. Biochem. Biophys. 298,* 678–681.
Fierke, C. A., Calderone, T. L., & Krebs, J. F. (1991) *Biochemistry 30,* 11054–11063.
Finzel, B. C., Clancy, L. L., Holland, D. R., Muchmore, S. W., Watenpaugh, K. D., & Einspahr, H. M. (1989) *J. Mol. Biol. 209,* 779–791.
Goormaghtigh, E., Cabiaux, V., & Ruysschaert, J.-M. (1990) *Eur. J. Biochem. 193,* 409–420.
Griffiths, P. R., & Pariente, G. L. (1986) *Trends Anal. Biochem. 5,* 209–215.
King, J., Fane, B., Haase-Pettingell, C., Mitraki, A., Villafane, R., & Yu, M.-H. (1990) in *Protein Folding: Deciphering the Second Half of the Genetic Code* (Gierasch, L. M., & King, J. Eds.) pp 225–240, American Association for the Advancement of Science, Washington, DC.
Krueger, J. K., Stock, A. M., Schutt, C. E., & Stock, J. B. (1990) in *Protein Folding: Deciphering the Second Half of the Genetic Code* (Gierasch, L. M., & King, J. Eds.) pp 136–142, American Association for the Advancement of Science, Washington, DC.
London, J., Skrzynia, C., & Goldberg, M. E. (1974) *Eur. J. Biochem. 47,* 409–415.
Marston, R. A. O. (1986) *Biochem. J. 240,* 1–12.
Mitraki, A., & King, J. (1989) *Bio/Technology 7,* 690–697.
Myers, C. A., Johanson, K. O., Miles, L. M., McDevitt, P. J., Simon, P. L., Webb, R. L., Chen, M. J., Holskin, B. P., Lillquist, J. S., & Young, P. R. (1987) *J. Biol. Chem. 262,* 11176–11181.
Perry, L. J., & Wetzel, R. (1986) *Biochemistry 25,* 733–739.
Powell, J. R., Wasacz, F. M., & Jacobsen, R. J. (1986) *Appl. Spectrosc. 40,* 339–344.
Prestrelski, S. J., Byler, D. M., & Liebman, M. N. (1991) *Biochemistry 30,* 133–143.
Priestle, J. P., Schar, H.-P., & Grutter, M. G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 9667–9671.
Rinas, U., Tsai, L. B., Lyons, D., Fox, G. M., Stearns, G., Fieschko, J., Fenton, D., & Bailey, J. E. (1992) *Biotechnology 10,* 435–440.
Schein, C. H. (1989) *Biotechnology 7,* 1141–1149.
Surewicz, W. K., & Mantsch, H. H. (1998) *Biochim. Biophys. Acta 952,* 115–130.
Susi, H. (1969) Infrared Spectra of Biological Macromolecules and Related Systems. in *Structure and Stability of Biological Macromolecules,* Biological Macromolecules Series, Vol. 2 (Timascheff, S., & Fasman, G. Eds.) Marcel Dekker, New York.
Swedberg, S. A., Pesek, J. J., & Fink, A. L. (1990) *Anal. Biochem. 186,* 153–158.
Truong, H.-T. N., Pratt, E. A., Rule, G. S., Hsue, P. Y., & Ho, C. (1991) *Biochemistry 30,* 10722–10729.
Varley, P. Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., & Clore, G. M. (1993) *Science 260,* 1110–1113.
Wetzel, R. (1992a) in *Protein Engineering—A Practical Approach* (Rees, A. R., Sternberg, J. E., & Wetzel, R., Eds.). pp 191–219, IRL Press at Oxford University Press, Oxford, U.K.
Wetzel, R. (1992b) in *Stability of Protein Pharmaceuticals: In Vivo Pathways of Degradation and Strategies for Protein Stabilization* (Ahern, T. J., & Manning, M. C., Eds.) pp 43–88, Plenum Press, New York.
Wetzel, R., & Goeddel, D. V. (1983) in *The Peptides: Analysis, Synthesis, Biology* (Meienhofer, J., & Gross, E., Eds.) pp 1–64, Academic Press, New York.
Wetzel, R., & Chrunyk, B. A. (1992) In *Biocatalyst Design for Stability and Specificity* (Himmel, M. E., & Georgiou, G., Eds.) pp 116–125, American Chemical Society, Washington, DC.
Wetzel, R., Perry, L. J., & Veilleux, C. (1991) *Biotechnology 9,* 731–737.
Wilder, C. L., Friedrich, A. D., Potts, R. O., Daumy, G. O., & Francoeur, M. L. (1992) *Biochemistry 31,* 27–31.
Williams, D. C., Van Frank, R. M., Muth, W. L., & Burnett, J. P. (1982) *Science 215,* 687–689.