Raman Markers of Nonaromatic Side Chains in an α-Helix Assembly: Ala, Asp, Glu, Gly, Ile, Leu, Lys, Ser, and Val Residues of Phage *fd* Subunits[†,‡]

Stacy A. Overman and George J. Thomas, Jr.*

Division of Cell Biology and Biophysics, School of Biological Sciences, University of Missouri−Kansas City, Kansas City, Missouri 64110

Received December 9, 1998; Revised Manuscript Received February 1, 1999

ABSTRACT: The study of filamentous virus structure by Raman spectroscopy requires accurate band assignments. In previous work, site- and residue-specific isotope substitutions were implemented to elucidate definitive assignments for Raman bands arising from vibrational modes of the α-helical coat protein main chain and aromatic side chains in the class I filamentous phage, *fd* [Overman, S. A., and Thomas, G. J., Jr. (1995) *Biochemistry 34*, 5440−5451; Overman, S. A., and Thomas, G. J., Jr. (1998) *Biochemistry 37*, 5654−5665]. Here, we extend the previous methods and expand the assignment scheme to identify Raman markers of nonaromatic side chains of the coat protein in the native *fd* assembly. This has been accomplished by Raman analysis of 11 different *fd* isotopomers selectively incorporating deuterium at specific sites in either alanine, aspartic acid, glutamic acid, glycine, isoleucine, leucine, lysine, serine, or valine residues of the coat protein. Raman markers are also identified for the corresponding deuterated side chains. In combination with previous assignments, the results provide a comprehensive understanding of coat protein contributions to the Raman signature of the *fd* virion and validate Raman markers assigned to the packaged single-stranded DNA genome. The findings described here show that nonaromatic side chains contribute prolifically to the *fd* Raman signature, that marker bands for specific nonaromatics differ in general from those observed in corresponding polypeptides and amino acids, and that the frequencies and intensities of many nonaromatic markers are sensitive to secondary and higher-order structures. Nonaromatic markers within the 1200−1400 cm^{-1} interval also interfere seriously with the diagnostic Raman amide III band that is normally exploited in secondary structure analysis. Implications of these findings for the assessment of protein conformation by Raman spectroscopy are considered.

The Raman spectrum of a protein consists of several dozen bands that are diagnostic of the covalent structure, conformation, and environment of the protein main chain and its amino acid side chains. Up to the present, protein Raman spectra have been most often exploited for analyses of main chain secondary structures and to assess structural details of aromatic and sulfur-containing side chains. This is a consequence of the inherently high spectral intensities associated with Raman bands of these protein groups and the large number of studies that have been carried out on related model compounds. Surveys of both off-resonance Raman and ultraviolet-resonance Raman (UVRR)[1] applications to proteins and their assemblies have been reviewed recently (*1*, *2*). On the other hand, with the exception of cysteine and cystine, relatively little is known about the contributions of nonaromatic side chains to the Raman spectrum of a typical protein, and virtually nothing is known about correlations that may exist between Raman bands of nonaromatic groups and their structural context. Here, we describe and implement an empirical approach for elucidating contributions of specific nonaromatic side chains to a protein Raman spectrum and for evaluating their structural significance.

Interpretation of protein Raman bands for structural purposes rests ultimately upon definitive band assignments (*3*−*5*). However, several factors may complicate the development of a reliable assignment scheme. For example, the occurrence of similar groups in chemically different side chains (such as methylene groups in Lys and Leu) or of the same group in closely related side chains (such as carboxyl groups in Asp and Glu) can generate Raman bands that overlap one another. A second complicating factor for globular proteins is the likelihood that identical side chains may generate nonidentical Raman markers, when located in regions of very different secondary and/or tertiary structure. Third, vibrational coupling between main chain and side chain groups may severely hinder spectral interpretations. Many examples involving CH bending vibrations have been cited (*6*−*8*). Finally, Raman bands associated with side chains that are neither aromatic nor sulfur-containing may be inherently very weak and therefore very difficult to detect in comparison to other Raman bands of the protein. We have addressed these difficulties by focusing on the filamentous bacterial virus *fd*, which is a well-characterized quaternary assembly of identical protein subunits (pVIII), each consisting of a single, small secondary structure domain (50-residue α-helix). The pVIII protein is the predominant contributor to the Raman spectrum of *fd* (*3*).

[†] Supported by NIH Grant GM50776.

[‡] Paper LXIII in the series Structural Studies of Viruses by Raman Spectroscopy.

* To whom correspondence should be addressed.

[1] Abbreviations: UVRR, ultraviolet-resonance Raman; pVIII, coat protein encoded by viral gene VIII; D, deuterium (^2H); ssDNA, single-stranded DNA.

10.1021/bi982901e CCC: $18.00
Published on Web 03/13/1999

To circumvent the problem of Raman band overlap and to identify Raman bands from individual nonaromatic side chains, we have incorporated specifically labeled amino acids singly into the *fd* coat protein using previously established methods (*4*, *9*). The present emphasis is on determination of the qualitative and quantitative contributions of Ala, Asp, Glu, Gly, Ile, Leu, Lys, Ser, and Val side chains of pVIII to the *fd* Raman spectrum. The nine amino acids selected for labeling in this study together represent 38 of the 50 residues of pVIII (sequence ^{1}AEGDDPAKAA^{11}FDSLQASATE^{21}-YIGYAWAMVV^{31}VIVGATIGIK^{41}LFKKFTSKAS). Because the *fd* virion is an assembly of uniformly packed α-helical domains (*10*−*12*), it is expected that the results of this study will be applicable to the same nonaromatic amino acids in other α-helical proteins and their assemblies. The findings described here may also provide a basis for understanding the contributions of these nonaromatic side chains to Raman spectra of globular proteins containing other types of secondary, tertiary, or quaternary structures.

The cataloging of Raman markers of nonaromatic side chains of pVIII completes a comprehensive assignment scheme for Raman bands of the *fd* coat protein. The results described here, together with assignments developed previously for the coat protein main chain (*13*) and aromatic side chains (*4*, *14*), confirm that the packaged *fd* genome contributes only feebly to the Raman spectrum of the virus. Raman markers deduced here for the packaged single-stranded (ss) DNA molecule are consistent with those proposed recently by UVRR spectroscopy of *fd* using excitation at 257 nm (*15*). The *fd* DNA assignments are also in accord with earlier analyses of the Raman spectra of several filamentous viruses (*3*, *16*).

MATERIALS AND METHODS

Sample Preparation. The *fd* variants used in this study were prepared from stocks of the wild-type virus obtained originally from L. A. Day (Public Health Research Institute, New York). Phage were grown on *Escherichia coli* strain Hfr3300 as the host. Growth medium and standard reagents were obtained from Sigma Chemical Co. (St. Louis, MO) and Fisher Scientific (St. Louis, MO).

Unlabeled viruses were grown in MS medium containing 1% glucose and 4 mM $CaCl_2$. Mature viral particles, extruded through the bacterial cell membrane and into the growth medium, were collected by precipitation with poly(ethylene glycol) (20 g/L) and NaCl (0.5 M) followed by low-speed centrifugation (*4*). The virus precipitate was resuspended in 10 mM Tris (pH 7.8 ± 0.2) and pelleted by centrifugation at 330000*g* for 1.5 h at 4 °C. The resulting virus pellet was purified by three additional cycles of resuspension in Tris buffer and repelleting under the same conditions. This procedure removed excess NaCl from the virus pellet (*3*). Typically, 30−40 mg of purified virus was obtained from a 1 L preparation.

To incorporate labeled amino acids into the coat protein of *fd*, the host and virus were grown in M9 minimal medium containing the appropriately labeled L-amino acid (Table 1) at a concentration of 0.1 mM per residue per subunit. Labeled amino acids were obtained from Cambridge Isotope Laboratories (Woburn, MA). The medium also contained all other L-amino acids, each at 0.1 mM, with 10 μg/mL thiamine hydrochloride and 1% glucose. Excess NH_4Cl (3 g) was added to the medium to reduce unwanted metabolic transfer of deuterium (D) labels to unlabeled amino acids. Mature viral particles were isolated as described above. The resulting virus solution was purified on a continuous KBr gradient (initial density of 1.257 g/mL) in a swinging bucket rotor at 247000*g* for 48 h at 4 °C. The virus band was collected, diluted 2-fold with 10 mM Tris, and subjected to the four cycles of centrifugation and resuspension described above. This procedure accomplished removal of excess KBr. Typically, 20 mg of purified virus was obtained from a 1 L preparation. The specific *fd* isotopomers prepared for this study are listed in Table 1.

Table 1: Deuterium-Labeled Amino Acids Incorporated in *fd* Virus

labeled amino acid	chemical formula[a]	abbreviation	nomenclature for virus
L-alanine-2-d_1	$CH_3C^{\alpha}D$	A_{d1}	*fd*($10A_{d1}$)
L-alanine-3,3,3-d_3	$CD_3C^{\alpha}H$	A_{d3}	*fd*($10A_{d3}$)
L-alanine-d_4	$CD_3C^{\alpha}D$	A_{d4}	*fd*($10A_{d4}$)
L-aspartic acid-d_3	$HOOCCD_2C^{\alpha}D$	D_{d3}	*fd*($3D_{d3}$)
L-glutamic acid-d_5	$HOOC(CD_2)_2C^{\alpha}D$	E_{d5}	*fd*($2E_{d5}$)
glycine-d_2	$C^{\alpha}D_2$	G_{d2}	*fd*($4G_{d2}$)
L-isoleucine-d_{10}	$CD_3CD_2CD(CD_3)C^{\alpha}D$	I_{d10}	*fd*($4I_{d10}$)
L-leucine-d_{10}	$(CD_3)_2CDCD_2C^{\alpha}D$	L_{d10}	*fd*($2L_{d10}$)
L-lysine-3,3,4,4,5,5,6,6-d_8	$NH_2(CD_2)_4C^{\alpha}H$	K_{d8}	*fd*($5K_{d8}$)
L-serine-3,3-d_2	$HOCD_2C^{\alpha}H$	S_{d2}	*fd*($4S_{d2}$)
L-valine-d_8	$(CD_3)_2CDC^{\alpha}D$	V_{d8}	*fd*($4V_{d8}$)

[a] Carboxyl and amino substituents of C^{α} are omitted.

Raman Spectroscopy. Solutions of viruses (∼80 mg/mL) and amino acids (∼20−100 mg/mL) were sealed in standard glass capillaries (KIMAX no. 34507) for Raman analysis. Typically, a 2 μL aliquot was sufficient to fill the portion of the sample cell exposed to laser illumination. All spectral data were collected from samples thermostated at 12 °C (*17*).

Raman spectra were excited with 400 mW of 514.5 nm excitation of an argon ion laser (Innova 70, Coherent Inc., Santa Clara, CA). Spectra of labeled (Ala, Asp, Glu, Gly, Ile, Leu, Lys, Ser, or Val isotopomers) and unlabeled *fd* were recorded on a computer-controlled scanning double spectrometer (Ramalog V/VI, Spex Industries, Edison, NJ), equipped with a photon-counting detector (model R928P, Hamamatsu, Middlesex, NJ). Data at intervals of 1 cm^{-1} were obtained with a 1.5 s integration time and an 8 cm^{-1} spectral slit width. Additional spectra of Ala-labeled isotopomers and unlabeled *fd* were collected on a more recently acquired single monochromator system of superior performance, incorporating a notch filter and charged-coupled-device (CCD) detector (Spex Industries). Frequencies cited are accurate to ±1 cm^{-1} for sharp bands and ±2 cm^{-1} for very broad bands or shoulders. Further details about the instrumentation and data collection protocols are given elsewhere (*4*, *18*).

Raman difference spectra were computed (SpectraCalc, Galactic Industries, Salem, NH) prior to correcting raw data for contributions of solvent or background. This differentiated subtle changes that might otherwise have been difficult to detect and avoided introduction of artifacts from incomplete solvent compensation in either minuend or subtrahend (*4*). In each case, the spectrum of the labeled virus is the minuend and that of the unlabeled virus is the subtrahend. A difference band is considered significant if it exhibits a signal-to-noise

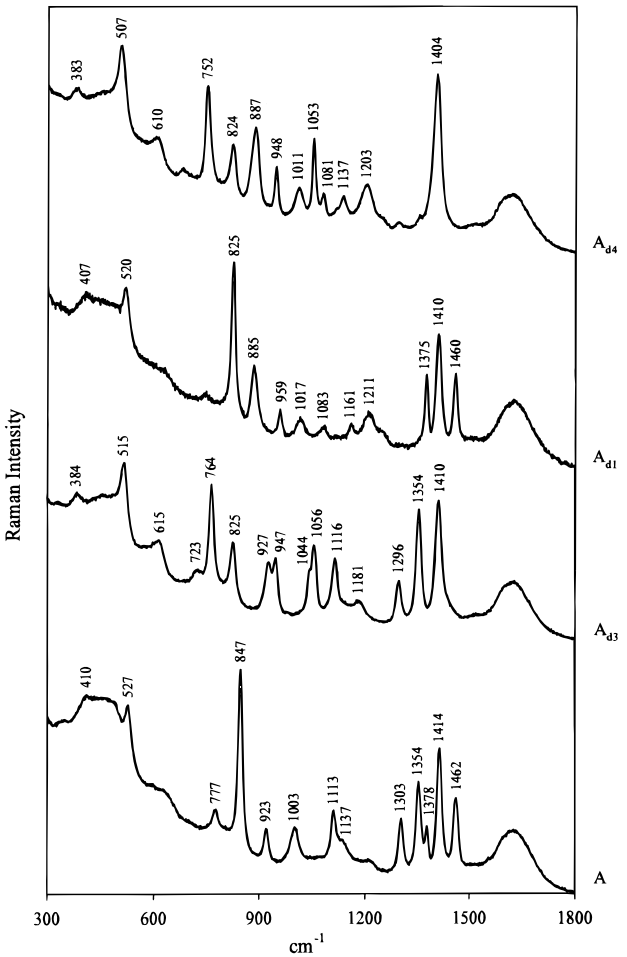

FIGURE 1: Raman spectra (300−1800 cm^{-1}) of L-alanine and C-deuterated isotopomers (Table 1). Frequencies of prominent bands are indicated in wavenumber units (cm^{-1}). Each amino acid was dissolved to saturation (20−100 mg/mL) in H_2O at pH 7.0. Instrument conditions were as follows: excitation wavelength, 514.5 nm; laser power, 400 mW; sample temperature, 12 °C; spectral slit width, 8 cm^{-1}; increment, 1 cm^{-1}; and integration time, 1.5 s. Each spectrum is the average of two to eight scans. No smoothing or solvent corrections were employed.

ratio of at least 2:1 and an intensity equal to at least 10% of that of the parent Raman band.

RESULTS

Raman spectra were collected from several deuterium-labeled variants of *fd*. Each incorporated coat subunits (protein pVIII) containing one of the following carbon-deuterated L-amino acids: Ala-2-d_1, Ala-3,3,3-d_3, Ala-d_4, Asp-d_3, Glu-d_5, Gly-d_2, Ile-d_{10}, Leu-d_{10}, Lys-3,3,4,4,5,5,6,6-d_8, Ser-d_2, and Val-d_8. In this nomenclature, the labeled side chain is perdeuterated, unless indicated otherwise. Additional nomenclature employed for the labeled viruses is summarized in Table 1.

Raman spectra of the deuterium-labeled amino acids of Table 1 are compared with spectra of corresponding unlabeled amino acids in Figures 1−3. In Figure 4, we compare the Raman signature of unlabeled *fd* with the Raman signatures of the following *fd* viruses, each of which incorporates the indicated deuterated alanines in all coat subunits: *fd*(10A$_{d1}$), containing 10 Ala-d_1 residues; *fd*(10A$_{d3}$), containing 10 Ala-d_3; and *fd*(10A$_{d4}$), containing 10 Ala-d_4.

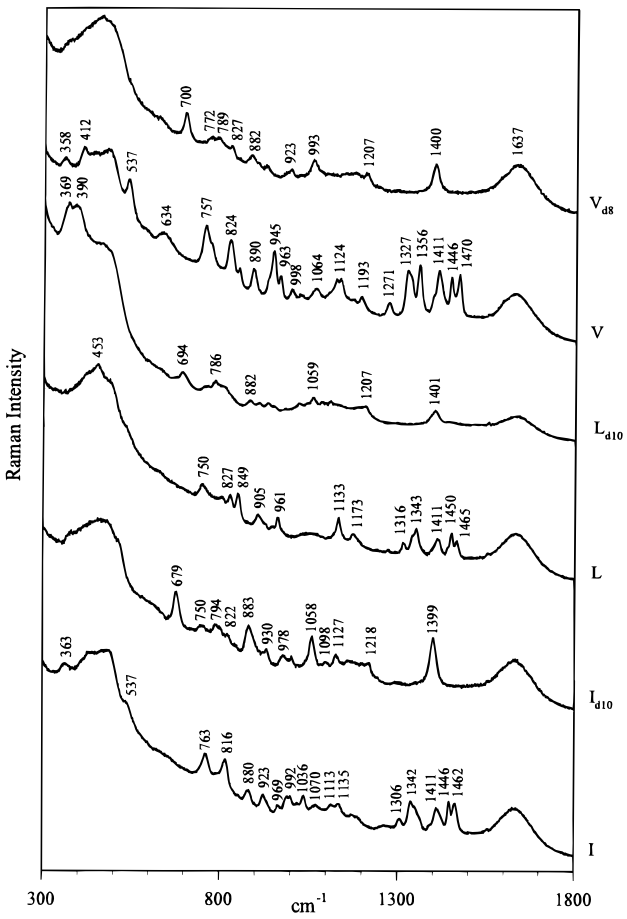

FIGURE 2: Raman spectra (300−1800 cm^{-1}) of L-valine, L-leucine, and L-isoleucine and C-deuterated isotopomers (Table 1). Other conditions are given in the legend of Figure 1.

In Figure 5, unlabeled *fd* is similarly compared with *fd*(4I$_{d10}$), which contains four Ile-d_{10}, with *fd*(2L$_{d10}$), which contains two Leu-d_{10}, and with *fd*(4V$_{d8}$), which contains four Val-d_8. Corresponding data for glycine, lysine, and serine variants are shown in Figure 6. Similar comparisons were made for aspartic acid and glutamic acid variants, which exhibited only very weak deuteration-shifted Raman bands (data not shown). Figures 4−6 include difference spectra computed by subtracting the Raman spectrum of unlabeled *fd* from that of the labeled virus. The difference spectra thus reveal the principal Raman markers of the unlabeled side chains as troughs and those of the labeled side chains as peaks. For each difference computation, the constituent spectra were normalized to eliminate (or minimize) Raman difference bands known to be unrelated to the modified side chain. Accordingly, the difference spectra of Figures 4−6 are virtually free of bands from either the peptide main chain or aromatic side chains. Further discussions about the spectral normalization and subtraction procedures have been given elsewhere (*4*, *13*, *14*).

Figures 4−6 reveal that the Raman spectrum of *fd* contains many bands common to more than one type of nonaromatic side chain, as well as bands that are unique to a given type of side chain. An example of the former is the band near 1458 cm^{-1}, which can be attributed to methyl CH bending modes of Ala, Ile, Leu, and Val (Figures 4 and 5); an example of the latter is the band near 890 cm^{-1}, which can be attributed to CC and C$^{\epsilon}$N$^{\zeta}$ stretching modes of Lys (Figure

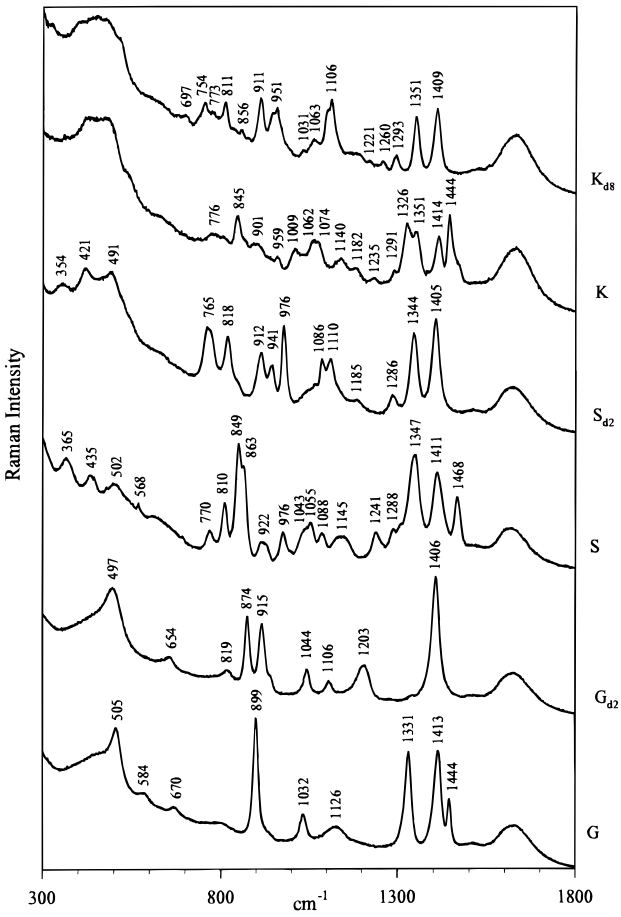

FIGURE 3: Raman spectra (300−1800 cm^{-1}) of L-lysine, L-serine, and glycine and C-deuterated isotopomers (Table 1). Other conditions are given in the legend of Figure 1.

6). A comprehensive summary of Raman assignments developed from data of Figures 1−6 is given in Tables 2 and 3, and more detailed discussion is given below. For convenience, Raman side chain markers discussed in the following sections are rounded to the nearest 5 cm^{-1}.

DISCUSSION

Alanine. The difference spectra of Figure 4 show that the alanine residues of *fd* produce eight distinct Raman markers (troughs near ∼760, ∼910, ∼1070, ∼1105, ∼1340, ∼1455, and ∼1470 cm^{-1}). Comparison of these results with studies of aliphatic hydrocarbons and related model peptides (*8*, *13*, *19*−*22*) suggests the following as the most plausible assignments: 760 cm^{-1}, C$^{\beta}$H$_3$ rock and skeletal bend; 910 cm^{-1}, C$^{\alpha}$C$^{\beta}$ stretch and skeletal bend; 1070 cm^{-1}, C$^{\alpha}$C$^{\beta}$ stretch; 1105 cm^{-1}, C$^{\alpha}$C$^{\beta}$ stretch and C$^{\beta}$H$_3$ rock; 1340 cm^{-1}, C$^{\alpha}$H bend and C$^{\alpha}$C stretch; and 1455 and 1470 cm^{-1}, C$^{\beta}$H$_3$ asymmetric bend. These assignments are supported additionally by studies of the *fd* isotopomer incorporating ^{13}C at each alanyl C site (*13*). Normal coordinate calculations and Raman spectra of alanyl peptides also support assignment of the band near 1340 cm^{-1} to a vibration involving the C$^{\alpha}$H bend (*6*−*8*). The two bands assigned to C$^{\beta}$H$_3$ asymmetric bending (1455 and 1470 cm^{-1}) are somewhat sensitive to C$^{\alpha}$ deuteration (Figure 4), and therefore, the methyl group vibrations may be coupled weakly to the main chain through the C$^{\alpha}$−C$^{\beta}$ bond. In the absence of such coupling, the pair of bands would presumably be replaced by a single band that could be assigned to a pure C$^{\beta}$H$_3$ bending mode.

Valine. Raman spectra of *fd*(4V$_{d8}$) and *fd* are compared in Figure 5. The valine side chains of pVIII generate numerous troughs in the difference spectrum (trace b-a), for which the following assignments are proposed: ∼520 (C$^{\gamma1}$C$^{\beta}$C$^{\gamma2}$ bend), ∼865 (C$^{\gamma1}$C$^{\beta}$C$^{\gamma2}$ symmetric stretch), ∼940, ∼1130, and ∼1160 (C$^{\alpha}$C$^{\beta}$ or C$^{\beta}$C$^{\gamma}$ stretch), ∼1320 (C$^{\alpha}$H bend), ∼1355 (C$^{\beta}$H bend), and ∼1460 cm^{-1} (C$^{\gamma}$H$_3$ asymmetric bend). Analogous assignments have been proposed for related model compounds (*8*, *19*, *21*, *22*).

Isoleucine. Raman spectra of *fd*(4I$_{d10}$) and *fd* are compared in Figure 5. The difference spectrum (trace d-a) reveals isoleucine markers of pVIII, assigned as follows: ∼740 (C$^{\gamma1}$H$_2$ rock), ∼845, ∼870, ∼990, ∼1130, and ∼1155 (various coupled CC stretch modes involving C$^{\alpha}$C$^{\beta}$, C$^{\beta}$C$^{\gamma1}$, C$^{\beta}$C$^{\gamma2}$, and C$^{\gamma1}$C$^{\delta1}$), ∼1320 (C$^{\alpha}$H bend), 1445 (C$^{\gamma2}$H$_3$ and C$^{\delta1}$H$_3$ asymmetric bend), and ∼1465 cm^{-1} (C$^{\gamma1}$H$_2$ scissor). Additionally, the difference spectrum reveals a weak and broad trough in the 1338−1354 cm^{-1} interval, which could be due to bending vibrations associated with C$^{\beta}$H and C$^{\gamma1}$H$_2$ groups. These assignments are consistent with previous studies of related model compounds (*8*, *19*, *21*−*23*).

Leucine. The *fd*(2L$_{d10}$) minus *fd* difference spectrum (Figure 5, trace c-a) reveals the following leucine markers for pVIII: ∼745 (C$^{\beta}$H$_2$ rock), ∼935, ∼965, and ∼1130 (various coupled CC stretch modes involving C$^{\alpha}$C$^{\beta}$, C$^{\beta}$C$^{\gamma}$, C$^{\gamma}$C$^{\delta1}$, and C$^{\gamma}$C$^{\delta2}$), ∼1445 (C$^{\delta1}$H$_3$ and C$^{\delta2}$H$_3$ asymmetric bend), and ∼1470 cm^{-1} (C$^{\beta}$H$_2$ scissor). As in the case of isoleucine (above), additional difference bands are observed in the 1260−1360 cm^{-1} interval, which are likely due to various CH bending modes.

Lysine. The *fd*(5K$_{d8}$) minus *fd* difference spectrum (Figure 6, trace c-a) reveals numerous lysine marker bands, resulting from perdeuteration of C$^{\beta}$H$_2$, C$^{\gamma}$H$_2$, C$^{\delta}$H$_2$, and C$^{\epsilon}$H$_2$ and assigned as follows: ∼745 (various CH$_2$ rocks), ∼890, ∼935, ∼965, ∼1055 and ∼1070 (various CC and C$^{\epsilon}$N$^{\zeta}$ stretches), ∼1325 and 1345 (various CH$_2$ twists or rocks), and ∼1445 cm^{-1} (various CH$_2$ scissors). Again, these assignments are consistent with studies of model compounds (*8*, *19*, *21*−*23*).

Glycine. The *fd*(4G$_{d2}$) minus *fd* difference spectrum (Figure 6, trace d-a) reveals one clear-cut glycine marker for pVIII near 1345 cm^{-1}, which is assigned to the C$^{\alpha}$H$_2$ twist or rock vibration. Additional weak difference bands are observed in the 800−1000 and 1400−1500 cm^{-1} intervals, which may be due to C$^{\alpha}$C stretching and C$^{\alpha}$H$_2$ bending modes, respectively.

Serine. The *fd*(4S$_{d2}$) minus *fd* difference spectrum (Figure 6, trace b-a) identifies a prominent serine marker near 1340−1345 cm^{-1}, which can be assigned confidently to the C$^{\beta}$H$_2$ twist or rock vibration. In the free amino acid (Figure 3), it is clear that C$^{\beta}$H$_2$ deuteration greatly reduces the peak intensity and bandwidth near 1345 cm^{-1}. Somewhat more complicated is the absence in Figure 6 (trace b-a) of a trough corresponding to the intense doublet at 849 and 863 cm^{-1} in the free amino acid (Figure 3), which suggests that the latter originates from a delocalized vibration involving coupling between the side chain methylene group and the amino terminal group. Additional weak difference bands are observed in the 1400−1500 cm^{-1} interval, which may be

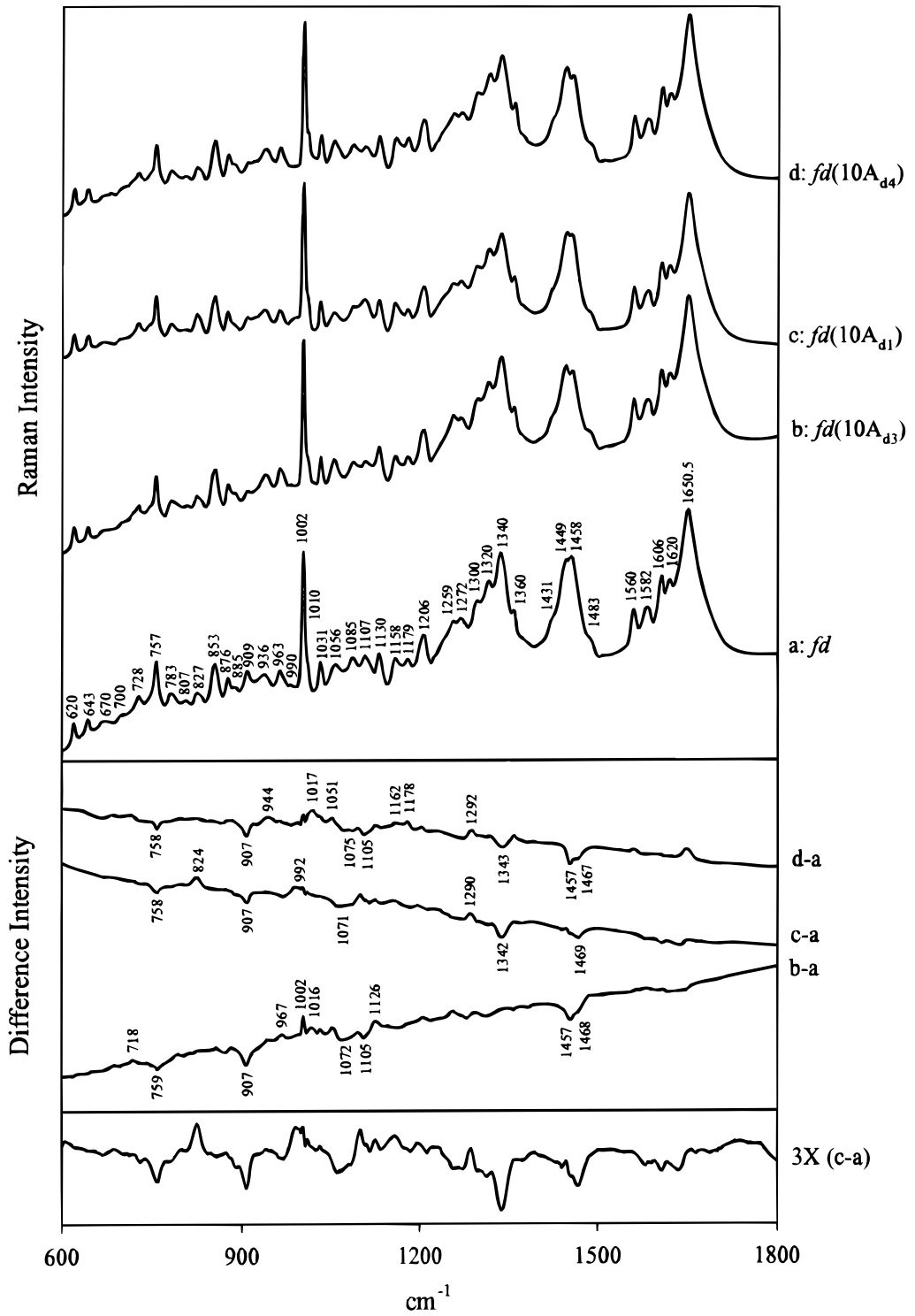

FIGURE 4: (Top) Raman spectra (300−1800 cm^{-1}, accumulation time of 20 min) of unlabeled and isotopically labeled viruses. Sample concentrations were 80 mg/mL in 10 mM Tris at pH 7.8. (Bottom) Normalized difference spectra, computed with the isotopically labeled virus as minuend and unlabeled virus as subtrahend. Difference intensities (peaks and troughs) are normalized to equate intensities of the 643 cm^{-1} Raman band of tyrosine. Other conditions are given in the legend of Figure 1.

due to C$^{\beta}$H$_2$ bending. Other seryl vibrations, including C$^{\beta}$O stretch and bend modes near 1000−1250 and 1600−1700 cm^{-1}, respectively, are not highly sensitive to C$^{\beta}$H$_2$ deuteration and are not detected in the difference spectrum of Figure 6.

Aspartic Acid and Glutamic Acid. The *fd*(3D$_{d3}$) minus *fd* and *fd*(2E$_{d5}$) minus *fd* difference spectra (data not shown) exhibit very weak troughs near 1345 and 1470 cm^{-1}, consistent with the deuteration shifts noted above for other methylene-containing pVIII side chains, particularly, isoleucine and leucine.

Other pVIII Side Chains. The *fd* subunit contains a single residue each of proline, glutamine, and methionine. Bands from these pVIII side chains are considered to contribute marginally to the *fd* Raman spectrum, with the possible exception of CS stretching modes of methionine expected near 650−750 cm^{-1} (*24*). The latter are not likely, however, to be shifted significantly by methylene or methyl deutera-

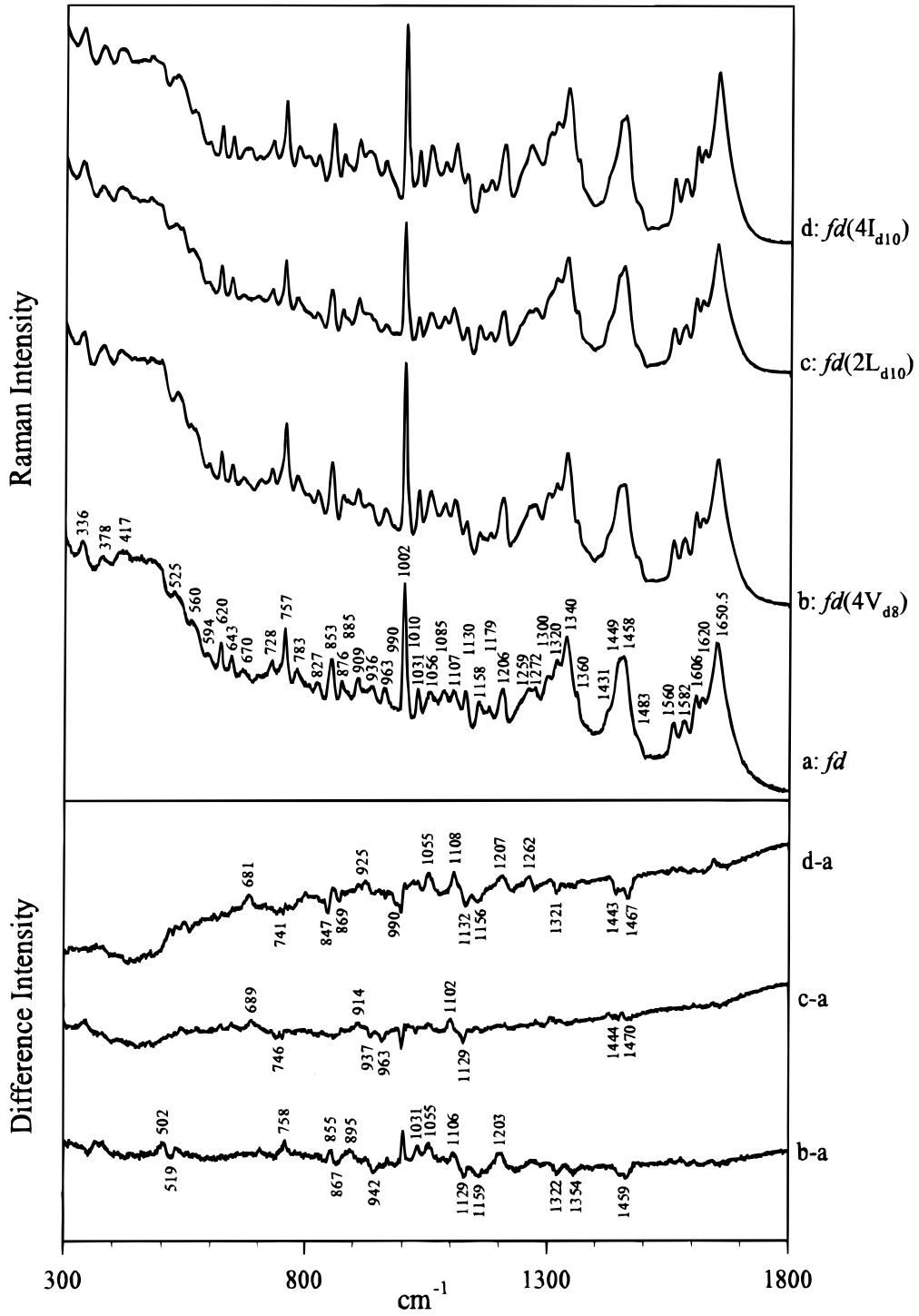

FIGURE 5: (Top) Raman spectra (300−1800 cm^{-1}, average of eight scans) of unlabeled and isotopically labeled viruses. Sample concentrations were 80 mg/mL in 10 mM Tris at pH 7.8. (Bottom) Normalized difference spectra were computed as described in the legend of Figure 4. Other conditions are given in the legend of Figure 1.

tions. (The remaining amino acid constituent of pVIII, threonine, though present in three copies per subunit, could not be obtained in deuterated form.)

Relationship to Raman Markers of fd Aromatics and ssDNA. Raman assignments determined here for the majority of nonaromatic amino acids of the *fd* coat protein are largely distinct from the Raman markers established previously for aromatic amino acids and for the packaged ssDNA genome (*3*, *4*, *13*−*16*, *25*). A comprehensive list is given in Table 2. The relative contributions of nonaromatic side chains to the Raman signature of *fd* and a comparison with contributions from aromatic side chains and encapsidated DNA are shown graphically in Figure 7.

SUMMARY AND CONCLUSIONS

The study presented here completes the series of recent investigations designed to assign all major Raman bands of the *fd* virus to specific main chain and side chain vibrations of coat subunits and to nucleoside vibrations of packaged DNA (*4*, *13*, *14*, *26*). Deuterated derivatives of 38 of the 44 nonaromatic pVIII residues have been incorporated into the *fd* virion. Spectra obtained from the resulting *fd* isotopomers

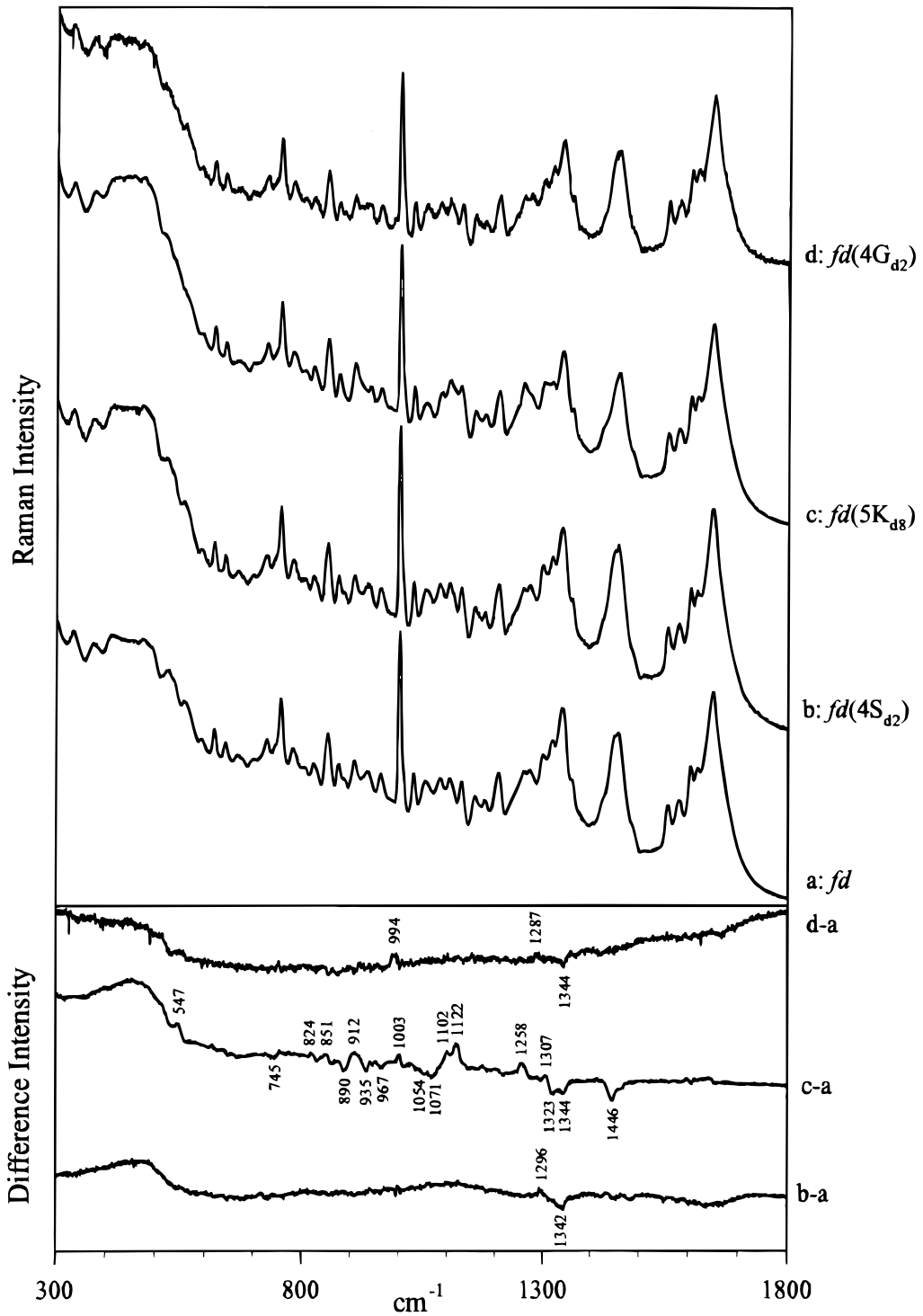

FIGURE 6: (Top) Raman spectra (300−1800 cm^{-1}, average of eight scans) of unlabeled and isotopically labeled viruses. Sample concentrations were 80 mg/mL in 10 mM Tris at pH 7.8. (Bottom) Normalized difference spectra were computed as described in the legend of Figure 4. Other conditions are given in the legend of Figure 1.

permit definitive assignment of Raman bands contributed by alanine, glycine, isoleucine, leucine, lysine, serine, and valine residues in the native virus structure (Table 2 and Figure 7). Several of the nonaromatic marker bands have been assigned to distinct vibrations of specific chemical groups within the amino acid residues (Table 3).

Raman markers associated with bending vibrations of hydrogenic groups (side chain methyne, methylene and methyl, and main chain $C^{\alpha}H$) are generally sensitive to the identity of neighboring groups. For example, a $C^{\alpha}H$ marker is observed near 1340 cm^{-1} when C^{β} is a methyl carbon (alanine), but near 1320 cm^{-1} when C^{β} is either a methyne or methylene carbon (valine, leucine, or isoleucine). Similarly, the side chain CH_2 marker (scissor mode) is observed near 1450 cm^{-1} for a nonbranched chain (lysine), but otherwise near 1470 cm^{-1} (isoleucine, leucine, etc.). This could be due to the fact that C^{ϵ} of lysine is bonded to a quaternary ammonium group, the effect of which is a perturbation of scissor modes of adjacent CH_2 substituents (*22*). Additionally, the CH_3 asymmetric bend marker is

Table 2: Raman Frequencies and Assignments of *fd*[a]

frequency (cm^{-1})	assignment	frequency (cm^{-1})	assignment
336	Y	1002	F
378	pVIII main chain	1010	W
		1031	F, Y
417	Y	1056	**K, A,** F
476	F	1085	**K,** F
525	W, **V**	1107	A
560	pVIII main chain	1130	**I, V, L,** W
		1158	**I, V**
594	pVIII main chain	1179	Y, F
		1206	F, Y, W
620	F	1243	Thy, Cyt
643	Y	1259	W, Thy, Ade
670	Thy, Gua	1272	amide III, Y
728	Ade	1300−1350[b]	amide III, **A**, **K**, **I**, **V**, **S**, **G**, W, Y, pVIII main chain ($C^{\alpha}H$ bend and $C^{\alpha}C$ stretch)
746	**I**, **L**, **K**, amide IV, F, Thy		
757	W, F, **A**	1360	W, **V**, F
807	W, DNA backbone	1431	W
		1449	**K**, **I**, **L**
827	F	1458	A, **I**, **V**, **L**, W, Y
853	Y, **I**	1483	Ade, Gua, W, F
876	W, **I**, **V**	1560	W
885	**K**	1582	W, F, Ade, Gua
909	A	1606	F, Y
936	**K**, **V**, **L**	1620	Y, W
963	**K**, **L**	1651	amide I
990	**I**		

[a] Assignments in boldface type are derived from the work described here; others are from Aubrey and Thomas (*3*) or Overman and Thomas (*13, 14*). Standard one-letter symbols are used for amino acids and three-letter abbreviations for DNA bases. [b] Although band overlap in this region is extensive, amide III contributions are greatest near 1300 cm^{-1} (*8*). A main chain mode ($C^{\alpha}H$ bend and $C^{\alpha}C$ stretch) dominates near 1340 cm^{-1}. Numerous nonaromatics contribute in this interval, producing peaks near 1320 and 1340 cm^{-1}.

Table 3: Vibrational Assignments for Raman Markers of Nonaromatic Side Chains in *fd*

coat protein residue	Raman marker (cm^{-1})	vibrational assignment
alanine	760	$C^{\beta}H_3$ rock + skeletal bend
	910	$C^{\alpha}C^{\beta}$ stretch + skeletal bend
	1070	$C^{\alpha}C^{\beta}$ stretch
	1105	$C^{\alpha}C^{\beta}$ stretch + $C^{\beta}H_3$ rock
	1340	$C^{\alpha}H$ bend + $C^{\alpha}C$ stretch
	1455 and 1470	$C^{\beta}H_3$ asymmetric bend
glycine	1345	$C^{\alpha}H_2$ twist/rock
isoleucine	740	$C^{\gamma 1}H_2$ rock
	845	CC stretch
	870	CC stretch
	990	CC stretch
	1130	CC stretch
	1155	CC stretch
	1320	$C^{\alpha}H$ bend
	1445	$C^{\gamma 2}H_3$ and $C^{\delta 1}H_3$ asymmetric bend
	1465	$C^{\gamma 1}H_2$ scissor
leucine	745	$C^{\beta}H_2$ rock
	935	CC stretch
	965	CC stretch
	1130	CC stretch
	1445	$C^{\delta 1}H_3$ and $C^{\delta 2}H_3$ asymmetric bend
	1470	$C^{\beta}H_2$ scissor
lysine	745	CH_2 rock
	890	CC and $C^{\epsilon}N^{\zeta}$ stretch
	935	CC and $C^{\epsilon}N^{\zeta}$ stretch
	965	CC and $C^{\epsilon}N^{\zeta}$ stretch
	1055	CC and $C^{\epsilon}N^{\zeta}$ stretch
	1070	CC and $C^{\epsilon}N^{\zeta}$ stretch
	1325	CH_2 twist/rock
	1345	CH_2 twist/rock
	1445	CH_2 scissor
serine	1345	$C^{\beta}H_2$ twist/rock
valine	520	$C^{\gamma 1}C^{\beta}C^{\gamma 2}$ bend
	865	$C^{\gamma 1}C^{\beta}C^{\gamma 2}$ symmetric stretch
	940	$C^{\alpha}C^{\beta}$ or $C^{\beta}C^{\gamma}$ stretch
	1130	$C^{\alpha}C^{\beta}$ or $C^{\beta}C^{\gamma}$ stretch
	1160	$C^{\alpha}C^{\beta}$ or $C^{\beta}C^{\gamma}$ stretch
	1320	$C^{\alpha}H$ bend
	1355	$C^{\beta}H$ bend
	1460	$C^{\gamma}H_3$ asymmetric bend

observed near 1460 cm^{-1} in short aliphatic side chains (alanine and valine), but near 1440 cm^{-1} in longer chains (isoleucine and leucine).

Assignments of Tables 2 and 3 constitute the Raman signature for nonaromatic side chains in α-helical coat protein subunits of the class I filamentous assembly. It is reasonable to conclude that similar Raman signatures would occur for nonaromatic side chains in class II filamentous viruses, *Pf1* and *Pf3*, as well as in other α-helical proteins and their assemblies.

Raman markers of the nonaromatic side chains in *fd* subunits are generally different from those of corresponding side chains in free amino acids. Notable examples are valyl and isoleucyl side chains. Both exhibit a prominent deuteration-sensitive Raman band near 1155−1160 cm^{-1} in *fd*, whereas no comparably intense deuteration-sensitive Raman band occurs in the spectrum of either L-Val or L-Ile. This implies in each case a coupling of the aliphatic group vibration (CC stretch) with a protein main chain vibration (probably CN stretch). Also, in the case of alanine, the band observed near 1340 cm^{-1} in *fd*, which is assigned on the basis of the present and previous (*13*) results to a complex mode involving the $C^{\alpha}H$ bend and main chain $C^{\alpha}C$ stretch, exhibits a counterpart in L-Ala only at a considerably higher frequency (1355 cm^{-1}).

Raman signatures of *fd* side chains also differ appreciably from those of synthetic polypeptides. For example, alanyl side chains of *fd* contribute a broad Raman band near 1072 cm^{-1}, whereas no comparable band occurs in α-helical poly-L-alanine (*20*). Interestingly, the alanyl side chain marker observed in *fd* at 1340 cm^{-1} has a counterpart in α-helical poly-L-alanine, but not in β-stranded poly-L-alanine (*20*). Likewise, comparison of the Raman signature of lysyl side chains of *fd* with that of α-helical poly-L-lysine reveals many differences. Lysyl side chains of *fd* contribute a moderately intense Raman band at 1345 cm^{-1}, whereas only a much weaker shoulder is observed in the spectrum of α-helical poly-L-lysine (*23*). Additionally, the lysyl markers near 965, 1055, and 1325 cm^{-1} in *fd* have only a single counterpart near 1310 cm^{-1} in α-helical poly-L-lysine (*23*). Numerous other examples can be seen by comparison of the results presented here (Tables 2 and 3) with previously reported Raman spectra of polypetides (*8*, *20*, *23*, *24*, *27*). These comparisons show that not only primary structures but also secondary and higher-order structures greatly influence the Raman signatures of nonaromatic side chains.

On the basis of alanine isotope shifts demonstrated in Figure 4, the $C^{\alpha}H$ bend and $C^{\alpha}C$ stretch mode for all pVIII

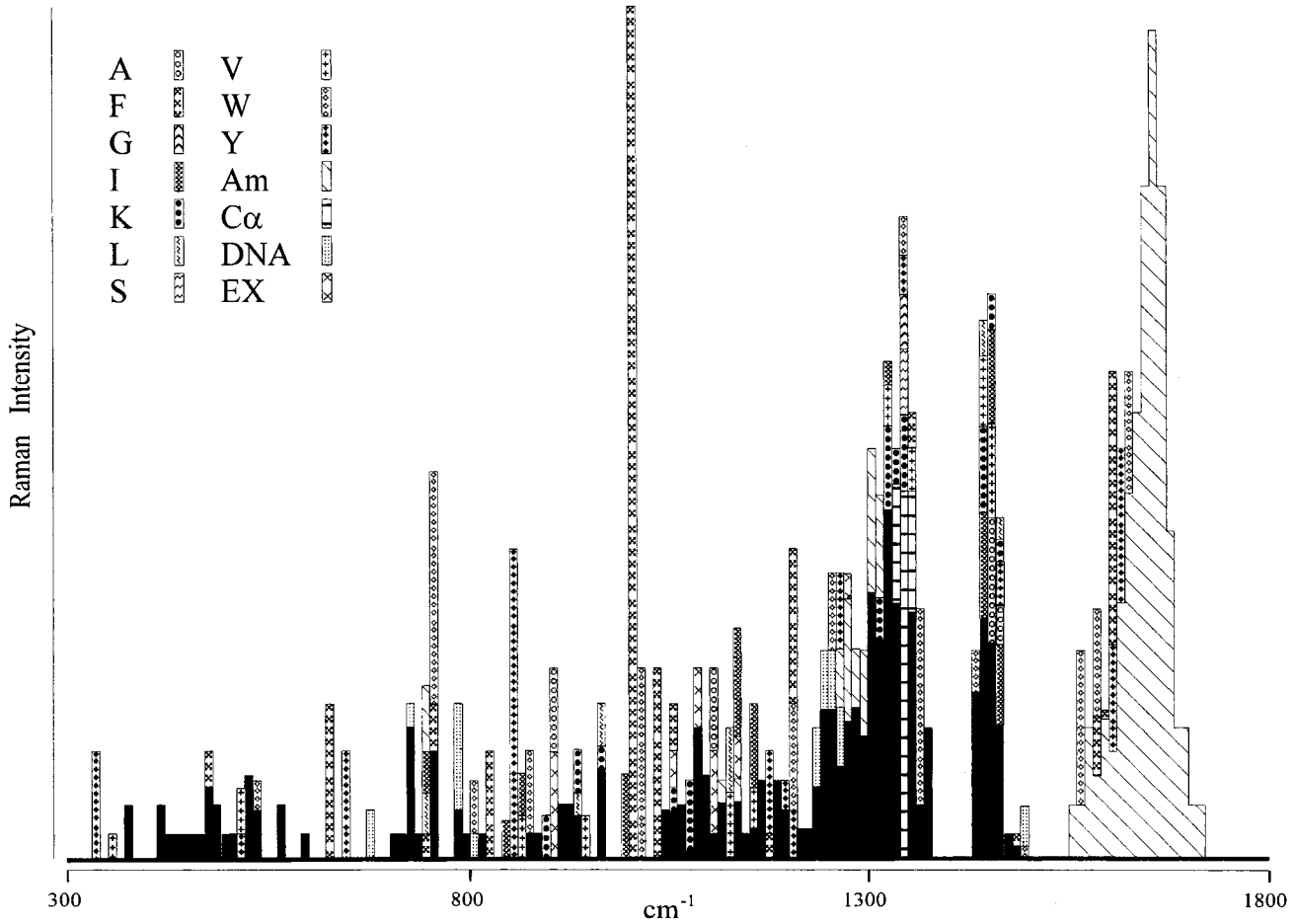

FIGURE 7: Representation of Raman band assignments of *fd* deduced from deuterium labeling studies. Assignments to amino acid side chain modes (one-letter abbreviations), amide modes (Am), the main chain $C^{\alpha}H$ bend and $C^{\alpha}C$ stretch mode (Cα), and DNA modes are indicated by the key at the upper left. The notation EX indicates additional assignments to modes of Gln, Lys, Ser, Thr, Trp, and/or Tyr that are sensitive to deuterium exchange of side chain NH or OH sites (*13*). The black rectangles represent intensity not yet assigned to specific vibrational modes and believed to be due in part to uncompensated solvent.

residues combined is estimated to account for at least 60% of the Raman intensity at 1340 cm^{-1}. [This refines our earlier estimate of an upper limit of 65% (*13*).] Additionally, glycine, lysine, and serine side chains contribute about 10% each and tryptophan and tyrosine side chains approximately 5% each. Thus, this study confirms that the Raman spectrum of *fd* contains no significant amide III intensity near 1340 cm^{-1}, and demonstrates further that the majority (60−65%) of the observed 1340 cm^{-1} intensity originates from the conformation-dependent $C^{\alpha}H$ bend and $C^{\alpha}C$ stretch mode of the subunit main chain. This conclusion is in accordance with that reached earlier from examination of $^{13}C{=}O$ and N-^{2}H isotope-edited Raman spectra of *fd* (*13*).

Finally, we note that many of the Raman bands assigned to specific nonaromatic side chains (Table 3) overlap appreciably with Raman markers assigned to the protein main chain. The extent of such overlap is apparent in Figure 7. The results presented here suggest a need for caution in estimating protein conformations solely on the basis of Raman intensities observed in the so-called amide III region. The problem of band overlap is further accentuated by the fact that, as noted above, nonaromatic side chain contributions in the 1200−1400 cm^{-1} region of the Raman spectrum are sensitive to differences in conformational structure.

ACKNOWLEDGMENT

We thank our colleagues Drs. James M. Benevides (University of Missouri−Kansas City) and Masamichi Tsuboi (Iwaki-Meisei University) for helpful discussions.

REFERENCES

1. Miura, T., and Thomas, G. J., Jr. (1995) in *Subcellular Biochemistry* (Biswas, B. B., and Roy, S., Eds.) Vol. 24, pp 55−99, Plenum Press, New York.
2. Austin, J. C., Jordan, T., and Spiro, T. G. (1993) in *Biomolecular Spectroscopy* (Clark, R. J. H., and Hester, R. E., Eds.) Part A, pp 55−127, Wiley, London.
3. Thomas, G. J., Jr., Prescott, B., and Day, L. A. (1983) *J. Mol. Biol. 165*, 321−356.
4. Aubrey, K. L., and Thomas, G. J., Jr. (1991) *Biophys. J. 60*, 1337−1349.
5. Overman, S. A., Aubrey, K. L., Vispo, N. S., Cesareni, G., and Thomas, G. J., Jr. (1994) *Biochemistry 33*, 1037−1042.
6. Pitzer, K. S., and Kilpatrick, J. E. (1946) *Chem. Rev. 39*, 435−447.
7. Colthup, N. B. (1950) *J. Opt. Soc. Am. 40*, 397−400.
8. Krimm, S., and Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181−365.
9. Cross, T. A., and Opella, S. J. (1981) *Biochemistry 20*, 290−297.
10. Symmons, M. F., Welsh, L. C., Nave, C., Marvin, D. A., and Perham, R. N. (1995) *J. Mol. Biol. 245*, 86−91.

11. Overman, S. A., Tsuboi, M., and Thomas, G. J., Jr. (1996) *J. Mol. Biol. 259*, 331−336.
12. Tsuboi, M., Overman, S. A., and Thomas, G. J., Jr. (1996) *Biochemistry 35*, 10403−10410.
13. Overman, S. A., and Thomas, G. J., Jr. (1998) *Biochemistry 37*, 5654−5665.
14. Overman, S. A., and Thomas, G. J., Jr. (1995) *Biochemistry 34*, 5440−5451.
15. Wen, Z. Q., Overman, S. A., and Thomas, G. J., Jr. (1997) *Biochemistry 36*, 7810−7820.
16. Thomas, G. J., Jr., Prescott, B., Opella, S. J., and Day, L. A. (1988) *Biochemistry 27*, 4350−4357.
17. Thomas, G. J., Jr., and Barylski, J. R. (1970) *Appl. Spectrosc. 24*, 463−464.
18. Li, Y., Thomas, G. J., Jr., Fuller, M., and King, J. (1981) *Prog. Clin. Biol. Res. 64*, 271−283.
19. Snyder, R. G., and Schachtschneider, J. H. (1963) *Spectrochim. Acta 19*, 85−116.
20. Frushour, B. G., Painter, P. C., and Koenig, J. L. (1976) *J. Macromol. Sci., Rev. Macromol. Chem. 15*, 29−115.
21. Lin-Vien, D., Colthup, N. B., Fateley, W. G., and Grasselli, J. G. (1991) *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules*, Academic Press, San Diego, CA.
22. Bellamy, L. J. (1975) *The Infrared Spectra of Complex Molecules*, 3rd ed., Vol. 1, Chapman and Hall, Ltd., London.
23. Yu, T.-J., Lippert, J. L., and Peticolas, W. L. (1973) *Biopolymers 12*, 2161−2176.
24. Lord, R. C., and Yu, N. (1970) *J. Mol. Biol. 50*, 509−524.
25. Day, L. A., Casadevall, A., Prescott, B., and Thomas, G. J., Jr. (1988) *Biochemistry 27*, 706−711.
26. Overman, S. A. (1996) Raman and polarized Raman spectroscopy of the filamentous virus *Ff*: Structural conclusions and a molecular model, Ph.D. Thesis, University of Missouri—Kansas City, Kansas City, MO.
27. Simons, L., Bergstrom, G., Blomfelt, G., Forss, S., Stenback, H., and Wansen, G. (1972) *Commentat. Phys.-Math. 42*, 125−207.

BI982901E