

Distinct Cysteine Sulfhydryl Environments Detected by Analysis of Raman S-H Markers of Cys → Ser Mutant Proteins

Stephen W. Raso¹, Patricia L. Clark¹, Cameron Haase-Pettingell¹
Jonathan King¹ and George J. Thomas, Jr^{2*}

¹Department of Biology
Massachusetts Institute of
Technology, Cambridge
MA 02139, USA

²School of Biological Sciences
University of Missouri-Kansas
City, Kansas City
MO 64110-2499, USA

Very little is known about the character or functional relevance of hydrogen-bonded cysteine sulfhydryl (S-H) groups in proteins. The Raman S-H band is a unique and sensitive probe of the local S-H environment. Here, we report the use of Raman spectroscopy combined with site-specific mutagenesis to document the existence of five distinguishable hydrogen-bonded states of buried cysteine sulfhydryl groups in a native protein. The 666 residue subunit of the *Salmonella typhimurium* bacteriophage P22 tailspike contains eight cysteine residues distributed through the elongated structure. The tailspike cysteine residues display an unusual Raman S-H band complex (2500–2600 cm⁻¹ interval) indicative of diverse S-H hydrogen-bonding interactions in the native trimeric structure. To resolve specific Cys contributions to the complex Raman band we characterized a set of tailspike proteins with each cysteine replaced by a serine. The mutant proteins, once folded, were structurally and functionally indistinguishable from wild-type tailspikes, except for their Raman S-H signatures. Comparison of the Raman spectra of the mutant and wild-type proteins reveals the following hydrogen-bond classes for cysteine sulfhydryl groups. (i) Cys613 forms the strongest S-H...X bond of the tailspike, stronger than any heretofore observed for a protein. (ii) Cys267, Cys287 and Cys458 form robust S-H...X bonds. (iii) Moderate S-H...X bonding occurs for Cys169 and Cys635. (iv) Cys290 and Cys496 form weak hydrogen bonds. (v) It is remarkable that Cys287 contributes two Raman S-H markers, indicating the population of two distinct hydrogen-bonding states. The sum of the S-H Raman signatures of all eight mutants accurately reproduces the composite Raman band of the wild-type tailspike. The diverse cysteine states may be an outcome of the folding and assembly pathway of the tailspike, which though lacking disulfide bonds in the native state, utilizes transient disulfide bonds in the maturation pathway. This Raman study represents the first detailed assessment of local S-H hydrogen bonding in a native protein and provides information not obtainable directly by other structural probes. The method employed here should be applicable to a wide range of cysteine-containing proteins.

© 2001 Academic Press

Keywords: cysteine; sulfhydryl hydrogen bond; P22 tailspike structure; protein folding; Raman spectroscopy

*Corresponding author

Abbreviations used: FTIR, Fourier-transform infrared; H/D, hydrogen/deuterium; IR, infrared.

E-mail address of the corresponding author:
thomasgj@umkc.edu

Introduction

Cysteine thiol or sulfhydryl (S-H) groups are the most chemically reactive sites in proteins at physiological conditions.¹ The oxidation of two cysteine S-H bonds to form the cystine disulfide linkage (S-S) is a key step in the regulation of oxidoreductase chemistry and is important in the stabilization

of many protein native states.² The cysteine thiolate ion (S^-) may function as a potent nucleophile in active-site chemistry of thiol proteases, as well as in reactions with exogenous modifying agents.³ The cysteine sulfhydryl may function as either a hydrogen bond donor (e.g. $S-H \cdots OH_2$) or acceptor (e.g. $HOH \cdots S-H$) group.⁴ Alternatively, the sulfhydryl may release its proton to facilitate sulfur ligation at metal coordination sites in proteins.^{5,6} Sulfhydryls are also vulnerable to aberrant oxidations, which can compromise protein function or stability.^{7,8}

Sulfhydryl hydrogen bonding in proteins is not well understood, primarily because such interactions are difficult to detect experimentally. Owing to the low electron density of hydrogen, $S-H \cdots X$ interactions (where X indicates any suitable hydrogen bond acceptor group) cannot be detected in protein crystal structures. Therefore, the presence or absence of an $S-H \cdots X$ hydrogen bond is usually inferred from the interatomic distance between the atoms S and X. The validity of such an inference is difficult to assess, particularly in view of the greater tendency of S than O (or N) to participate in non-linear hydrogen bonds. Other factors contribute to the paucity of information on protein $S-H \cdots X$ interactions. For example, the feasibility of hydrogen bonding in a protein structure is usually deduced after optimizing the geometries of all covalent bonds except those involving hydrogen; also, electrostatic interactions are typically ignored in the refinement process. Even in cases where hydrogen bond structural constraints are considered, those involving a sulfur center are weighted less heavily than those involving oxygen or nitrogen. Structures refined in such a manner will tend to underestimate the extent of cysteine S-H hydrogen bonding. Although neutron crystallography is, in principle, more effective than X-ray crystallography in locating hydrogen atoms, the neutron technique is little used for proteins because of the requirements of extraordinarily large crystals and time-consuming data collection protocols.

Baburina and co-workers have demonstrated the utility of Fourier transform infrared (FTIR) spectroscopy in a systematic analysis of yeast pyruvate decarboxylase Cys \rightarrow Ser mutants. They were able to use this technique to determine the specific ionization state of each of the four cysteine residues of the protein.⁹

Raman spectroscopy provides a convenient alternative for investigating cysteine sulfhydryl sites in proteins.¹⁰ The Raman spectrum monitors the S-H group directly and is applicable to proteins in both crystalline and solution states. Previous studies of thiol model compounds demonstrate that the sulfhydryl Raman marker (S-H bond stretching vibration) is a highly sensitive probe of local S-H interaction.¹¹⁻¹³ The Raman band diagnostic of the S-H environment occurs within a spectral range ($2500-2600\text{ cm}^{-1}$) that is well removed from all other Raman bands of a protein

($300-1800$ and $2800-3600\text{ cm}^{-1}$ intervals). Spectral interferences from solvent water molecules and from possible overtone or combination bands of a protein are also much less problematic in Raman than in FTIR spectroscopy. Because no other chemical groups contribute to the sulfhydryl region of the Raman spectrum, the method offers an interference-free probe of cysteinyl structure and local environment.^{10,13} In previous applications, the active site of *Escherichia coli* thioredoxin,^{14,15} the dynamics of viral capsid assembly¹⁶ and the redox chemistry of γ II crystallin¹⁷ have been investigated. Here, we exploit Raman spectroscopy to identify and characterize an unexpected range of cysteine S-H environments in the trimeric P22 tailspike protein.

The 666 amino acid subunit of the P22 tailspike contains eight cysteine residues, Cys169, Cys267, Cys287, Cys290, Cys458, Cys496, Cys613 and Cys635,¹⁸ all of which are reduced in the native protein.¹⁹ The X-ray crystal structure shows that the six Cys residues proximal to the N terminus are located in the so-called "parallel β -helix" domain, which is an extended coil-like conformation stabilized by parallel β -strand secondary structure (Figure 1).^{20,21} All sulfhydryl groups of the tailspike are unreactive toward cysteine modifying agents in solution²² and all are inaccessible to water molecules in the crystal structure.²¹ Near the C-terminal end of the protein, the three polypeptide chains wrap around each other to form a unique prism of interdigitated β -strands.²¹ On the C-terminal side of the interdigitated region each subunit forms a five-stranded antiparallel β -sheet, which combine to form a triangular β -sheet prism, as shown in Figure 1.^{23,24} Cysteine residues 613 and 635, from each chain, form a distinctive ring of six cysteine residues in this C-terminal β -sheet prism (Figure 1(a)).

The tailspike folding and subunit assembly pathway is guided by unusual cysteine chemistry: intersubunit disulfide bonds involving a few of the cysteine residues are formed transiently in a trimeric folding intermediate, both *in vivo* and *in vitro*.^{22,25} There is evidence to suggest, but not conclusively demonstrate, that Cys496, Cys613 and Cys635 (Figure 1) are involved in the transient disulfide bond formation.^{25,26,28} The P22 tailspike provides the first and, as yet, only example of a protein that utilizes a cysteine-disulfide linked folding intermediate en route to a completely reduced and highly stable native state.^{25,27}

It is reasonable to assume that cysteine residues that play different roles in the folding process may occupy distinct final environments in the native protein. Here we report that the S-H region of the tailspike Raman spectrum is unusually complex, suggesting several distinct sulfhydryl environments. To determine the specific contribution of each cysteine sulfhydryl to the native-state Raman signature and to deduce the corresponding S-H hydrogen-bonding environments, we have constructed eight single-site Cys \rightarrow Ser mutants and

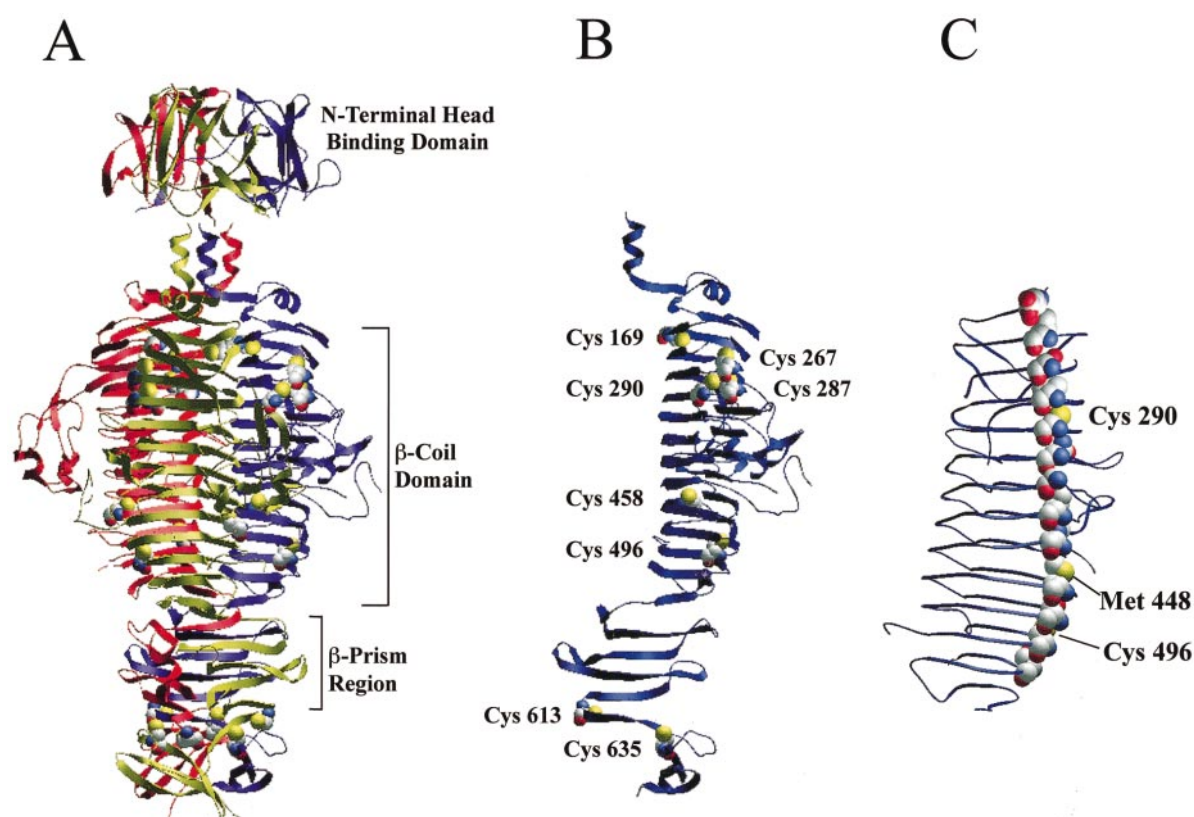


Figure 1. (a) Ribbon diagram of the X-ray crystal structure of the trimeric P22 tailspike protein, showing the N-terminal head-binding domain, the central β -coil (parallel β helix) domain and the C-terminal β -prism domain.^{20,21} Each subunit is shown in a different color. (b) The parallel β -helix and β -prism domains of a single subunit, showing the locations of the eight cysteine residues with space-filling atoms. (c) Locations of Cys290 and Cys496 within a predominantly hydrophobic stack of the parallel β -helix domain. Met448 is also labeled to avoid confusion with cysteine.

systematically analyzed their respective Raman S-H signatures. The Raman analysis reveals the contributions of individual S-H \cdots X hydrogen bonds to the stability of the native state. The results provide information about S-H interactions that currently cannot be obtained by any other structural technique.

Results

Raman signature of the recombinant tailspike protein

The *Escherichia coli*-expressed recombinant tailspike of phage P22 has been fully characterized with respect to composition, morphology and biological activity.²⁶ No differences were apparent between the recombinant tailspike and the tailspike isolated by phage infection of *Salmonella*. The Raman spectrum of the recombinant protein, shown in Figure 2, is indistinguishable from the Raman signature reported previously for the tailspike isolated from P22-infected *Salmonella*.^{19,28} The spectral identity of recombinant and *Salmonella*-produced tailspikes has also been confirmed in the present study (data not shown).

The complex pattern of S-H stretching bands of the tailspike protein, which extends from approximately 2520 to 2600 cm^{-1} , is shown in the inset of Figure 2. The raw data indicate at least four distinct sulfhydryl populations, with corresponding S-H stretching vibrations near 2530, 2550, 2565 and 2585 cm^{-1} . These are identical to the previously reported Raman S-H bands of the *Salmonella*-produced tailspike.²⁸ A comprehensive tabulation of the Raman bands of the tailspike and proposed assignments are given in Table 1.

Characteristics of tailspike sulfhydryl Raman markers

The P22 tailspike protein is highly thermostable, retaining its native structure and biological activity at considerably elevated temperatures.^{19,26,29} We find that the S-H stretching region of the tailspike Raman spectrum (Figure 2, inset) is invariant to changes in temperature between 4 and 70 $^{\circ}\text{C}$ (data not shown). This confirms that native-state interactions of the eight cysteine sulfhydryl groups are stable up to at least 70 $^{\circ}\text{C}$. Abstraction of a cysteine sulfhydryl proton to yield the thiolate anion ($\text{SH} \rightarrow \text{S}^{-} + \text{H}^{+}$) is expected at pH values above

Table 1. Raman bands and assignments for the P22 tailspike

Band (cm ⁻¹)	Relative Intensity ^a	Residue (vibrational) assignment ^b
620	0.7	Phe (ring str)
643	0.8	Tyr (ring str)
650	sh	Met, Cys (C-S str)
667	0.2	Met, Cys (C-S str)
757	2.1	Trp (ring str)
827	1.5	Tyr (ring str)
851	1.8	Tyr (ring str)
878	1.2	Trp (ring str)
888	1.0	-
895	sh	Lys (C-C, C-N str)
929	1.5	al (C-C str, CH ₃ def)
956	0.9	al (C-C str, CH ₃ def)
1002	5.5	Phe (ring str)
1011	sh	Trp (ring str)
1031	1.9	Phe (ring str)
1054	1.2	nar (C-N, C-C str)
1073	1.1	nar (C-N, C-C str)
1099	0.6	Ala (C-C str)
1126	1.3	nar (C-C str), Trp (ring str)
1154	0.6	al (CH ₃ def)
1174	0.7	al (CH ₃ def)
1206	2.3	Tyr (ring str)
1238	4.1	Parallel β -helix (amide III)
1262	3.2	Tyr (ring str), non- β -helix (amide III)
1316	2.7	nar (CH ₂ def)
1340	3.1	nar (CH ₂ def)
1347	2.4	Trp (ring str)
1401	1.9	Asp, Glu (CO ₂ ⁻ str)
1419	2.2	al (CH ₃ def), Gly (CH ₂ def)
1448	4.2	nar (CH ₂ def)
1460	3.7	nar (CH ₂ def)
1550	1.8	Trp (ring str)
1577	sh	Trp (ring str)
1586	2.3	Trp (ring str)
1605	sh	Phe (ring str)
1617	sh	Tyr (ring str)
1667	10.0	Parallel β -helix (amide I)

^a sh, shoulder.^b Assignments are based on studies of related proteins and model compounds. (see Miura & Thomas,¹⁰ Sargent *et al.*¹⁹ and Overman & Thomas,³¹ and citations therein).

Abbreviations: str, stretch; al, unspecified aliphatic side-chain(s); def, deformation; nar, unspecified non-aromatic side-chain(s).

the sulfhydryl pK_a (normally 8.4). The thiolation reaction is accompanied by elimination of the corresponding Raman S-H stretching band.¹⁰ We examined the pH dependence of the Raman S-H profile of tailspike solutions equilibrated in high pH buffers at 20 °C. No significant change in the S-H band pattern could be detected up to pH 10.5. The onset of significant pH-induced change in the Raman spectrum was detected only at pH 11, at which point denaturation was also indicated by substantial broadening of Raman amide I (1667 cm⁻¹) and amide III markers (1238 cm⁻¹). Accordingly, in the native state of the tailspike, all eight cysteine residues exhibit a local pK_a value of at least 10.5, i.e. two pK_a units greater than expected for an exposed cysteine side-chain.

Deuterium exchange of the cysteine sulfhydryl (SH → SD) is manifested by a shift of Raman intensity from the S-H stretching region (2500–2600 cm⁻¹) to the S-D stretching region (1800–1900 cm⁻¹) of the Raman spectrum.¹² We subjected the tailspike to prolonged incubation (48 hours at 4 °C, followed by 12 hours at 25 °C) in ²H₂O solution, but found little evidence of H/D exchange at these conditions.

Resistance of tailspike sulfhydryls to both thiolation and deuteration is consistent with the X-ray crystal structure,²¹ which locates all eight sulfhydryl groups in buried regions of the fold. Conversely, cysteine sulfhydryl groups of *E. coli* thioredoxin are fully titrated at pH 8.5 and rapidly deuterated in ²H₂O solutions.^{14,15} Similar sulfhydryl accessibility is observed for the single cysteine residue of the capsid subunit of bacteriophage P22.¹⁶

Although resistant to denaturation in detergents at physiological temperature, the tailspike protein unfolds irreversibly when heated (90 °C, 30 minutes) in 0.5% (w/v) SDS solution. To assess the effect of such denaturation on the cysteine Raman signature, tailspike was unfolded as described in

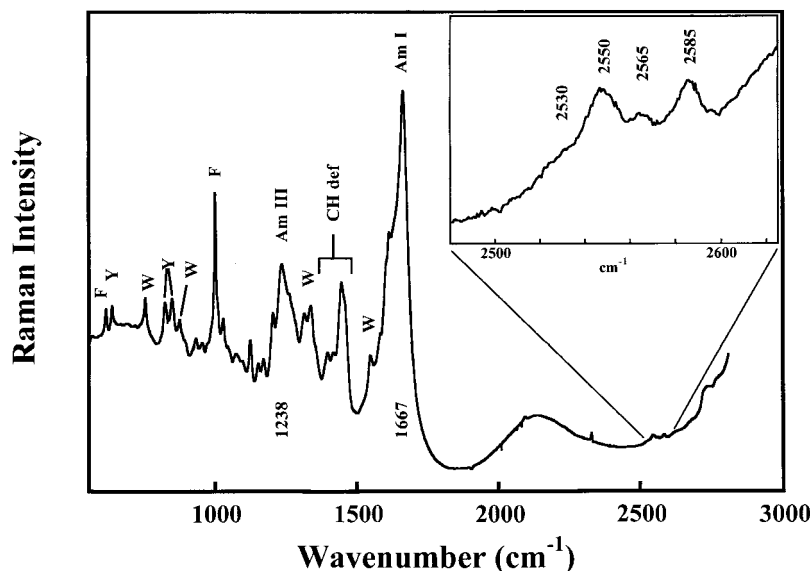


Figure 2. Raman spectrum (670–2800 cm⁻¹) of the P22 trimeric tailspike protein at 10 °C. The tailspike was dissolved at 92 μ g/ μ l in 50 mM Tris, 2 mM EDTA (pH 7.0). The inset at upper right, which shows an amplification of the spectral interval 2480–2620 cm⁻¹, exhibits the composite S-H stretching profile (bands at 2530, 2550, 2565 and 2585 cm⁻¹) of the eight cysteine residues per subunit. The data are uncorrected for contributions of the solvent and scattering background.

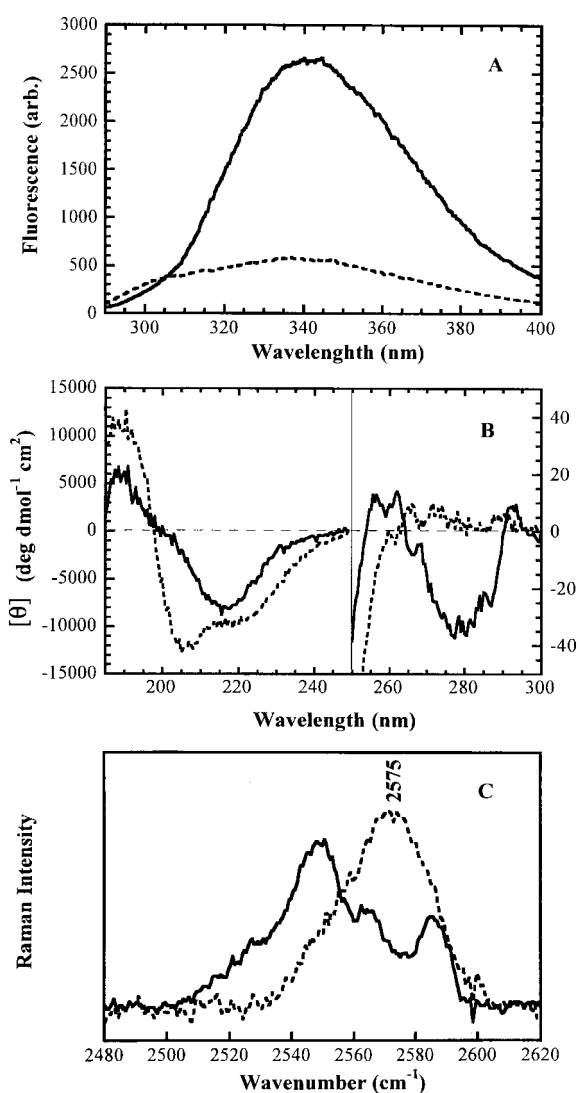


Figure 3. Spectra of native (—) and SDS-denatured (---) tailspike solutions. Native buffer is as given in the legend to Figure 2; SDS-denatured protein was obtained by heating for 30 minutes at 90°C with 0.5% SDS in the native buffer. (a) Fluorescence emission spectra, 250–400 nm (~1 µg/µl concentration). (b) CD spectra in the near-UV, 250–300 nm, and far-UV, 190–250 nm (~1 µg/µl concentration). (c) Raman spectra in the S-H stretching region after baseline correction (~85 µg/µl concentration).

Materials and Methods. Figure 3 compares the fluorescence (a), circular dichroism (b) and Raman sulfhydryl (c) spectra of native and denatured tailspikes. The tryptophan fluorescence (Figure 3(a)) and near-UV CD (Figure 3(b), right panel) profiles of the native tailspike are completely attenuated at the denaturation conditions employed. This demonstrates loss of native structure in tryptophan-containing domains. It is interesting that the far-UV CD spectrum (Figure 3(b), left panel) shows that the native parallel β -helix of the tailspike polypeptide chain is not converted by this

denaturation to a random coil but to a non-native structure that retains substantial α -helix conformation. This reflects the induction of non-native α -helix by the detergent.³⁰ Indeed, the fluorescence and near-UV CD signatures of Figure 3 are indistinguishable from those of tailspike denatured in 5 M urea.

Figure 3(c) shows that tailspike denaturation leads to coalescence of the complex S-H band-shape of the native structure into a single broad, although slightly asymmetric, band centered near 2575 cm^{-1} . This indicates that the multiple S-H environments of the native state are replaced essentially by a common S-H environment in the denatured state. The single broad Raman band of the denatured tailspike represents an average (non-specific) cysteinyl environment of the denatured (random coil + α -helix) polypeptide. It is interesting to note that the Raman S-H signature of L-cysteine in 0.5% SDS solution also consists of a single broad band centered near 2575 cm^{-1} , very similar to that of the denatured tailspike.

Raman spectra of mutant (Cys \rightarrow Ser) tailspikes

Raman spectra of each of the eight single-site Cys \rightarrow Ser mutant tailspikes (C169S, C267S, C287S, C290S, C458S, C496S, C613S, C635S) were obtained at experimental conditions comparable to those employed for the wild-type protein. The Raman signature (600–1800 cm^{-1} interval) of a representative mutant, C169S, is compared with that of the wild-type tailspike in Figure 4. The two spectra are indistinguishable with respect to the key Raman markers of main-chain conformation and side-chain environments. (Note that Cys and Ser side-chains do not contribute Raman bands of appreciable intensity in the 600–1800 cm^{-1} region.³¹ See also Table 1. In addition, electrospray mass spectra indicate no covalent modification other than the Cys \rightarrow Ser mutation.) Figure 4 demonstrates that the native secondary, tertiary and quaternary structures of the tailspike are conserved with the Cys169 \rightarrow Ser mutation. Similar results were obtained for all mutants. The data of Figure 4 are consistent with the results of Haase-Pettingell *et al.*²⁶ showing that each Cys \rightarrow Ser mutant, once folded, exhibits the same biological activity and thermostability as the wild-type protein.

On the other hand, the mutant tailspikes have dramatically different Raman signatures in the region of S-H stretching vibrations (2480–2620 cm^{-1}). Figure 5 compares the Raman S-H profile of the C169S variant with that of the wild-type tailspike. The vertical arrow in Figure 5 identifies the sulfhydryl marker (2565 cm^{-1}) of the Cys169 side-chain, which is present in the wild-type protein but absent from the mutant.

The Raman S-H profiles of the eight Cys \rightarrow Ser mutants are compared with one another and with the wild-type profile in Figure 6(a) (left panel). Corresponding difference spectra in which the wild-type spectrum is subtracted from the mutant

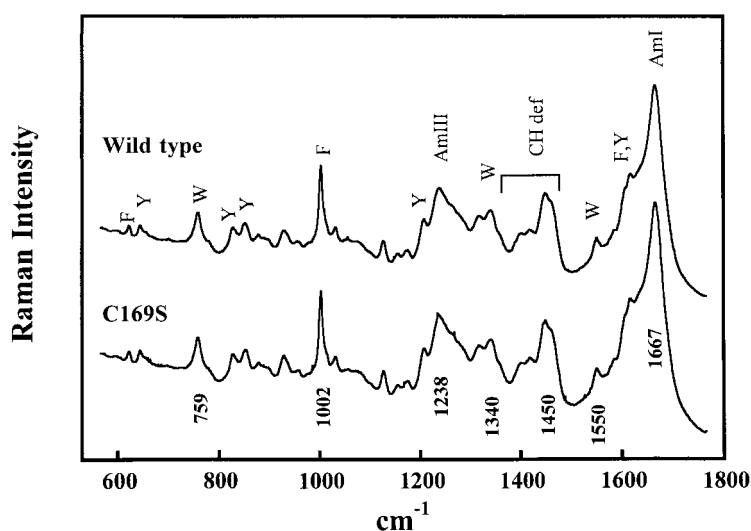


Figure 4. Raman spectra (600–1800 cm^{-1}) of wild-type and C169S-mutant tailspikes under the conditions given for Figure 2. Labels indicate the principal Raman bands and their residue assignments. See also Table 1.

spectrum are shown in Figure 6(b) (right panel). In each case, the Raman signature of the mutated cysteine site appears as a negative difference profile.

Assignment of Raman markers for tailspike cysteine residues

Figure 6(b) identifies a specific Raman difference spectrum for each cysteine sulfhydryl of the tailspike protein. With the exception of Cys290 and Cys496, which produce the same S-H Raman signature (2585 cm^{-1} band), every other cysteine sulfhydryl exhibits a unique Raman signature. The most unusual is Cys287, which consists of *two* comparably intense Raman S-H markers at 2550 and 2585 cm^{-1} . Cys267 may also have a bimodal signature; however, the 2585 cm^{-1} Raman band associated with Cys267 is roughly tenfold weaker

than the dominant 2550 cm^{-1} band and only marginally above the data noise level. The results for Cys267 and Cys287 were repeatedly observed for independently prepared samples of the mutant proteins, all of which exhibited the same native-state Raman signature (600–1800 cm^{-1} region), as well as the electrophoretic and mass spectrometric properties expected for a pure single-site mutant protein without additional covalent modification. Biological activity and thermostability of the C267S and C287S tailspikes were also similar to those of wild-type.²⁶

The deduced Raman signatures of the tailspike sulfhydryl groups differ by virtue of different strengths of S-H \cdots X hydrogen bonding.¹² On the basis of a Raman study of sulfhydryl model compounds in a variety of hydrogen-bonding environments,¹¹ the tailspike cysteine signatures can be correlated with S-H \cdots X hydrogen-bond strength, as shown in Table 2.

It is also of interest that the Raman signatures of the individual mutant proteins (Figure 6(a)) can be summed together to reconstruct the Raman S-H profile of the wild-type tailspike, as shown in Figure 7. This suggests first, that each cysteine sulfhydryl contributes equal Raman intensity to the wild-type profile irrespective of hydrogen bond strength, and second, that the mutation of one cysteine does not alter the hydrogen-bonding environments of other cysteine residues.

Discussion

Raman signature of the native tailspike

The Raman spectrum of the tailspike is atypical. The sulfhydryl region (2500–2600 cm^{-1}) is distinguished by a rich pattern of S-H stretching bands (Figure 2, inset), indicating that cysteine sulfhydryl groups populate a variety of local environments.^{11,12} Conversely, Raman amide I

Table 2. Spectral contributions and hydrogen-bond strengths of cysteine sulfhydryl groups of the native P22 tailspike protein

Residue	Raman S-H band ^a	Hydrogen-bond strength ^b
Cys169	2565	Moderate
Cys267	2550 (90%) 2585 (10%)	Strong Very weak
Cys287	2550 (63%) 2585 (37%)	Strong Very weak
Cys290	2585	Very weak
Cys458	2550	Strong
Cys496	2585	Very weak
Cys613	2530	Very strong
Cys635	2576	Weak

^a Raman band center (in cm^{-1} units). Numbers in parentheses are the percentages of total intensity contributed by the specified cysteine sulfhydryl at the indicated cm^{-1} value.

^b Based upon the results reported by Li & Thomas.¹¹

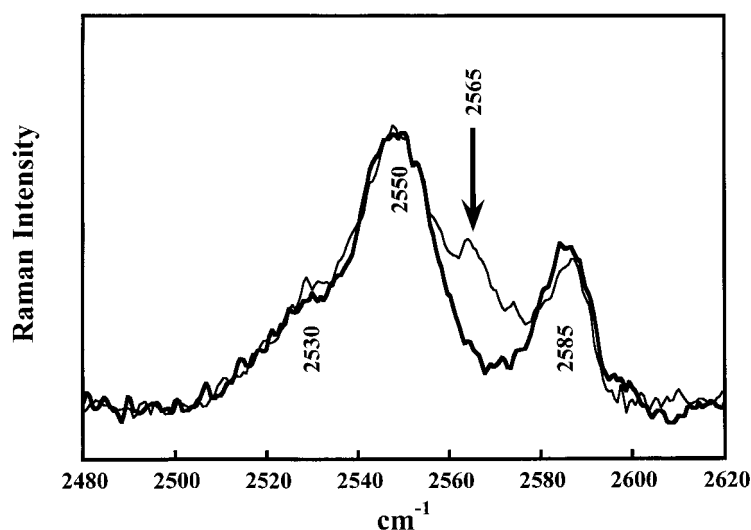


Figure 5. Raman spectra (2480-2620 cm^{-1}) of wild-type (—) and C169S-mutant (---) tailspikes under the conditions given for Figure 2. Labels indicate the major peaks of the wild-type S-H profile. The vertical arrow identifies the S-H Raman signature of Cys169 in the wild-type protein. Spectral traces are baseline-corrected.

(1667 cm^{-1}) and amide III (1238 cm^{-1}) markers of the native tailspike (Figure 2) are remarkably sharp, diagnostic of the prevalent parallel β -strand secondary structure that characterizes the robust parallel

β -helix of the native trimer. These data demonstrate that the complex Raman S-H profile results not from heterogeneity in protein conformation but from diversity in local S-H environments of the native trimer subunit. Consistent with the crystal

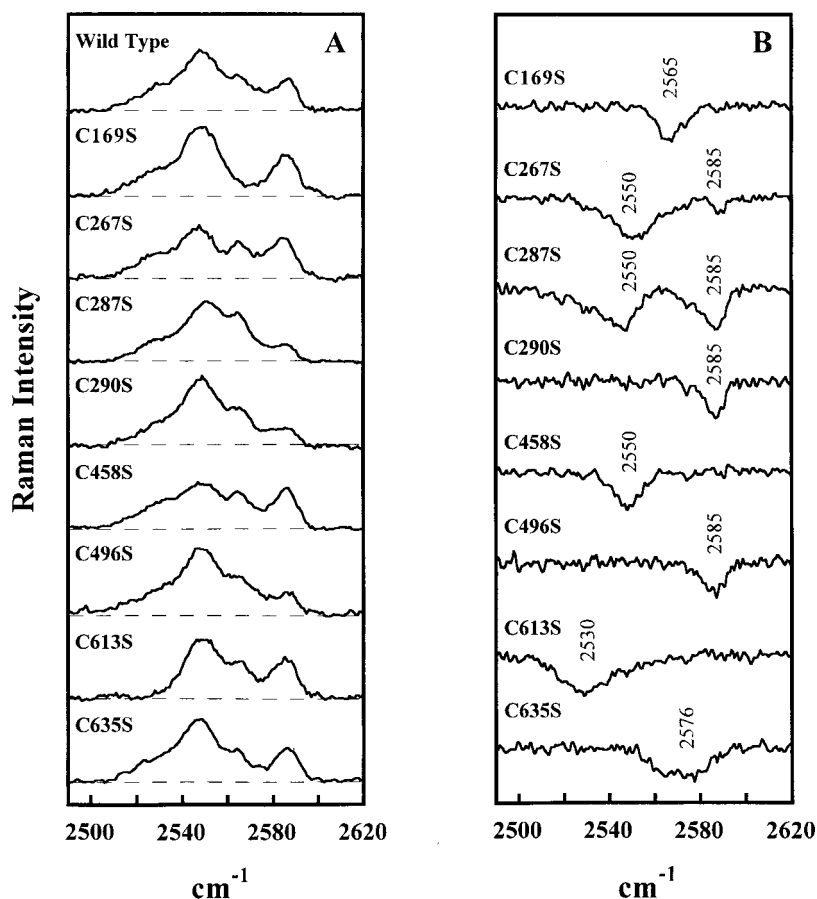


Figure 6. S-H Raman signatures (2480-2620 cm^{-1}) of tailspike cysteine residues. (a) The Raman S-H profiles observed for the wild-type tailspike and for each of the eight Cys \rightarrow Ser mutants, as labeled. (b) Raman difference spectra, computed as mutant minus wild-type, for each of the eight Cys \rightarrow Ser mutants. In each trace, the S-H Raman signature of the mutated Cys site is revealed as a negative spectrum, in accordance with Table 2. Tailspike concentrations range from 81 to 127 $\mu\text{g}/\mu\text{l}$ and other experimental conditions are given in the legend to Figure 5.

structure (Figure 1),²¹ the present experiments also show that all cysteine residues of the native state are solvent inaccessible and well protected from both pH titration and H/D exchange (data not shown). This is in accord with previously demonstrated non-reactivity of the cysteine residues to sulfhydryl reagents.²² It is interesting that all cysteine sulfhydryl groups exhibit comparable protection against pH titration and H/D exchange. This can be attributed to the steric inaccessibility of all cysteine residues to hydroxyl ions and water molecules rather than to comparable strengths of sulfhydryl hydrogen bonding.

Tailspike proteins with each of the eight cysteine residues individually substituted by serine were expressed in *E. coli*. The mutant chains exhibit a range of kinetic defects in their *in vivo* folding and assembly.²⁶ However, sufficient quantities of the native states were recovered and purified to obtain Raman spectra. The Raman conformation markers of the 600-1800 cm^{-1} spectral interval show that recombinant wild-type and all mutant proteins are structurally indistinguishable from phage-derived tailspike throughout the temperature range of the native fold (4-70 °C).^{19,28} Conversely, in the spectral interval of cysteinyl S-H stretching vibrations (2500-2600 cm^{-1}), each mutant variant exhibits a distinctively different Raman signature than wild-type tailspike. Analysis of the Raman S-H data provides insight into the differences in S-H...X local hydrogen-bonding environments among cysteine residues of the native state.

Haase-Pettingell and co-workers have used biological activity and thermal denaturation to also demonstrate that native states of the Cys → Ser mutant and wild-type tailspikes have no significant functional differences.²⁶

Structural significance of cysteine sulfhydryl markers

The combined use of site-directed cysteine mutagenesis and Raman spectroscopy has provided a specific sulfhydryl Raman signature for each of the eight cysteine residues of the native P22 tailspike (Figure 6). On the basis of previously established spectra-structure correlations,^{11,12} the following sequence of S-H...X hydrogen-bond strengths can be assigned to the cysteine sulfhydryl groups (see also Table 2):

Strongest H-bond	Weakest H-bond
Cys613 > Cys458 > Cys267 > Cys287 > Cys169 > Cys635 > Cys290 = Cys496	

This novel approach has yielded information about the protein sulfhydryl environments that cannot be obtained by any other technique. The results are surprising in several respects.

First, S-H...X hydrogen bonding among tailspike cysteine residues covers an extraordinary

range, from the strongest to the weakest extremes of hydrogen bonding observed in model compounds.¹¹

Second, the sulfhydryl hydrogen bond of Cys613 is more robust than encountered in any protein examined previously. Indeed, the Raman marker for Cys613 (2530 cm^{-1}) is indicative of a much stronger sulfhydryl interaction than occurs for either of the two cysteine molecules (2542 and 2556 cm^{-1} markers) in the unit cell of crystalline L-cysteine. Only the crystal of glutathione (2528 cm^{-1} marker) exhibits an S-H...X hydrogen bond of strength comparable to Cys613.¹¹ It is also somewhat surprising that the strong S-H...X interaction of Cys613 is located not in the highly thermostable parallel β -helix but in the C-terminal domain of the tailspike.

Third, we have found that Cys287 significantly populates two different hydrogen-bonding states simultaneously. The state of higher population (~63%) is characterized by strong S-H...X interaction (2550 cm^{-1} marker), whereas the state of lower population (~37%) is characterized by weak interaction (2585 cm^{-1} marker), comparable to that observed in the case of S-H...OH₂. The data obtained for Cys267 also allow for population of two hydrogen bonding states; however, in the case of Cys267, the weaker hydrogen-bonding state is much less favored (~10%) than in the case of Cys287. A precedent for such two-state occupancy by a sulfur-containing side-chain within a native protein structure was reported recently by Petsko and co-workers.³² In the crystal structure of aminopeptidase from *Aeromonas proteolytica* at ultrahigh resolution (~1.0 Å), a methionine side-chain has been found to co-exist in two discrete conformations, in which the methyl thioether moiety is directed toward different hydrophobic partners (G. Petsko, W. Desmarais, R. Holz & D. Ringe, personal communication). Raman spectroscopy, like crystallography, has the capability to resolve such structural polymorphism, as long as the occupancy of each site exceeds the period of the molecular vibration serving as the spectral probe ($\tau \sim 1.3 \times 10^{-14}$ s for the S-H stretch).

The Raman S-H bands of the wild-type tailspike can be very well represented as the sum of four Gauss-Lorentz functions, centered at 2530, 2550, 2565 and 2585 cm^{-1} . In the absence of the

spectroscopic data on the mutants, it might therefore be concluded that no more than four distinct hydrogen-bonding environments exist for the eight sulfhydryl groups. However, analysis of the mutant spectra demonstrates a fifth Raman S-H marker (2576 cm^{-1}), as well as a breakdown of the

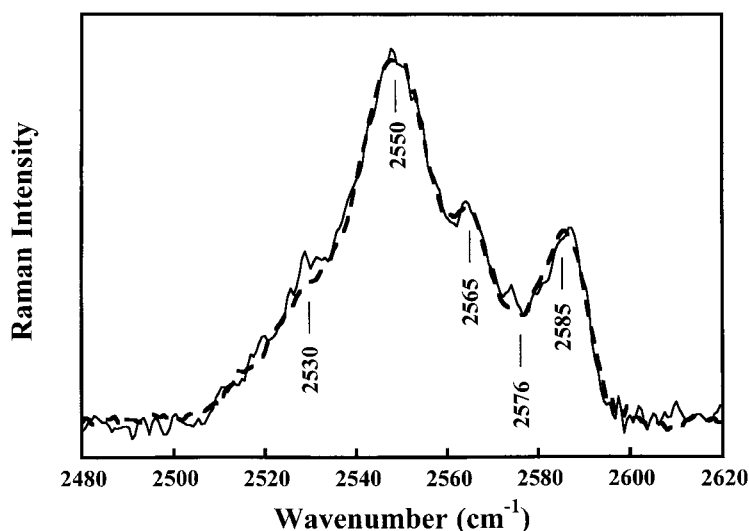


Figure 7. A comparison of the observed Raman spectrum of wild-type tailspike (—) with the normalized sum of observed spectra of the eight Cys → Ser mutants (---) in the region of cysteinyl S-H stretching bands (2480–2620 cm^{-1}). Experimental conditions are given in the legend to Figure 2.

overly simplistic concept “one sulfhydryl moiety = one S-H Raman band.” Nevertheless, the mutant data confirm that all of the tailspike Raman intensity in the 2500–2600 cm^{-1} interval arises only from S-H stretching vibrations (Figure 6), and further, that each sulfhydryl contributes equal intensity to that spectral interval (Figure 7).

Relationship to the tailspike crystal structure

In protein crystal structures cysteine sulfhydryl groups are often chemically modified for the purpose of providing X-ray phasing information. Even if chemically unmodified, a sulfhydryl proton is virtually transparent to X-rays, unless the structure is solved at a resolution considerably better than 1.5 Å.³² In addition, the refinement of an X-ray crystal structure may be undertaken without the imposition of constraints on the coordinates of cysteine sulfhydryl groups. In such a case, the refined crystal structure may not accurately reflect the location of a sulfhydryl in relation to potential hydrogen bond acceptor atoms. For all of these reasons there is a paucity of information on sulfhydryl group interactions, including S-H...X hydrogen bonds, in protein crystal structures. The tailspike structure is no exception. Accordingly, it is not possible to accurately identify the hydrogen-bonding partner (acceptor atom X) for any tailspike cysteine from the existing crystal structure refined to 1.8 Å resolution (Protein Data Bank identifier: 1TYU).³³ Nevertheless, it seems appropriate to examine the published crystal structure as a first approximation to locating plausible acceptor atoms for the S-H...X hydrogen bonds characterized by Raman spectroscopy. Putative hydrogen bonds involving cysteine sulfhydryl groups were predicted using the HBPLUS software package, version 3.06.³⁴ The algorithm involves building hydrogen atoms into the structure using standard bond lengths and angles.

Analysis of the tailspike crystal structure with HBPLUS suggests the following hydrogen-bonding possibilities for the cysteine S-H donor groups (calculated sulfur-to-acceptor distance and bond angle are given in parentheses). Relatively long and presumably weak S-H...O bonds, where O is in each case a main-chain carbonyl oxygen atom, are predicted to occur between Cys287 and Arg285 (3.70 Å, 163.7°), Cys458 and Glu456 (3.69 Å, 167.7°) and Cys635 and Ser633 (3.76 Å, 175.7°). Somewhat shorter and presumably more robust S-H...O bonds are predicted between Cys267 and Arg265 (3.44 Å, 169.0°), Cys290 and Val270 (3.51 Å, 134.5°), and Cys496 and Thr467 (3.47 Å, 130.0°). A very short and apparently very strong hydrogen bond is predicted between Cys613 and His610 (3.25 Å, 147.9°). In every case, the acceptor is a main-chain carbonyl oxygen, except for Cys496/Thr467, where the side-chain O γ atom is the predicted acceptor. From the X-ray structure, HBPLUS predicts the sulfhydryl of Cys169 to have no hydrogen-bonding partner. If it is assumed that the S-H...O distance is the primary determinant of hydrogen-bonding strength, the HBPLUS results imply the following order of decreasing strength of hydrogen bonding: Cys613 > Cys267 ~ Cys496 ~ Cys290 > Cys458 ~ Cys287 > Cys635 > Cys169, which is in partial agreement with the hydrogen-bonding environments determined by Raman spectroscopy (Table 2). The less than perfect agreement may reflect the above noted limitations of current crystallographic methods in assessing S-H...X interactions, influences of bond angle upon the strength of sulfhydryl hydrogen-bonding interaction, and other factors. For example, the hydrogen-bonding states assessed by Raman spectroscopy are expected to reflect cysteine sulfhydryl interactions with acceptor groups that may not be located unambiguously in the crystal

structure (such as oxygen of bound H₂O). This could be the case for Cys169.

It is interesting that the Raman data and X-ray structure are consistent for Cys287 and Cys613. In addition to the Cys287/Arg285 hydrogen bond predicted by the HBPLUS program, the sulfur atom of Cys287 is only 3.31 Å from the carbonyl oxygen of Ile264. This provides at least one good candidate for an alternate hydrogen-bond partner of Cys287, consistent with the observation of two S-H vibrational peaks (Figure 6(b)). The HBPLUS program did not identify a Cys287/Ile264 interaction because it does not predict hydrogen bonds based on the interatomic distance of heteroatoms alone, but by the distance and angle of built-in hydrogen atoms and a putative hydrogen-bond acceptor. The positions of the hydrogen atoms assigned by the program are in the calculated lowest-energy state and may not be accurate.

The Raman S-H marker of Cys613 at 2530 cm⁻¹ (Figure 6(b), Table 2) represents the lowest energy S-H vibration ever observed in either an aqueous protein or thiol solution.^{10,11} In the crystal structure,²¹ Cys613 is located at a tight turn and its sulfhydryl has the potential for two strong hydrogen bonds, one with the carbonyl oxygen of His610

(S-H...O distance of 3.25 Å, as noted above), and the other with the amide nitrogen of Gly615 (S-H...N distance of 3.45 Å). Because the sulfhydryl proton cannot be located experimentally, further mutagenesis would be required to unambiguously identify the Cys613 hydrogen-bonding partner(s).

Roles of cysteine in native tailspike and folding intermediates

The Raman spectra do not reveal whether any serine substituted for cysteine can participate in local hydrogen bonding. Given the high thermal and detergent stability of the trimeric 210,000 Da native tailspike, it is not likely that the loss of three local hydrogen bonds (one per monomer) would seriously perturb the native trimeric state. More than 70 tailspike proteins with single amino acid substitutions have been characterized to date and none exhibits a detectable alteration in the stability of the native state.^{29,35} The loss of the cysteine residues did result in kinetic deficiencies in chain folding or assembly.²⁶ In addition, Haase-Pettingell *et al.*²⁶ were unable to recover native states of the double mutant C613S/C635S. This suggests that loss of more than one of the cysteine residues in the same chain may result in more severe deficiencies in folding or stability.

The most striking difference between cysteine and serine protein chemistry is the capability of the former to form the disulfide bond. Sather & King²² first drew attention to the possible role of disulfide chemistry in tailspike maturation by demonstrating that blockage of S-S bond formation inhibits tailspike production *in vivo*. Subsequently,

Robinson & King²⁵ provided evidence for a transient disulfide-crosslinked folding intermediate *in vitro*. Several additional lines of evidence suggest that the three C-terminal cysteine residues may be involved in transient S-S bond formation; however, the most likely candidates are Cys613 and Cys635.^{22,25,27}

The chemical pathway for the oxidation or reduction of the transient disulfide bonds has not been elucidated. Because there is no net change in the redox state of the protein from the beginning to the end of the folding and assembly process, Haase-Pettingell *et al.*²⁶ suggested that the reactions might involve intramolecular oxidation and reduction. If that were the case, the strongly hydrogen-bonded cysteine residues might reflect the outcome of the intramolecular reduction of the disulfide bonds in the protrimmer.

Although many parallel β-helical proteins contain native disulfide bonds, it is not common to find them playing a dominant structural role in the β-helix itself.³⁶⁻³⁸ However, two recently reported structures of insect antifreeze proteins show extensive and regular disulfide bonding within a β-helix domain.^{39,40} In one instance, S-S bonds are propagated from rung-to-rung throughout the motif,³⁹ while in the other, many regularly spaced S-S bonds extend across the protein, directly through the hydrophobic core.⁴⁰ Because these antifreeze proteins perform the highly specialized task of binding ice crystals, it seems likely that the regular network of S-S bonds does not serve primarily to stabilize the parallel β-helix, but to provide proper spacing and rigidity for function.

Stacks of chemically similar amino acid side-chains, hydrophobic, aromatic or polar, are commonly observed in parallel β-helices.^{36,37,41} These stacks extend from rung-to-rung throughout the parallel β-helix motif and can be found usually within the hydrophobic core or, in the case of polar stacks, on the surface of the protein. Cysteine sulfhydryl groups are often found participating in either hydrophobic or polar stacks.^{36,42,43} Unlike serine, cysteine can make polar interactions and easily exist in a hydrophobic environment; therefore, cysteine is better suited to form the core stacks that may stabilize parallel β-helices. Based on the S to X distances reported in the tailspike crystal structure, most of the cysteine residues appear to be making polar interactions; however, Cys290 and Cys496 also appear to reside within a stack of non-polar residues (Figure 1(c)). The Raman markers for Cys290 and Cys496 are both at 2585 cm⁻¹, which would indicate either very weak hydrogen bonding (interaction with solvent water) or an environment devoid of hydrogen bonding (completely non-polar).¹¹ We can conclude either that their hydrogen-bonding partners within the protein core represent S-H...X interactions of similar strength and geometry to those of S-H...OH₂ interaction, or that Cys290 and Cys496 exist in a completely non-polar environment.

Despite the unique capability of Raman spectroscopy to probe cysteine sulfhydryl chemistry, the technique has been underutilized for this purpose. Here, the results reveal a distinctive Raman S-H signature for each cysteinyl side-chain of the native tailspike, diagnostic of the strength of sulfhydryl hydrogen bonding. The present investigation demonstrates that an extraordinarily broad range of S-H...X interactions is accessible to a native protein. Not all of these are evident from the crystal structure. The data show further that diversity in S-H hydrogen bonding can apply to a single cysteinyl side-chain in a native protein: Cys287 of the tailspike simultaneously and significantly populates two distinct hydrogen-bonding states in the native structure. A complex sulfhydryl band pattern in the Raman spectrum of a protein is not unique to the P22 tailspike.^{14,44,45} The methodology developed here should be applicable to other proteins containing multiple sulfhydryl groups.

Materials and Methods

Tailspike protein and reagents

P22-derived tailspike protein was isolated from phage infection of *Salmonella typhimurium*, using a previously published method.⁴⁶ Recombinant tailspike proteins, including wild-type and the eight Cys → Ser mutants, were overexpressed in *E. coli* and purified as described by Haase-Pettingell *et al.*²⁶ SDS, EDTA, Tris, mono and dibasic phosphate salts, L-cysteine and deuterium oxide (99.9% ²H₂O) were purchased from Sigma Chemical Co. (St. Louis, MO). Centricon protein concentrators were obtained from Millipore (Bedford, MA).

Sample preparations for Raman spectroscopy

For Raman spectroscopy tailspike variants were concentrated to approximately 50–100 µg/µl in Tris buffer (50 mM Tris (pH 7.5), 2 mM EDTA) by centrifugation using Centricon-30 (30,000 M_r cutoff) protein concentrators. Precise protein concentrations were determined spectrophotometrically using the extinction coefficient at 278 nm, $\epsilon_{278} = 0.983 \text{ mg}^{-1} \text{ cm}^{-1} \text{ ml}^{18}$.

Deuterium exchange (SH → SD) of tailspike cysteine residues was attempted at several pD values in the range 5.3 < pD < 10. Samples were exchanged into deuterium buffer (20 mM phosphate in ²H₂O at the desired pD) by centrifugation through a pre-packed MicroSpin G-25 column (Pharmacia) equilibrated with ²H₂O buffer. A 20 µl aliquot of the aqueous sample was loaded onto the column and centrifuged for one minute at 800 g in a microfuge at 4 °C. The newly exchanged sample was immediately sealed in a capillary cell for Raman analysis. Efficiency and purity of buffer exchange were confirmed by the absence of O-H stretching bands (3200–3600 cm⁻¹) in the Raman spectrum.

Hydroxyl titration of tailspike sulfhydryl groups was attempted by equilibrating the tailspike in a 20 mM phosphate buffer system up to pH 11. Raman spectra were recorded immediately after pH elevation.

Denatured tailspike was also prepared for Raman analysis. An 0.5 ml aliquot of tailspike protein solution (12.4 µg/µl in Tris buffer containing 0.5% SDS) was heated to 90 °C for 30 minutes. Subunit dissociation was

demonstrated by SDS-PAGE and unfolding was assessed by CD and fluorescence spectroscopy. After cooling, the protein was concentrated to 73.8 µg/µl for Raman analysis.

Circular dichroism and fluorescence spectroscopy

CD spectra were recorded on an Aviv 60DS spectrometer with spectral resolution of 0.5 nm and averaging time of five seconds. Sample cell paths were 0.1 mm and 1.0 cm for the far and near-UV regions, respectively. Fluorescence emission spectra were obtained on a Hitachi F-4500 spectrometer using 280 nm excitation and a slit width of 1 nm. Emission (290–400 nm) was measured with a scan speed of 60 nm/minute and slit width of 2.5 nm. Samples for CD and fluorescence spectroscopy were prepared at 1 mg/ml and data were collected at 20 °C.

Raman spectroscopy

Tailspike solutions were sealed in 1.0 mm glass capillary tubes (Kimax 34507), which were mounted in a thermostatable jacket (10 °C) in the sample compartment of the Raman spectrometer. Spectra were excited at 514.5 nm using an argon laser (model Innova 70, Coherent, Inc., Santa Clara, CA) and were collected on a single grating spectrometer (Spex 500 M, Instruments S. A., Edison, NJ) equipped with a liquid-nitrogen cooled charge-coupled device detector. The effective spectral resolution is 3.5 cm⁻¹. Raman frequencies were calibrated to ±1 cm⁻¹ with indene and CCl₄ liquids. Laser power at the sample was maintained at or below 200 mW. The total acquisition time for each spectrum was about 25 minutes. Other details of the Raman instrumentation have been described by Movileanu *et al.*⁴⁷

Data were collected in the spectral intervals 380–1922 cm⁻¹ and 1380–2760 cm⁻¹ with acquisition times of 25 minutes per interval. Spectral analyses, including baseline corrections and generation of difference spectra, were carried out using Grams software (Galactic, Salem, NH) and procedures described by Movileanu *et al.*⁴⁷

Mass spectrometry

Electrospray mass spectrometry was performed with a Perkin-Elmer Sciex 365 instrument at the MIT Biopolymers Facility. For mass analysis tailspike samples were prepared at approximately 10 µM in 50% (v/v) acetonitrile, 0.1% (v/v) formic acid solution.

Acknowledgments

Support of this research by NIH grants GM50776 (to G.J.T.) and GM17980 (to J.K.) and by Amgen, Inc. is gratefully acknowledged. S.W.R. and P.L.C. thank Drs James M. Benevides (UMKC) and Stacy A. Overman (UMKC) for their assistance with data collection and analyses and recommendations to improve the text. We also thank Dr Marilyn D. Yoder (UMKC) for reviewing the manuscript prior to submission and carrying out calculations with HBPLUS (MacDonald & Thornton, 1994).

References

- Creighton, T. E. (1993), *Proteins*, 2nd edit., W.H. Freeman & Co., New York.
- Darby, N. & Creighton, T. E. (1995). Disulfide bonds in protein folding and stability. *Methods Mol. Biol.* **40**, 219-252.
- Shaw, E. (1990). Cysteinyl proteinases and their selective inactivation. *Advan. Enzymol. Relat. Areas Mol. Biol.* **63**, 271-347.
- Burley, S. K. & Petsko, G. A. (1988). Weakly polar interactions in proteins. *Advan. Protein Chem.* **39**, 125-189.
- McCall, K. A., Huang, C. & Fierke, C. A. (2000). Function and mechanism of zinc metalloenzymes. *J. Nutr.* **130**, 1437S-1446S.
- Sticht, H. & Rosch, P. (1998). The structure of iron-sulfur proteins. *Prog. Biophys. Mol. Biol.* **70**, 95-136.
- Konz, J. O., King, J. & Cooney, C. L. (1998). Effects of oxygen on recombinant protein expression. *Biotechnol. Prog.* **14**, 393-409.
- Stadtman, E. R. (1988). Protein modification in aging. *J. Gerontol.* **43**, B112-B120.
- Baburina, I., Moore, D. J., Volkov, A., Kahyaoglu, A., Jordan, F. & Mendelsohn, R. (1996). Three of four cysteines, including that responsible for substrate activation, are ionized at pH 6.0 in yeast pyruvate decarboxylase: evidence from Fourier transform infrared and isoelectric focusing studies. *Biochemistry*, **35**, 10249-10255.
- Miura, T. & Thomas, G. J., Jr (1995). Raman spectroscopy of proteins and their assemblies. *Subcell. Biochem.* **24**, 55-99.
- Li, H. & Thomas, G. J., Jr (1991). Cysteine conformation and sulfhydryl interactions in proteins and viruses. 1. Correlation of the Raman S-H band with hydrogen bonding and intramolecular geometry in model compounds. *J. Am. Chem. Soc.* **113**, 456-462.
- Li, H., Wurry, C. J. & Thomas, G. J., Jr (1992). Cysteine conformation and sulfhydryl interactions in proteins and viruses. 2. Normal coordinate analysis of the cysteine side-chain in model compounds. *J. Am. Chem. Soc.* **114**, 7463-7469.
- Tuma, R., Vohník, S., Li, H. & Thomas, G. J., Jr (1993). Cysteine conformation and sulfhydryl interactions in proteins and viruses. 3. Quantitative measurement of the Raman S-H band intensity and frequency. *Biophys. J.* **65**, 1066-1072.
- Li, H., Hanson, C., Fuchs, J. A., Woodward, C. & Thomas, G. J., Jr (1993). Determination of the pK_a values of active-center cysteines, cysteines-32 and -35, in *Escherichia coli* thioredoxin by Raman spectroscopy. *Biochemistry*, **32**, 5800-5808.
- Vohník, S., Hanson, C., Tuma, R., Fuchs, J. A., Woodward, C. & Thomas, G. J., Jr (1998). Conformation, stability, and active-site cysteine titrations of *Escherichia coli* D26A thioredoxin probed by Raman spectroscopy. *Protein Sci.* **7**, 193-200.
- Tuma, R., Prevelige, P. E., Jr & Thomas, G. J., Jr (1998). Mechanism of capsid maturation in a double-stranded DNA virus. *Proc. Natl Acad. Sci. USA*, **95**, 9885-9890.
- Pande, J., McDermott, M. J., Callender, R. H. & Spector, A. (1989). Raman spectroscopic evidence for a disulfide bridge in calf γ II crystallin. *Arch. Biochem. Biophys.* **269**, 250-255.
- Sauer, R. T., Krovatin, W., Poteete, A. R. & Berget, P. B. (1982). Phage P22 tail protein: gene and amino acid sequence. *Biochemistry*, **21**, 5811-5815.
- Sargent, D., Benevides, J. M., Yu, M. H., King, J. & Thomas, G. J., Jr (1988). Secondary structure and thermostability of the phage P22 tailspike. XX. Analysis by Raman spectroscopy of the wild-type protein and a temperature-sensitive folding mutant. *J. Mol. Biol.* **199**, 491-502.
- Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R. & Huber, R. (1997). Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *J. Mol. Biol.* **267**, 865-880.
- Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. & Reinemer, P. (1994). Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science*, **265**, 383-386.
- Sather, S. K. & King, J. (1994). Intracellular trapping of a cytoplasmic folding intermediate of the phage P22 tailspike using iodoacetamide. *J. Biol. Chem.* **269**, 25268-25276.
- Kreisberg, J. F., Betts, S. D. & King, J. (2000). β -Helix core packing within the triple-stranded oligomerization domain of the P22 tailspike. *Protein Sci.* **9**, 2338-2343.
- Seckler, R. (1998). Folding and function of repetitive structure in the homotrimeric phage P22 tailspike protein. *J. Struct. Biol.* **122**, 216-222.
- Robinson, A. S. & King, J. (1997). Disulfide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein. *Nature Struct. Biol.* **4**, 450-455.
- Haase-Pettingell, C., Betts, S. D., Raso, S. W., Stuart, L., Robinson, A. S. & King, J. (2001). The role for cysteine residues in the *in vivo* folding and assembly of the phage P22 tailspike. *Protein Sci.* **10**, 397-410.
- Betts, S. & King, J. (1999). There's a right way and a wrong way: *in vivo* and *in vitro* folding, misfolding and subunit assembly of the P22 tailspike. *Structure Fold. Des.* **7**, R131-R139.
- Thomas, G. J., Jr, Becka, R., Sargent, D., Yu, M. H. & King, J. (1990). Conformational stability of P22 tailspike proteins carrying temperature-sensitive folding mutations. *Biochemistry*, **29**, 4181-4187.
- Sturtevant, J. M., Yu, M. H., Haase-Pettingell, C. & King, J. (1989). Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. *J. Biol. Chem.* **264**, 10693-10698.
- Parker, W. & Song, P. S. (1992). Protein structures in SDS micelle-protein complexes. *Biophys. J.* **61**, 1435-1439.
- Overman, S. A. & Thomas, G. J., Jr (1999). 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. *Biochemistry*, **38**, 4018-4027.
- Desmarais, W., Ringe, D. & Petsko, G. A. (2000). The trouble with hydrogen. *FASEB J.* **14**, A1581.
- Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R. & Huber, R. (1996). Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc. Natl Acad. Sci. USA*, **93**, 10584-10588.
- MacDonald, K. A. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777-793.
- Haase-Pettingell, C. & King, J. (1997). Prevalence of temperature sensitive folding mutations in the

- parallel β -coil domain of the phage P22 tailspike endorhamnosidase. *J. Mol. Biol.* **267**, 88-102.
36. Jenkins, J., Mayans, O. & Pickersgill, R. (1998). Structure and evolution of parallel β -helix proteins. *J. Struct. Biol.* **122**, 236-246.
37. Petersen, T. N., Kauppinen, S. & Larsen, S. (1997). The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel β -helix. *Structure*, **5**, 533-544.
38. Yoder, M. D., Lietzke, S. E. & Jurnak, F. (1993). Unusual structural features in the parallel β -helix in pectate lyases. *Structure*, **1**, 241-251.
39. Graether, S. P., Kuiper, M. J., Gagne, S. M., Walker, V. K., Jia, Z., Sykes, B. D. & Davies, P. L. (2000). β -Helix structure and ice binding properties of a hyperactive antifreeze protein from an insect. *Nature*, **406**, 325-328.
40. Liou, Y.-C., Tocilj, A., Davies, P. L. & Jia, Z. (2000). Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. *Nature*, **406**, 322-324.
41. Yoder, M. D. & Jurnak, F. (1995). Protein motifs. 3. The parallel β -helix and other coiled folds. *FASEB J.* **9**, 335-342.
42. Huang, W., Matte, A., Li, Y., Kim, Y. S., Linhardt, R. J., Su, H. & Cygler, M. (1999). Crystal structure of chondroitinase B from *Flavobacterium heparinum* and its complex with a disaccharide product at 1.7 Å resolution. *J. Mol. Biol.* **294**, 1257-1269.
43. Raetz, C. R. & Roderick, S. L. (1995). A left-handed parallel β -helix in the structure of UDP-*N*-acetylglucosamine acyltransferase. *Science*, **270**, 997-1000.
44. Pande, J., McDermott, M. J., Callender, R. H. & Spector, A. (1991). The calf γ -crystallins: a Raman spectroscopic study. *Exp. Eye Res.* **52**, 193-197.
45. Rodriguez-Casado, A., Moore, S., Prevelige, P. E., Jr & Thomas, G. J., Jr (2000). Structural properties of the portal protein subunit of bacteriophage P22 in relation to quaternary assembly. *Biophys. J.* **78**, 171A.
46. Yu, M.-H. & King, J. (1988). Surface amino acids as sites of temperature-sensitive folding mutations in the P22 tailspike proteins. *J. Biol. Chem.* **263**, 1424-1431.
47. Movileanu, L., Benevides, J. M. & Thomas, G. J., Jr (1999). Temperature dependence of the Raman spectrum of DNA. I. Raman signatures of premelting and melting transitions of poly(dA-dT)-poly(dA-dT). *J Raman Spectrosc.* **30**, 637-649.

Edited by P. E. Wright

(Received 13 October 2000; received in revised form 12 January 2001; accepted 15 January 2001)