

Conformation Dependence of the SH and CS Stretch Frequencies of the Cysteine Residue

WEILI QIAN and SAMUEL KRIMM

Biophysics Research Division and Department of Physics, University of Michigan, Ann Arbor, Michigan 48109

SYNOPSIS

In order to relate the observed SH and CS stretch frequencies of the cysteine residue in proteins more closely to its conformation, we have done normal mode calculations on a model for this structure, viz., (CCONH)(CNHCO)CHCH₂SH. A range of χ^1 and χ^2 were studied, combined with the ϕ, ψ of the α -helix, β -sheet, glutathione, and extended-helix conformations. The force field was a combination of a scaled ab initio force field of the -CH₂SH group, obtained from ethanethiol and tested on 1-propanethiol and 3-thiol-N-methylpropionamide, and our empirical force field for the peptide group. The results provide more detailed structure-spectra correlations than are possible from experimental studies of model compounds. © 1992 John Wiley & Sons, Inc.

INTRODUCTION

The cysteine residue side chain, -CH₂SH, is an important one in proteins, both because of its inherent properties as well as its ability to react with similar groups to form a disulfide bridge, -CH₂SSCH₂-, between polypeptide chains. It is therefore important to develop detailed spectroscopic correlations for characterizing its conformation.

It has long been known that the SH stretch (s) mode, $\nu(\text{SH})$, is generally found in the range of 2500–2600 cm⁻¹, being weak in the ir and strong in the Raman spectrum. Since $\nu(\text{SH})$ is sensitive to the presence of SH groups and to their environment, particularly hydrogen bonding, it has been studied in various proteins, such as hemoglobin,^{1,2} eye lens proteins,^{3–5} β -lactoglobulin,⁶ and virus proteins.^{7,8} Conformational information with respect to the C ^{α} -C ^{β} -S-H dihedral angle (χ^2) has been sought through experimental studies of model alkanethiols^{9–11} as well as normal mode calculations on such molecules.^{12,13} Low-frequency torsional modes have also been studied^{14–16} in order to obtain information on rotational barriers.

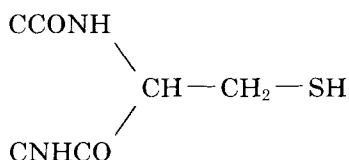
The $\nu(\text{CS})$ mode has been known to give rise to Raman bands in the 600–800-cm⁻¹ region. Confor-

mational assignments with respect to the C-C ^{α} -C ^{β} -S dihedral angle (χ^1) have been based on experimental studies of thiols,^{17,18} dialkyl disulfides,^{19,20} and alkyl sulfides,²¹ as well as normal mode calculations on these molecules.^{22,23}

While these experimental studies on simple model compounds have provided an important base for categorizing conformational states, the results cannot be satisfactorily transferred to a detailed understanding of the structure-spectra correlations in proteins, where, as in the case of the disulfide bridge,^{24–26} the conformations of the adjoining peptide groups are probably important. To achieve such insights requires normal mode analyses with a reliable force field on the relevant structures. In preparing for a vibrational study of glutathione,²⁷ we did not feel that current empirical alkanethiol force fields^{12,13} were satisfactory, in part because some of the band assignments in these studies were questionable. Nor did a force field developed for the amino acid cysteine²⁸ seem suitable, since it was based in part on the alkanethiol force field¹³ and in any case only considered the residue conformation in the crystal.

We have, therefore, by analogy with the disulfide case,^{29–31} obtained an ab initio force field for the -CH₂SH group. This was scaled using the experimental data on ethanethiol^{9,16} (ET), and was tested on 1-propanethiol^{12,17} (1PT), and 3-thiol-N-

methylpropionamide³² (3TNMP), $\text{CH}_3\text{NHCOCH}_2\text{CH}_2\text{SH}$, in the latter case using our ab initio force field for hydrogen-bonded N-methylacetamide³³ (NMA). The resulting satisfactory agreement justified using the $-\text{CH}_2\text{SH}$ force field, together with our empirical force field for the peptide group,³⁴ to calculate the $\nu(\text{SH})$ and $\nu(\text{CS})$ modes of a model for the cysteine residue in proteins,



as a function of χ^2, χ^1 , and the ϕ, ψ of the adjacent peptide groups (ϕ and ψ are the $\text{CNC}^{\alpha}\text{C}$ and $\text{NC}^{\alpha}\text{CN}$ dihedral angles, respectively). The results provide greater insight into exactly how these frequencies depend on the complete conformation of the cysteine residue.

AB INITIO FORCE FIELD OF CH_2SH GROUP

As noted above, we have based the force field of the $-\text{CH}_2\text{SH}$ group on the scaled ab initio force field of ET. The calculational procedures were similar to those used for the disulfide group.³⁰ The structure is determined by χ^2 , which is known to be $\sim 60^\circ$ (*gauche*, G) and $\sim 180^\circ$ (*trans*, T) for the two stable conformers.⁹ However, since χ^2 is more variable in proteins, we have obtained the force field not only at the above minima but also at the energy maxima of $\chi^2 \sim 0^\circ$ (*cis*, C) and $\sim 120^\circ$ (*skew*, S). This provides, as in the case of the disulfide group,³¹ a conformation-dependent force field.

The restricted Hartree-Fock calculations were done using Gaussian 86 with a 3-21G* basis set, the same basis set as was used for the disulfide calculations.³⁰ (A calculation was also done with a 4-31G* basis set, but the results were not significantly different.) Force constants were calculated in cartesian coordinates by the analytical method at the optimized C, G, S, and T geometries. They were then transformed into internal coordinates using a numerical method for determining the redundant coordinates.³⁰ Force constants for negative χ^2 , G' and S', can easily be obtained from their positive counterparts.²⁴ Since force constants obtained with such limited basis sets are generally too high, we have scaled them according to a standard procedure.³⁵ Because of the similarity of parts of the ET molecule

to the disulfide group, we initially transferred related scale factors from the latter³⁰ and subsequently refined all scale factors to improve agreement with the experimental data.

The structures of the four conformers are shown in Figure 1 and their optimized geometric parameters are given in Table I. A dihedral angle of $\chi^2 = 64.8^\circ$ has been obtained for the G conformer from analysis of higher level torsional transitions,¹⁶ which compares very well with our calculated value of 64.5° . The relative energies, in kcal/mole, are G = 0, T = 0.317, S = 1.619, and C = 1.906. (The 4-31G* relative energies are G = 0, T = 0.546, S = 1.666, and C = 2.054.) This leads to $\Delta E(\text{G-T}) = 0.317$ (compared to an experimental value of 0.336¹⁶) and barriers of $\text{G} \rightarrow \text{T} = 1.619$ (1.410¹⁶), $\text{T} \rightarrow \text{G} = 1.302$ (1.028¹⁶), and $\text{G} \rightarrow \text{G}' = 1.906$ (1.764¹⁶), which are all in acceptable agreement with experiment.

The internal and local symmetry coordinates are defined in Tables II and III, respectively. The band assignments for the G conformer of gaseous ET are quite secure,^{9,13-16} and therefore the refinement of scale factors was straightforward; these are given in Table IV. The scaled ab initio force field (for force constants ≥ 0.03) is given in Table V. The calculated frequencies based on the full ab initio force field, with their potential energy distributions (PED), are given in Table VI for the G and T conformers. They are compared with observed gas-phase

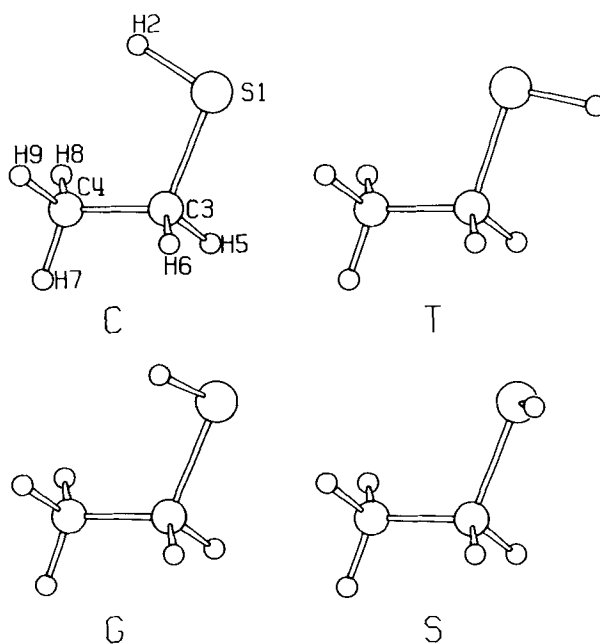


Figure 1. Conformations and atom numberings of C, T, G, and S conformers of ET.

Table I Optimized Geometric Parameters of Ethanethiol Conformers

Parameter ^a	Conformer			
	G	T	S	C
CC	1.538	1.541	1.539	1.539
CS	1.826	1.829	1.838	1.838
SH	1.328	1.327	1.325	1.326
CH5	1.082	1.081	1.080	1.080
CH6	1.081	1.081	1.081	1.080
CH7	1.085	1.084	1.085	1.084
CH8	1.082	1.083	1.083	1.083
CH9	1.084	1.083	1.082	1.083
CSH	97.4	97.9	98.1	98.0
CCS	113.5	109.2	111.9	113.4
SCH5	105.1	109.4	107.9	107.5
SCH6	109.3	109.4	108.4	107.6
CCH7	110.1	109.8	110.0	109.8
CCH8	110.5	110.7	110.7	110.8
CCH9	110.8	110.7	110.5	110.8
CCSH	64.5	181.5	119.8	0.7
H5CSH	-175.1	-58.1	-118.6	122.4
H6CSH	-59.8	61.1	-2.0	-121.1
H7CCS	176.7	180.5	179.4	179.9
H8CCS	56.4	60.7	59.7	60.3
H9CCS	-63.7	-59.8	-60.3	-60.5

^a Bond lengths in Å, bond angles in degrees. Numbered H atoms refer to Figure 1.

frequencies^{9,14-16} and with condensed phase data.⁹ Calculated ir and Raman intensities are also given.

The agreement between observed gas-phase frequencies and those calculated for the G conformer is excellent. The one exception is the CH₂ rock (r) mode, calculated at 718 cm⁻¹ and observed at 736 cm⁻¹ (unless the very weak 725-cm⁻¹ band¹⁶ is assignable to this mode). Its observed shift to 781 cm⁻¹ in the T conformer,⁹ however, is reasonably well reproduced by the calculation. [We should note that the observed difference between $\nu_G(\text{SH})$ and $\nu_T(\text{SH})$ can be accounted for by the difference in force constants $f(\text{SH})$ that results from differences in electronic structure between the conformers (cf. Table V), since an unchanged $f(\text{SH})$ results in an unchanged $\nu(\text{SH})$. It is therefore unnecessary to invoke an interaction between SH s and CCSH torsion,¹⁰ (t) which in any case is very small (cf. Table V), in order to account for this difference. That $\Delta\nu(\text{calc}) = 3 \text{ cm}^{-1}$ compared to $\Delta\nu(\text{obs}) = 9 \text{ cm}^{-1}$ may be due to the limited basis set used in our ab initio calculation or to anharmonicity effects.] For the few bands assignable to the T conformer, the agreement is good and the assignments are reasonable. How-

ever, it has been proposed¹⁶ that some of these bands may be combinations or hot bands (e.g., 855 = CS s + 0 → 1 CS t, 232 = 1 → 2 CC t, 157 = 1 → 2 CS t). There was no such assignment for the 675-cm⁻¹ band,¹⁶ and it and the 781-cm⁻¹ band are convincingly assigned to the T conformer,⁹ so it may be necessary to reexamine such alternate assignments.¹⁶ After all, if $\Delta E(\text{G-T})$ is 0.336 kcal/mole,¹⁶ there should be ~ 36% of the T conformer in the sample.

The ab initio force field also accounts convincingly for the frequency shifts on deuteration (Table VII). The $\nu(\text{SH})$ mode of the G conformer shifts from 2594 (calculated at 2594) to 1881 (1866) cm⁻¹; this discrepancy is undoubtedly due to the different (and large) anharmonicities of the SH and SD modes (the ab initio calculation of course gives harmonic frequencies). The shift of the CH₃r/CSH bend (b) mode from 1098 (1100) to 1069 (1064) cm⁻¹ is well reproduced. The CSH b mode at 870 (870) cm⁻¹ moves to 625 (610) cm⁻¹, a reasonable prediction in view of the different anharmonicities. What is particularly gratifying is that shifts of modes not involving significant SH contributions are well predicted, viz., CS s for G [662 (662) to 676 (670) cm⁻¹] and T [675 (672) to 696 (704) cm⁻¹] conformers, and CS t for G [191 (191) to 150 (144)

Table II Internal Coordinates of ET^a

R1 = $\Delta r(\text{S1-H2})$
R2 = $\Delta r(\text{S1-C3})$
R3 = $\Delta r(\text{C3-C4})$
R4 = $\Delta r(\text{C3-H5})$
R5 = $\Delta r(\text{C3-H6})$
R6 = $\Delta r(\text{C4-H7})$
R7 = $\Delta r(\text{C4-H8})$
R8 = $\Delta r(\text{C4-H9})$
R9 = $\Delta\theta(\text{C3-S1-H2})$
R10 = $\Delta\theta(\text{C4-C3-S1})$
R11 = $\Delta\theta(\text{S1-C3-H5})$
R12 = $\Delta\theta(\text{S1-C3-H6})$
R13 = $\Delta\theta(\text{C4-C3-H5})$
R14 = $\Delta\theta(\text{H5-C3-H6})$
R15 = $\Delta\theta(\text{C4-C3-H6})$
R16 = $\Delta\theta(\text{C3-C4-H9})$
R17 = $\Delta\theta(\text{C3-C4-H7})$
R18 = $\Delta\theta(\text{C3-C4-H8})$
R19 = $\Delta\theta(\text{H7-C4-H9})$
R20 = $\Delta\theta(\text{H7-C4-H8})$
R21 = $\Delta\theta(\text{H8-C4-H9})$
R22 = $\Delta\tau(\text{C3S1})$
R23 = $\Delta\tau(\text{C3C4})$

^a For atom numbering, see Figure 1.

Table III Local Symmetry Coordinates of ET

Symmetry Coordinates	Description ^a
S1 = R1	SH s
S2 = R2	CS s
S3 = R3	CC s
S4 = R4 + R5	CH ₂ ss
S5 = R4 - R5	CH ₂ as
S6 = R6 + R7 + R8	CH ₃ ss
S7 = R6 - R7	CH ₃ as1
S8 = 2R8 - R6 - R7	CH ₃ as2
S9 = R9	CSH b
S10 = 5R10 - R11 - R12 - R13 - R14 - R15	CCS d
S11 = 4R14 - R11 - R12 - R13 - R15	CH ₂ b
S12 = R13 + R15 - R11 - R12	CH ₂ w
S13 = R13 + R11 - R12 - R15	CH ₂ tw
S14 = R11 + R15 - R12 - R13	CH ₂ r
S15 = R19 + R20 + R21 - R16 - R17 - R18	CH ₃ sb
S16 = 2R20 - R19 - R21	CH ₃ ab1
S17 = R19 - R21	CH ₃ ab2
S18 = 2R16 - R17 - R18	CH ₃ r1
S19 = R17 - R18	CH ₃ r2
S20 = R22	CS t
S21 = R23	CC t
S22 = R10 + R11 + R12 + R13 + R14 + R15	Red 1
S23 = R16 + R17 + R18 + R19 + R20 + R21	Red 2

^a s: stretch; ss: symmetric stretch; as: antisymmetric stretch; b: bend, sb: symmetric bend; ab: antisymmetric bend; d: deformation; w: wag; r: rock; tw: twist; t: torsion; Red: redundancy.

cm⁻¹] and T [157 (166) to 136 (125) cm⁻¹] conformers. Thus, this force field should be transferable to other molecules with SH groups.

FORCE FIELD OF CYSTEINE RESIDUE

In order to properly predict the modes of the cysteine residue in proteins, we need to be sure that the force field can satisfactorily reproduce the dependence of $\nu(\text{SH})$ and $\nu(\text{CS})$ on χ^1 , and, as found in the disulfide case,²⁶ the differences introduced by adjacent peptide groups. We have examined the first point by transferring our ab initio force field to calculations of the modes of 1PT, the simplest relevant molecule that can exhibit a χ^1 dependence, and 3TNMP, a molecule containing an adjacent peptide group and for which the effects of hydrogen bonding can also be examined.

1-Propanethiol

The force field used in the calculation of the normal modes of 1PT consisted of two parts: the ab initio

force field for the -C ^{β} H₂SH portion of the molecule, taken from that of ET, and an empirical force field for the CH₃C ^{α} H₂- hydrocarbon part of the molecule.³⁶ This combined force field gave reasonably good frequency agreement with the frequencies of the expected predominant $\chi^1\chi^2 = \text{TG}$ conformer,¹³ particularly in the low-temperature liquid¹² (we agree with this TG assignment¹³ but not with the proposed TT or GT assignments¹²). We then did a least-squares refinement of the empirical force constants (which is not unreasonable in merging two such force fields) to optimize the agreement. The force constant changes were very small, and we then applied this force field to the calculation of the normal modes of the five possible conformers. The results are given in Table VIII.

In making the assignments proposed in the table, we were guided by the consideration that, while the energy of the TG conformer is the lowest, those of GG and TT may well be about equal to each other because of the roughly comparable $\Delta E(\text{T-G})$ for χ^1 and χ^2 . We find that some spectral results, particularly in comparing band intensities between gas (g) room-temperature liquid (l_R), and low-temper-

Table IV Scale Factors for ET and 3TNMP Force Constants

Internal Coordinate ^a	ET	3TNMP	Ratio ^b
SH s	0.8084	0.7826	0.9681
CS s	0.9045	0.9061	1.0018
CC s	0.9771	0.9281	0.9499
C ^β H s	0.8255		
C ^α H s	0.8189		
CSH b	0.7841	0.7626	0.9726
CCS b	0.8451	0.9470	1.1206
SCH b	0.7602	0.8431	1.1090
C ^α C ^β H b	0.7904	0.6718	0.8499
C ^β C ^α H b	0.7685	0.8679	1.1294
HC ^β H b	0.7558	0.7410	0.9804
HC ^α H b	0.7528	0.7405	0.9837
CC t	0.8590	0.8590	1.0000
CS t	0.6246	0.6246	1.0000
CN s		0.8318	0.9442
CO s		0.6962	1.0423
CC ^α and NC(H ₃) s		0.8284	0.9827
CNC, NCC ^α , NCO, and C ^α CO d		0.7881	0.8651
CNH b		0.8118	1.0667
NCH b		0.7474	0.9558
HC (H ₃)H b		0.7686	0.9879
NH ob		0.8849	1.0056
CO ob		1.0179	1.1567
CN t		0.8752	0.9946
CC and NC t ^c		1.0000	1.0000

^a s: stretch; b: bend, ob: out-of-plane bend; d: deformation; t: torsion.

^b Ratio of 3TNMP to ET or N-methylacetamide³³ scale factors.

^c From empirical force field.³⁴

ature liquid (*l_L*), indicate the presence of TT as well as GG conformers (contrary to previous assumptions¹³), while there is no evidence for the necessary presence of G'G or GT.

The predominance of the TG conformer in the *l_L* phase is attested to by the excellent agreement between the observed band frequencies and intensities and the frequencies calculated for the various conformers. This is particularly true in view of the presence of unique TG bands at 961, 898, and (probably) 781 cm⁻¹, and the weakened band at ~ 790 cm⁻¹ and absent band at 655 cm⁻¹. The presence of the GG conformer in the *g* and *l_R* phases is supported by bands at 1210, ~ 920, 880, 793, and 655 cm⁻¹ (although some of these coincide with bands of the G'G conformer). The intensity change at ~ 1085 cm⁻¹ indicates the likely presence of the TT conformer in these phases. The strong bands at 706 and 736 cm⁻¹ undoubtedly result from a Fermi resonance between CS s and the overtone of CCC deformation (d) (as previously suggested¹⁷), and

several other combination bands are readily indicated. Incidentally, the slightly lower observed $\nu(\text{SH})$ (compared to ET), with a necessarily smaller $f(\text{SH})$ to give closer agreement, probably reflects the presence of hydrogen bonding, at least in the liquid.

Thus, reasonable assignments can be made for 1PT, the results of which show that the ab initio force field provides a reliable description of the conformation dependence of frequencies associated with the -CH₂SH group.

3-Thiol-N-Methylpropionamide

No crystal structure has been reported for 3TNMP, and we have therefore used aqueous solution data³² in refining the scale factors for this molecule (these frequencies are in fact close to those of the solid³²). In order to provide a relevant model for the normal mode calculations, we have placed H and O atoms in hydrogen-bonding positions appropriate to the peptide group, taken from NMA,³³ and to the SH

Table V Scaled Ab Initio Force Field of ET for Different Conformers (Force Constants ≥ 0.03)

Force Constant ^a	Value ^b				Force Constant ^a	Value ^b			
	C	G	S	T		C	G	S	T
SH	3.903	3.875	3.914	3.882	CH7/CCH9	-0.067	-0.067	-0.068	-0.067
SH/CCS	-0.043	-0.011	0.010	0.026	CH7/CCH7	0.052	0.054	0.055	0.054
SC	3.012	3.064	3.016	3.044	CH7/CCH8	-0.067	-0.068	-0.068	-0.067
SC/CC	0.145	0.181	0.194	0.181	CH7/HCHa	0.079	0.080	0.077	0.079
SC/CH5	0.076	0.079	0.074	0.070	CH7/HCHb	0.079	0.077	0.079	0.079
SC/CH6	0.076	0.068	0.030	0.072	CH7/HCHc	-0.072	-0.072	-0.072	-0.072
SC/CSH	0.226	0.171	0.209	0.146	CH8	4.773	4.811	4.770	4.786
SC/CCS	0.208	0.221	0.209	0.158	CH8/CH9	0.048	0.048	0.048	0.048
SC/SCH5	0.177	0.133	0.178	0.185	CH8/CCS	-0.021	-0.033	-0.015	-0.016
SC/SCH6	0.178	0.182	0.167	0.187	CH8/CCH6	0.034	0.036	0.026	0.032
SC/CCH5	-0.193	-0.187	-0.198	-0.186	CH8/CCH9	-0.066	-0.065	-0.067	-0.065
SC/HCH6	-0.180	-0.180	-0.176	-0.154	CH8/CCH7	-0.067	-0.065	-0.067	-0.068
SC/CCH6	-0.192	-0.173	-0.181	-0.186	CH8/CCH8	0.050	0.049	0.052	0.054
SC/CCH9	-0.027	-0.031	-0.028	-0.018	CH8/HCHa	-0.070	-0.068	-0.072	-0.070
SC/CCH7	0.058	0.065	0.063	0.045	CH8/HCHb	0.076	0.072	0.077	0.072
CC	4.379	4.383	4.380	4.369	CH8/HCHc	0.078	0.079	0.080	0.078
CC/CH5	0.070	0.075	0.080	0.071	CH9	4.771	4.751	4.808	4.782
CC/CH6	0.070	0.077	0.074	0.071	CH9/CCS	-0.021	-0.017	-0.033	-0.015
CC/CH7	0.075	0.077	0.076	0.073	CH9/CCH5	0.034	0.031	0.036	0.031
CC/CH8	0.074	0.065	0.074	0.076	CH9/CCH9	0.050	0.049	0.047	0.055
CC/CH9	0.074	0.073	0.065	0.077	CH9/CCH7	-0.067	-0.067	-0.064	-0.068
CC/CSH	-0.049	-0.026	0.021	0.053	CH9/CCH8	-0.066	-0.068	-0.065	-0.066
CC/CCS	0.149	0.139	0.144	0.192	CH9/HCHa	0.077	0.080	0.073	0.073
CC/SCH5	-0.153	-0.150	-0.145	-0.157	CH9/HCHb	-0.070	-0.072	-0.068	-0.071
CC/SCH6	-0.152	-0.143	-0.148	-0.157	CH9/HCHc	0.078	0.081	0.080	0.079
CC/CCH5	0.137	0.141	0.132	0.137	CSH	0.850	0.815	0.824	0.794
CC/HCH6	-0.143	-0.142	-0.147	-0.156	CSH/CCS	-0.051	-0.061	0.062	0.176
CC/CCH6	0.137	0.126	0.143	0.136	CSH/SCH5	0.058	0.154	0.061	-0.060
CC/CCH9	0.161	0.159	0.170	0.156	CSH/SCH6	0.056	-0.058	-0.067	-0.058
CC/CCH7	0.113	0.100	0.099	0.124	CSH/CCH7	-0.013	-0.006	0.019	0.038
CC/CCH8	0.161	0.169	0.164	0.156	CCS	0.768	0.737	0.740	0.745
CC/HCHa	-0.154	-0.154	-0.147	-0.153	CCS/SCH5	-0.150	-0.145	-0.156	-0.142
CC/HCHb	-0.154	-0.146	-0.154	-0.153	CCS/SCH6	-0.149	-0.141	-0.133	-0.143
CC/HCHc	-0.143	-0.144	-0.146	-0.145	CCS/CCH5	-0.177	-0.175	-0.179	-0.155
CH5	4.855	4.829	4.861	4.848	CCS/HCH6	-0.148	-0.145	-0.149	-0.151
CH5/CH6	0.058	0.051	0.051	0.056	CCS/CCH6	-0.176	-0.171	-0.153	-0.157
CH5/CCS	-0.052	-0.057	-0.046	-0.051	CCS/CCH9	-0.034	-0.036	-0.041	-0.036
CH5/SCH5	0.008	0.052	0.005	0.003	CCS/CCH7	0.085	0.091	0.093	0.091
CH5/SCH6	-0.044	-0.059	-0.052	-0.045	CCS/CCH8	-0.035	-0.041	-0.040	-0.036
CH5/CCH5	0.056	0.055	0.064	0.057	CCS/CS t	-0.000	0.000	0.045	-0.002
CH5/HCH6	0.098	0.090	0.091	0.094	SCH5	0.544	0.549	0.540	0.523
CH5/CCH6	-0.059	-0.066	-0.058	-0.058	SCH5/SCH6	-0.115	-0.110	-0.105	-0.099
CH6	4.856	4.841	4.868	4.847	SCH5/CCH5	-0.066	-0.063	-0.069	-0.083
CH6/CSH	0.014	-0.010	-0.040	-0.002	SCH5/HCH6	-0.080	-0.071	-0.074	-0.076
CH6/CCS	-0.052	-0.047	-0.049	-0.051	SCH5/CCH6	-0.121	-0.126	-0.122	-0.118
CH6/SCH5	-0.044	-0.052	-0.050	-0.045	SCH5/CS t	0.045	-0.011	-0.036	0.004
CH6/CCH5	-0.059	-0.057	-0.057	-0.058	SCH6	0.543	0.526	0.526	0.524
CH6/HCH6	0.098	0.092	0.094	0.094	SCH6/CCH5	-0.121	-0.120	-0.120	-0.118
CH6/CCH6	0.056	0.063	0.058	0.057	SCH6/HCH6	-0.080	-0.089	-0.080	-0.076
CH7	4.747	4.719	4.724	4.762	SCH6/CCH6	-0.066	-0.064	-0.076	-0.083
CH7/CH8	0.049	0.046	0.050	0.048	SCH6/CS t	-0.045	0.011	-0.003	-0.003
CH7/CH9	0.049	0.051	0.046	0.048	CCH5	0.543	0.555	0.556	0.547
CH7/CCS	0.033	0.043	0.044	0.031	CCH5/HCH6	-0.074	-0.061	-0.074	-0.081

Table V (Continued)

Force Constant ^a	Value ^b				Force Constant ^a	Value ^b			
	C	G	S	T		C	G	S	T
CCH5/CCH6	-0.116	-0.126	-0.120	-0.113	CCH7/CCH8	-0.126	-0.126	-0.125	-0.126
CCH5/CCH9	0.087	0.087	0.089	0.089	CCH7/HCHa	-0.084	-0.087	-0.086	-0.085
CCH5/CCH7	-0.037	-0.039	-0.041	-0.038	CCH7/HCHb	-0.084	-0.085	-0.087	-0.085
CCH5/CCH8	-0.040	-0.038	-0.038	-0.040	CCH7/HCHc	-0.099	-0.096	-0.095	-0.097
HCH6	0.490	0.482	0.475	0.465	CCH8	0.524	0.524	0.528	0.527
HCH6/CCH6	-0.075	-0.078	-0.073	-0.081	CCH8/HCHa	-0.099	-0.100	-0.098	-0.099
CCH6	0.543	0.530	0.534	0.548	CCH8/HCHb	-0.094	-0.090	-0.094	-0.094
CCH6/CCH9	-0.040	-0.038	-0.039	-0.040	CCH8/HCHc	-0.089	-0.092	-0.092	-0.090
CCH6/CCH7	-0.037	-0.038	-0.037	-0.038	HCHa	0.454	0.453	0.447	0.452
CCH6/CCH8	0.087	0.087	0.088	0.089	HCHa/HCHb	-0.079	-0.081	-0.080	-0.078
CCH9	0.524	0.525	0.525	0.527	HCHa/HCHc	-0.085	-0.082	-0.082	-0.084
CCH9/CCH7	-0.126	-0.125	-0.127	-0.125	HCHb	0.454	0.448	0.452	0.452
CCH9/CCH8	-0.128	-0.128	-0.128	-0.129	HCHb/HCHc	-0.085	-0.083	-0.081	-0.084
CCH9/HCHa	-0.094	-0.093	-0.090	-0.095	HCHc	0.451	0.451	0.450	0.450
CCH9/HCHb	-0.099	-0.098	-0.099	-0.099	CS t	-0.029	0.035	-0.024	0.028
CCH9/HCHc	-0.089	-0.090	-0.093	-0.090	CC t	0.116	0.098	0.098	0.092
CCH7	0.515	0.510	0.511	0.513					

^a XY: X—Y stretch; XYZ: X—Y—Z bend; t: torsion. HCHa: H7C4H9; HCHb: H7C4H8; HCHc: H8C4H9.

^b Values included only if force constant for one conformer ≥ 0.03 .

group, taken from the crystal structure of glutathione.³⁷

Since it is not possible to know a priori which of the many conformers of 3TNMP may be most prevalent in aqueous solution, we have attempted to answer this question spectroscopically. The conformation of 3TNMP depends on three dihedral angles, $\text{NCC}^\alpha\text{C}^\beta$ (ψ), $\text{CC}^\alpha\text{C}^\beta\text{S}$ (χ^1), and $\text{C}^\alpha\text{C}^\beta\text{SH}$ (χ^2) (see Figure 2). We have, by analogy with the observations for ET and IPT, set $\chi^2 = \text{G}$. We then examined the normal modes of the 15 conformers obtained by setting $\chi^1 = \text{T}$, G , and G' and $\psi = \text{S}'$, G' , G , S , and T (the $\psi = \text{C}$ structure is obviously of high energy). Using an initial force field (see below), we did a least squares refinement of scale factors for each conformer to the observed non-CH s NH, SH and (with half the weight) ND, SD frequencies, and examined the goodness of fit using several criteria. We assume that, in general, a relatively complete starting force field will refine most satisfactorily for the conformer(s) that actually gives rise to the observed frequencies. Such selectivity seems to be indicated here.

The initial force field was obtained by combining the ab initio force field of hydrogen-bonded NMA³³ with our ab initio force field for ET (except that empirical values³⁴ were used for CC t and NC t). Since both of these molecules have a $\text{C}^\alpha\text{H}_3$ rather than a $\text{C}^\alpha\text{H}_2$ group, it was necessary to modify the

force constants at the C^α atom, where the two molecules "merge." This was done as follows. First, diagonal force constants relating to the merged region were retained from each force field, for example, $f(\text{C}^\beta\text{C}^\alpha\text{H})$ from the ET force field and $f(\text{CC}^\alpha\text{H})$ from the NMA force field. For $f(\text{C}^\alpha\text{H})$ and $f(\text{HC}^\alpha\text{H})$ the ET values were initially used, and an appropriate starting value was taken for $f(\text{CC}^\alpha\text{C}^\beta)$.³⁶ Of course, common force constants in each force field were required to have the same values; thus, what was the corresponding $f(\text{C}^\alpha\text{H})$ in the NMA force field must take the value of $f(\text{C}^\alpha\text{C}^\beta)$ of the ET force field and what was the $f(\text{C}^\alpha\text{H})$ of ET must take the value of $f(\text{CC}^\alpha)$ of NMA. Second, although all off-diagonal terms in each force field were retained, it was assumed that there were no interaction terms between the two parts; thus, for example, $f(\text{NCC}^\alpha, \text{CC}^\alpha\text{C}^\beta)$ would be included (with an initial value from the NMA force field) but not $f(\text{NCC}^\alpha, \text{C}^\alpha\text{C}^\beta\text{S})$. Relevant off-diagonal constants not in either force field were approximated by using the standard scaling procedure³⁵ on the corresponding adjusted diagonal scale factors and applying the resulting scale factor to the appropriate off-diagonal constant in the ET or NMA force fields. Third, hydrogen-bond force constants were chosen from NMA³³ to give approximate $\text{H} \cdots \text{O}$ s frequencies (below $\sim 150 \text{ cm}^{-1}$) in order to incorporate this effect, without expecting any accuracy in these modes since no experimental

Table VI Observed and Calculated Frequencies (in cm^{-1}) of ET

ν (obs) ^a		ν (calc)		$I(\text{ir})^b$	$I(\text{R})^b$	PED ^c
Gas	Cond	$\nu(\text{G})$	$\nu(\text{T})$			
	2980	2988		24.9	11.0	CH ₂ as(80) CH ₃ as2(11)
			2993	25.4	19.7	CH ₂ as(88)
2971vw	2966	2961		6.7	98.7	CH ₃ as1(67) CH ₂ as(17) CH ₃ as2(12)
			2961	28.8	72.4	CH ₃ as1(75) CH ₂ ss(17)
2946vs		2950		31.6	78.5	CH ₃ as2(54) CH ₂ ss(38)
			2958	4.8	106.0	CH ₃ as2(78) CH ₂ as(12) CH ₃ as1(10)
2940vs	2931	2937		7.3	109.5	CH ₂ ss(60) CH ₃ as2(24) CH ₃ as1(15)
			2942	14.9	105.2	CH ₂ ss(83) CH ₃ as1(13)
2886m	2876	2895		23.6	100.6	CH ₃ ss(97)
			2902	19.3	97.2	CH ₃ ss(100)
2594s	2561	2594		13.6	117.0	SH s(100)
2603w			2597	17.0	148.7	SH s(100)
		1450		4.1	16.6	CH ₃ ab2(88)
1450w	1451	1453		4.3	16.6	CH ₃ ab2(72) CH ₃ ab1(21)
		1447		10.7	22.7	CH ₃ ab1(87)
			1449	9.9	25.0	CH ₃ ab1(71) CH ₃ ab2(21)
	1434	1434		6.1	21.6	CH ₂ b(99)
			1443	6.7	18.7	CH ₂ b(100)
	1377	1375		6.0	3.2	CH ₃ sb(105)
			1375	6.8	2.9	CH ₃ sb(106)
1275w	1272	1279		12.9	2.0	CH ₂ w(92)
			1278	32.4	1.9	CH ₂ w(94)
	1250	1255		4.3	12.7	CH ₂ tw(63) CH ₃ r1(10)
			1246	0.6	10.8	CH ₂ tw(57) CH ₃ r1(15)
1098w	1095	1100		13.1	11.9	CH ₃ r1(30) CSH b(23) CH ₂ r(15)
			1094	3.9	13.9	CC s(28) CH ₃ r2(26) CSH b(19)
1045w	1052	1048		3.4	5.0	CC s(39) CH ₃ r2(38)
			1021	0.0	6.5	CH ₂ tw(43) CH ₃ r1(27) CH ₂ r(21) CH ₃ r2(10)
985w	971	985		4.8	8.2	CC s(49) CH ₃ r2(28)
			990	6.4	6.8	CC s(56) CH ₃ r2(19)
870w	868	870		12.0	7.4	CSH b(57) CH ₃ r1(34) CH ₂ tw(22)
855w			859	1.7	12.7	CSH b(77) CC s(13) CH ₃ r2(11)
736w	738	718		2.5	3.0	CH ₂ r(72) CSH b(17) CH ₃ r1(10)
	781		770	5.7	0.4	CH ₂ r(73) CH ₃ r1(27)
662m	660	662		5.1	25.3	CS s(89)
675w			672	1.9	30.0	CS s(92)
330w	333	330		1.2	2.0	CCS d(88)
			308	2.8	0.7	CCS d(95)
246vww		246		6.1	2.2	CC t(90)
232vww			240	2.8	2.1	CC t(86)
191w ^d		191		18.8	7.9	CS t(95)
157vww ^e			166	22.1	9.6	CS t(98) CC t(14)

^a Gas-phase Raman frequencies: Ref. 16, except for 2594 and 2603, from Ref. 10. Condensed-phase frequencies (average of liquid and glass phases): Ref. 9.

^b $I(\text{ir})$ = calculated ir intensity, $I(\text{R})$ = calculated Raman intensity.

^c Contributions to potential energy distribution ≥ 10 . s: Stretch; ss: symmetric stretch; as: antisymmetric stretch; b: bend; sb: symmetric bend; ab: antisymmetric bend; d: deformation; w: wag; tw: twist; r: rock; t: torsion.

^d Given as 190 cm^{-1} in Ref. 14 and 193 cm^{-1} in Ref. 15.

^e Given as 158 cm^{-1} in Refs. 14 and 15.

data are available (in fact, we do not report such modes).

The calculations showed that a number of modes were particularly sensitive to conformation, with

frequencies varying by up to $\sim 70 \text{ cm}^{-1}$ between conformers. In Table IX we present various measures of the goodness of fit between the observed frequencies of 3TNMP and the calculated frequen-

Table VII Observed and Calculated Frequencies (in cm^{-1}) of ET-SD

$\nu(\text{obs})^a$	ν (calc)		PED ^b
	ν (G)	ν (T)	
	2988		CH ₂ as(80) CH ₃ as2(11)
		2993	CH ₂ as(88)
2971vw	2961		CH ₃ as1(87) CH ₂ as(18) CH ₃ as2(12)
		2961	CH ₃ as1(75) CH ₂ ss(16)
2945vs	2950		CH ₃ as2(54) CH ₂ ss(38)
2951vw		2958	CH ₃ as2(78) CH ₂ as(12) CH ₃ as1(10)
2940vs	2937		CH ₂ ss(60) CH ₃ as2(24) CH ₃ as1(15)
		2942	CH ₂ ss(83) CH ₃ as1(13)
2882m	2895		CH ₃ ss(97)
		2902	CH ₃ ss(100)
1881m	1866		SD s(100)
1886sh		1868	SD s(100)
	1451		CH ₃ ab2(87)
1450w		1453	CH ₃ ab2(72) CH ₃ ab1(21)
	1447		CH ₃ ab1(87)
		1449	CH ₃ ab1(71) CH ₃ ab2(21)
	1434		CH ₂ b(99)
		1443	CH ₂ b(100)
	1375		CH ₃ sb(105)
		1375	CH ₃ sb(106)
1271w	1277		CH ₂ w(94)
		1275	CH ₂ w(95)
	1253		CH ₂ tw(63) CH ₃ r1(12)
		1246	CH ₂ tw(57) CH ₃ r1(15)
1069w	1064		CC s(31) CH ₃ r1(29)
		1067	CC s(36) CH ₃ r2(32) CH ₃ r1(11)
1046w	1042		CH ₃ r2(30) CH ₂ tw(23) CC s(16) CH ₂ r(15)
		1021	CH ₂ tw(43) CH ₃ r1(28) CH ₂ r(21) CH ₃ r2(10)
978w	984		CC s(48) CH ₃ r2(26)
		990	CC s(58) CH ₃ r2(18)
	779		CH ₂ r(48) CH ₃ r1(33) CSD b(15)
		769	CH ₂ r(73) CH ₃ r1(28)
676m	670		CS s(89) CCS d(10)
696w		704	CS s(63) CSD b(33) CH ₃ r2(11)
625w	610		CSD b(80) CH ₂ r(31)
		620	CSD b(56) CS s(32)
325w	326		CCS d(87)
301vvw		302	CCS d(95) CSD b(13)
250vw	244		CC t(93)
232vw		238	CC t(92)
150w ^c	144		CS t(99)
136w ^d		125	CS t(101)

^a Gas-phase Raman frequencies from Ref. 16.

^b Contributions to potential energy distribution ≥ 10 . s: Stretch; ss: symmetric stretch; as: antisymmetric stretch; b: bend; sb: symmetric bend; ab: antisymmetric bend; d: deformation; w: wag; tw: twist, r: rock; t: torsion.

^c Given as 149 cm^{-1} in Ref. 15.

^d Given as 125 cm^{-1} in Ref. 15.

cies obtained from a least-squares refinement of scale factors for each conformer. Since the CS s frequency was well reproduced for ET and 1PT, we expect that it should be comparably well predicted in 3TNMP. Thus, the TGG and TG'G conformers,

with discrepancies of 84 and 91 cm^{-1} , respectively, are unacceptable possibilities for the actually existing structure(s). A similar potential conclusion about S'TG, G'TG, GTG, STG, and TTG is strengthened when we examine the average devia-

Table VIII Observed and Calculated Frequencies (in cm^{-1}) of IPT

ν (obs) ^a			ν (calc)					Mode ^b	
ν (g)	ν (l _R)	ν (l _L)	TG	GG	TT	G'G	GT		
			2985	2984	2991	2985	2992	C ^β H ₂ as	
2950bd	2950bd	2950bd	{	2955	2955	2955	2955	2955	CH ₃ as2
				2955	2955	2955	2955	2955	CH ₃ as1
				2946	2946	2949	2946	2948	C ^β H ₂ ss
				2934	2931	2936	2935	2936	C ^α H ₂ as
				2881	2880	2880	2880	2881	C ^α H ₂ ss
				2862	2862	2862	2862	2862	CH ₃ sb
2565s	2564s	2560s	2570	2570	2573	2570	2573	SH s	
	1458s	1458s	{	1468	1468	1468	1468	1468	CH ₃ ab2
1450s	1440s	1440s	{	1467	1467	1466	1467	1467	CH ₃ ab1
				1442	1435	1449	1440	1449	C ^β H ₂ b
1380s	1380s	1380m		1424	1424	1425	1426	1428	C ^α H ₂ b
1380s	1380s	1380m		1376	1377	1376	1376	1376	CH ₃ sb
1340w	1339m	1340m							{ 290 + 1030 = 1320
									{ 2 × 655 = 1310
1298s	1298s	1297s			1309				C ^α H ₂ w
			1294			1285	1283	1287	C ^α H ₂ w
						1285			C ^β H ₂ w
1240s	1246s	1249s		1259		1256	1257	1252	C ^β H ₂ w
				1228			1218		C ^β H ₂ tw
1225sh	1225s	1221s	{						C ^α H ₂ tw
			{		1209			1197	C ^β H ₂ tw
1210w	1208sh		{		1204				C ^α H ₂ tw
			{	1162		1141			C ^β H ₂ tw
					1126				CSH b C ^α H ₂ tw
							1100		C ^β H ₂ r
1105m	1105s	1108s	{						290 + 810 = 1100
			{	1085					CH ₃ r2
1085s	1088s	1082m	{			1091			CC ^α s
			{	1062		1054		1044	C ^α H ₂ tw
1060w	1060sh	1062m	{				1059	1069	CH ₃ r2
			{		1057				CC ^α s
			{	1029	1028	1026	1022	1029	CC ^α s
961vw	961vw	960w		961					CSH b
						937	945	940	CSH b
920w	925w	918w			928				CSH b
890s	892s	891vs		898					C ^α C ^β s
					883		884		C ^α C ^β s
880sh	880m	880sh							
						871		875	C ^β H ₂ r
						847		851	C ^α C ^β s
							837		CH ₃ r1
815w	815w	810s		806	815	802			C ^α H ₂ r
793s	793s	790sh			791				C ^β H ₂ r
780w	779sh	780w		781					C ^β H ₂ r
								781	C ^α H ₂ r
							744		C ^β H ₂ r

Table VIII (Continued from the previous page)

ν (obs) ^a			ν (calc)					Mode ^b	
ν (g)	ν (l _R)	ν (l _L)	TG	GG	TT	G'G	GT		
735m	736m	736vs	736		745			CS s	
710s	706m	706vs						2 × 363 = 726	
655s	655s				659		666	672	CS s
			417 (0)		472		478	467	CCC d
									290 + 118 = 408
	363 (5)			365		364			290 + 135 = 425
	290 (0.5)		274	274	252	297	278	CCS d	
	230 (0)		236	230	218	221	205	CS t	
			141	135	142	137	135	CC ^α t	
			118	110	114	111	106	C ^α C ^β t	

^a From Ref. 12, for gas, room-temperature liquid, and low-temperature (-120°C) liquid. Raman intensities of 417–230- cm^{-1} bands from Ref. 17.

^b Only major contributor is given, except for 1126 (GG), where both components are essentially equal. s: Stretch; ss: symmetric stretch; as: antisymmetric stretch; b: bend; sb: symmetric bend; ab: antisymmetric bend; d: deformation; w: wag; tw: twist, r: rock; t: torsion.

tion for frequencies $<1700\text{ cm}^{-1}$, which is highest for these conformers. Another criterion might be the number of modes that show very large deviation, say $\geq 20\text{ cm}^{-1}$. Such a "sieve" leaves three favorable cases: SGG, G'G'G, and SG'G. The last is disfavored by a large discrepancy for the amide V mode and the first by a large discrepancy for the CSH b mode. Thus, the G'G'G conformer, which is overall the most favored on these (as well as other) criteria, seems to be the most likely structure in solution. Its structure is shown in Figure 2, and we present a complete description of its normal modes in Table X, with scale factors being given in Table IV. While this procedure may not guarantee a unique structural solution, it should provide a test of whether the ET force field (or a very slight modification of it) can be transferred to a system in which there are peptide groups adjacent to the $-\text{CH}_2\text{SH}$ group, as in proteins.

As can be seen from Table X, the agreement between observed and calculated frequencies is quite good: the average discrepancy for bands below 1700 cm^{-1} is 4.3 cm^{-1} , compared to 5.7 for 1PT and 3.0 for ET. The assignments are reasonable and (not surprisingly) differ in some cases from those previously proposed on empirical grounds.³² We have also calculated the normal modes of the ND,SD molecule, and these compare well with the experimental data.³² The $\nu(\text{SH})$ frequency does not depend on χ^1 and has a small sensitivity to χ^2 . Since $\partial\nu(\text{SH})/\partial f(\text{H}\cdots\text{O})$ is positive, the lower $\nu(\text{SH})$ (compared to ET), with the required smaller $f(\text{SH})$,

reflects the hydrogen bonding of this group. The shift on deuteration to 1875 cm^{-1} (calculated at 1865 cm^{-1}) is reasonably well predicted considering the anharmonicity difference. Amide I at 1634 shifts to 1631 (1621) cm^{-1} and amide II at 1583 shifts to 1496 (1492) cm^{-1} on deuteration. The 1373-cm^{-1} band shifts to 1335 (1321 , C^αH₂ wag (w)) cm^{-1} on deuteration, a shift that is reasonably well predicted for this partially amide III mode. (It may be possible that a Fermi resonance with $2 \times 673 = 1346$ also contributes to the observed 1373- and 1316-cm^{-1} pair.) The three CH₂ w and CH₂ twist (tw) modes at 1292 , 1280 , and 1209 cm^{-1} are replaced by two

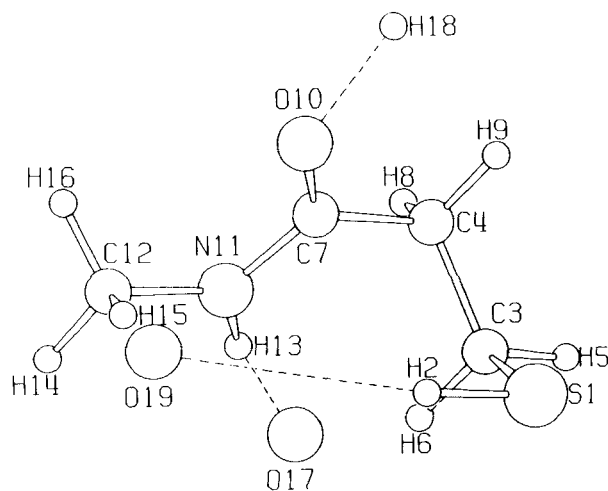
**Figure 2.** Structure of G'G'G conformer of 3TNMP.

Table IX Measures of Agreement Between Calculated^a and Observed^b Frequencies of Conformers of 3TNMP

Conformer ^c	$ \Delta\nu(\text{CS}) ^d$	$\langle \Delta\nu \rangle^e$	$N(\Delta\nu \geq 20)^f$	$ \Delta\nu(\text{V}) ^g$	$ \Delta\nu(\text{CSH}) ^h$
S'TG	14	14.4	3	6	20
G'TG	28	13.2	5	11	9
GTG	20	13.1	5	9	16
STG	19	11.6	3	9	8
TTG	14,126 ⁱ	10.6	3	13	7
S'GG	7	8.6	2	11	6
G'GG	3	10.1	4	9	2
GGG	8	9.7	3	5	6
SGG	1	6.0	0	8	14
TGG	84	7.9	2	9	3
S'G'G	1	8.4	3	6	1
G'G'G	2	4.3	0	7	0
GG'G	4	8.6	3	8	0
SG'G	6	8.3	0	47	10
TG'G	91	7.7	3	2	2

^a Based on least-squares refinement of force field for each conformer (see text).

^b From Ref. 32.

^c $\psi\chi^1\chi^2$ dihedral angles. Prime indicates negative angle.

^d Magnitude of difference between observed and calculated CS stretch mode, in cm^{-1} .

^e Average magnitude of difference between observed and calculated frequencies, in cm^{-1} .

^f Number of discrepancies whose magnitude is $\geq 20 \text{ cm}^{-1}$.

^g Magnitude of difference between observed and calculated amide V frequencies, in cm^{-1} .

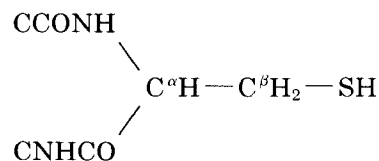
^h Magnitude of difference between observed and calculated CSH bend mode, in cm^{-1} .

ⁱ CS stretch divided between two modes.

bands on deuteration, 1287 (1275, $\text{C}^\alpha\text{H}_2$ tw, C^βH_2 w) and ~ 1232 (1219, C^βH_2 w) cm^{-1} , while bands at 1168, 1084, and 1025 cm^{-1} shift only slightly [although the 1091- (1107)- cm^{-1} band is now completely CH_3 r2]. The ND in-plane bend (ib) mode appears as a new medium band at 990 (981) cm^{-1} , while the 881-, 807-, and 770- cm^{-1} bands reorganize into modes at 895 (902, $\text{C}^\alpha\text{H}_2$ r), 848 (830, CH_3 r1, CN s, C^βH_2 r, CC^α s), and 767 (777, CO out-of-plane bend (ob)) cm^{-1} . The $\nu(\text{CS})$ mode shifts from 673 to 680 (679) cm^{-1} , and CCS d shifts from 369 to 355 (360) cm^{-1} . This general level of agreement for the deuterated molecule supports the validity of the force field and the assignments for 3TNMP, and provides the basis for a confident transfer of the ET force field to a model of the cysteine residue in proteins.

PROTEIN MODEL FOR CYSTEINE RESIDUE

The model system chosen to calculate the conformation dependence of $\nu(\text{SH})$ and $\nu(\text{CS})$ of the cysteine residue in proteins, following our experience with similar calculations on the disulfide bridge,²⁴ was



which is shown in Figure 3. The force field consisted of our ab initio force field for ET combined with our empirical force field for the peptide group,³⁴ a procedure that proved to be satisfactory in the analysis of the disulfide bridge.²⁴ The SH group was left non-hydrogen bonded, since this only affects $f(\text{SH})$ and we are interested here in the conformation dependence of $\nu(\text{SH})$. The structure is determined by χ^2 , χ^1 [taken as $\tau(\text{NC}^\alpha\text{C}^\beta\text{S})$], and (ϕ, ψ) . The calculations were done for $\chi^2 = 180^\circ$ (T), -150° (D'), -120° (S'), -90° (B'), -60° (G'), -30° (A'), 0° (C), 30° (A), 60° (G), 90° (B), 120° (S), and 150° (D); $\chi^1 = 180^\circ$ (N), -60° (C), and 60° (H); and $(\phi, \psi) = \alpha$ (-57.4° , -47.5°), γ (-86.9° , -3.0°), β (-138.4° , 135.7°), and ϵ (-80° , 142°) (the γ structure corresponds to that in glutathione,²⁷ and the ϵ structure is the average extended-helix conformation²⁶). The χ^2 dependence of the force

Table X Observed and Calculated Frequencies (in cm^{-1}) of G'G'G 3TNMP

ν (obs) ^a				
Solid				
IR	R	Solution R	ν (calc)	PED ^b
2572vs	2573vs	2582vs	2596	SH s(96)
1640s	1639s	1634m	1636	CO s(66) NH ib(21)
1560s	1559w	1583w	1586	NH ib(45) CN s(42) CO s(15)
	1469w	1461sh	1463	CH ₃ ab2(67) CH ₃ ab1(22)
			1451	C ^{α} H ₂ b(93)
	1432w	1432m	1433	CH ₃ ab1(73) CH ₃ ab2(24)
	1426s	1420m	1419	C ^{β} H ₂ b(85)
1407m			1407	CH ₃ sb(86)
1370w	1372w	1373m	1355	C ^{α} H ₂ w(33) CC ^{α} s(20) CN s(14)
				NH ib(13) CH ₃ sb(11)
		1316sh		598 + 720 = 1318
		1292m	1292	C ^{α} H ₂ w(53) CN s(11)
1275s	1272m	1280m	1277	C ^{α} H ₂ tw(27) C ^{β} H ₂ w(27) C ^{β} H ₂ tw(15)
1217s	1222w	1209w	1217	C ^{β} H ₂ w(74) C ^{α} H ₂ tw(19)
1162s	1162m	1168m	1167	C ^{β} H ₂ tw(54) C ^{α} H ₂ tw(24) CSH b(11)
				1149
			1108	CH ₃ r2(85)
1090sh	1090w	1084w	1084	NC s(50) CC ^{α} s(16)
1033sh	1032w	1025w	1025	C ^{α} C ^{β} s(71)
				981w
		945w		275 + 673 = 948
910s	911vs	912m	908	C ^{α} H ₂ r(63)
860m	860m	881m	874	CN s(17) CSH b(14) CH ₃ r1(13)
				804
	800sh	807w	804	CO ob(17) C ^{β} H ₂ r(17) CO ib(12)
765sh	767m	770w	766	C ^{β} H ₂ r(36) CO ob(35) CSH b(10)
720vs			713	CN t(37) NH ob(17) NH t(12)
673s	671m	673vs	675	CS s(71)
583s	584w	598m	592	CC ^{α} s(23) CO ib(22)
435s	437s	442m	455	C ^{α} CN d(32) CO ib(29) CCC d(10)
363w		369w	364	CCS d(58) CCC d(15)
278w	278vs	275m	288	CNC d(40) CS t(18) CCC d(12)
				251
			191	CS t(57) CNC d(20) C ^{α} CN d(13)
				CCC d(24) C ^{α} CN d(20) SH · · · O s(19)
			149	NC t(92)
			126	NH ob(61) CN t(45)
			63	C ^{α} C ^{β} t(33) NH · · · O b(17) NH t(15)
			42	NH t(62) CN t(13)
			24	CO t(94)
			8	SH t(79) CS t(11)

^a From Ref. 32.^b Contributions to potential energy distribution ≥ 10 . s: Stretch; ss: symmetric stretch; as: antisymmetric stretch; b: bend; sb: symmetric bend; ab: antisymmetric bend; ib: in-plane bend; ob: out-of-plane bend; d: deformation; w: wag; tw: twist; r: rock; t: torsion.

field was incorporated in the calculation, as in the case of the disulfide bridge.²⁴

The calculated $\nu(\text{SH})$ and $\nu(\text{CS})$ frequencies of

these 144 structures are given in Table XI. The calculated χ^2 dependence of $\nu(\text{SH})$ of ET is reproduced here, i.e., $\Delta\nu(\text{T-G}) = 3 \text{ cm}^{-1}$, but we see that the

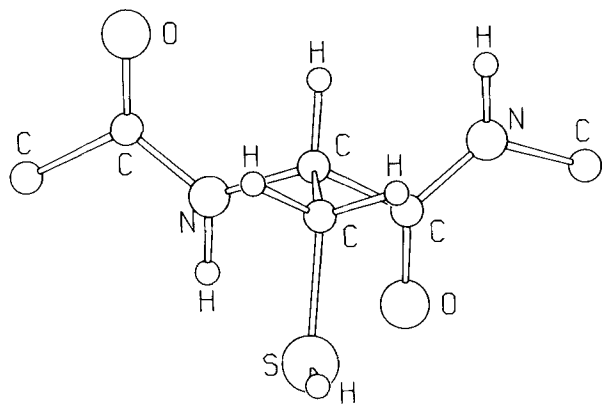


Figure 3. Structure of model for calculations of normal modes of cysteine residue in proteins.

maximum variation can be 14 cm^{-1} , viz., $\Delta\nu(\text{S-G})$. Since the observed ET difference is much larger, viz., 9 cm^{-1} , we expect that the observed conformational variation of $\nu(\text{SH})$ in proteins will be significantly larger than the predicted 14 cm^{-1} . Of course, this will be influenced by the effects of hydrogen bonding, so caution will be needed in inferring χ^2 from $\nu(\text{SH})$.

With respect to $\nu(\text{CS})$, we see that, as in the case of the disulfide bridge,²⁶ it depends strongly on χ^2 and χ^1 , and particularly on ϕ, ψ . The dependence is not exactly the same for the two structures, and the overlapping regions are different. In Table XII we present the χ^2 frequency ranges for the dominant $\nu(\text{CS})$ mode as a function of χ^1, ϕ, ψ , and it can be

Table XI Cysteine Residue SH and CS Stretch Frequencies (in cm^{-1})

χ^2 ^b	$\nu(\text{SH})$				$\nu(\text{CS})$ ^a								
	$N\alpha$	$N\gamma$	$N\beta$	$N\epsilon$	$C\alpha$	$C\gamma$	$C\beta$	$C\epsilon$	$H\alpha$	$H\gamma$	$H\beta$	$H\epsilon$	
T	2593	739 (66) ^c	739 (68)	758 (54)	737 (64)	680 (34)	805 (15)	784 (18)	725 (33)	683 (45)	711 (16)	673 (69)	668 (62)
D'	2598	738 (69)	739 (71)	759 (59)	739 (67)	830 (26)	826 (22)	812 (16)	814 (15)	681 (43)	711 (16)	669 (69)	666 (60)
						679 (31)	668 (50)	706 (37)	723 (34)	617 (16)	681 (52)		
S'	2604	734 (71)	736 (71)	756 (61)	737 (71)	650 (18)	818 (26)	800 (25)	805 (19)	680 (41)	710 (15)	665 (68)	664 (60)
						679 (27)	666 (43)	704 (37)	720 (33)	618 (19)	679 (53)		
B'	2597	739 (69)	746 (70)	765 (64)	737 (73)	651 (24)	825 (23)	817 (29)	801 (27)	798 (21)	684 (48)	711 (23)	675 (72)
						681 (30)	670 (51)	709 (40)	719 (33)	622 (16)	686 (51)		
G'	2590	736 (56)	750 (56)	763 (42)	733 (66)	655 (18)	826 (21)	815 (29)	805 (24)	793 (22)	686 (58)	709 (17)	683 (76)
						682 (36)	741 (16)	671 (52)	715 (40)	718 (35)	686 (53)	686 (24)	
A'	2595	736 (69)	740 (50)	758 (57)	731 (71)	666 (56)	823 (18)	803 (18)	789 (22)	788 (18)	681 (52)	711 (16)	672 (78)
						678 (35)	674 (15)	709 (39)	721 (38)	681 (58)			
C	2600	729 (68)	728 (71)	750 (55)	728 (69)	655 (15)	674 (31)	803 (18)	786 (18)	784 (15)	675 (45)	673 (58)	661 (74)
						652 (24)	660 (28)	704 (35)	720 (37)	617 (21)			
A	2595	739 (70)	740 (58)	753 (61)	732 (60)	657 (33)	836 (25)	832 (20)	821 (16)	722 (39)	681 (50)	711 (16)	670 (78)
						667 (55)	678 (37)	667 (55)	708 (40)	667 (23)	621 (18)	680 (57)	
G	2590	746 (62)	749 (64)	753 (56)	735 (54)	655 (16)	835 (30)	830 (29)	811 (31)	810 (19)	687 (51)	713 (20)	675 (72)
						682 (40)	678 (37)	673 (52)	711 (43)	723 (41)	684 (51)	675 (72)	664 (67)
B	2597	743 (72)	743 (74)	757 (62)	738 (66)	668 (21)	830 (25)	824 (26)	804 (27)	806 (18)	684 (45)	712 (18)	671 (73)
						667 (22)	682 (31)	672 (51)	708 (41)	723 (37)	621 (19)	683 (51)	
S	2604	734 (70)	736 (72)	756 (64)	737 (70)	655 (19)	825 (21)	819 (22)	798 (24)	802 (16)	680 (41)	679 (54)	665 (71)
						667 (22)	679 (27)	666 (45)	704 (38)	722 (34)	617 (18)	665 (71)	663 (62)
D	2598	738 (68)	741 (62)	761 (59)	739 (68)	651 (24)	827 (18)	808 (16)	795 (23)	791 (19)	681 (43)	711 (15)	670 (70)
						666 (22)	680 (30)	668 (50)	706 (37)	723 (33)	617 (16)	682 (54)	666 (62)

^a $\chi^1 = \tau(\text{NC}^\alpha\text{C}^\beta\text{S})$, N, C, H are atoms *trans* to S. $\alpha, \gamma, \beta, \epsilon = \phi, \psi$: $\alpha = -57.4^\circ, -47.5^\circ$; $\gamma = -86.9^\circ, -3.0^\circ$; $\beta = -138.4^\circ, 135.7^\circ$, $\epsilon = -80^\circ, 142^\circ$.

^b $\chi^2 = \tau(\text{C}^\alpha\text{C}^\beta\text{SH})$. T = 180° , D = 150° , S = 120° , B = 90° , G = 60° , A = 30° ; prime indicates negative angle.

^c Potential energy distribution in CS stretch ≥ 15 .

Table XII Frequency Ranges of CS Stretch Frequencies^a (in cm⁻¹) of Cysteine Residues in Proteins

$\chi^1\phi, \psi^b$	Range
N α	729–746
N γ	728–750
N β	750–765
N ϵ	728–739
C α	674–682
C γ	657–673
C β	704–715
C ϵ	718–725
H α	680–687
H γ	679–686
H β	661–683
H ϵ	663–670

^a For the dominant CS stretch mode from Table XI.

^b $\chi^1 = \tau(\text{NC}^\alpha\text{C}^\beta\text{S})$, N, C, H are atoms *trans* to S. $\alpha, \gamma, \beta, \epsilon = \phi$, ψ ; $\alpha = -57.4^\circ, -47.5^\circ$; $\gamma = -86.9^\circ, -3.0^\circ$; $\beta = -138.4^\circ, 135.7^\circ$, $\epsilon = -80^\circ, 142^\circ$.

seen that though there is overlap in other regions, these are unique for N β , C β , and C ϵ . Within each of these regions, χ^2 can be (though less uniquely) characterized from the dependence given in Table XI. Of course, as in the disulfide case,²⁶ not all $\chi^2\chi^1\phi, \psi$ may be found with equal probability in proteins. When sufficient structural data are available to compile such statistics, then, as with the disulfide bridge,²⁶ it will be possible to narrow the conformational probabilities on the basis of such information.

At present, detailed experimental data to compare to the predictions in Table XI are highly limited. The structure of glutathione is known,³⁷ and the cysteine residue in it has a conformation close to BH γ . Our predicted $\nu(\text{CS})$ is 682 cm⁻¹, and the spectra show a strong Raman band at 679 cm⁻¹ and a shoulder in the ir at 682 cm⁻¹,²⁷ in excellent agreement with our calculation. A neutron diffraction structure of orthorhombic L-cysteine is known,³⁸ and although it is not exactly relevant since it does not contain peptide groups, it probably can serve to verify the general range for $\nu(\text{CS})$. The unit cell has molecules whose conformation is close to GH and G'H, and the observed $\nu(\text{CS})$ mode is at ~ 690 cm⁻¹.^{18,28} Our predicted value, averaging over ϕ, ψ , is ~ 680 cm⁻¹, which is probably reasonable given the different structures. In monoclinic L-cysteine,³⁹ the unit cell has molecules whose conformations are close to $\chi^1 = \text{H}$ and $\chi^1 = \text{N}$ [the x-ray structure

determination did not identify the (S)H], and $\nu(\text{CS})$ Raman bands are observed at 677 and 737 cm⁻¹.¹⁸ The former frequency is clearly consistent with a $\chi^1 = \text{H}$ conformation and the latter with a $\chi^1 = \text{N}$ conformation (but not with $\chi^1 = \text{C}$, as possibly suggested¹⁸). These examples thus provide support for the validity of the calculated frequencies given in Table XI as being reliable indicators of cysteine residue geometry.

CONCLUSIONS

The development of a reliable ab initio force field for the -CH₂SH group has enabled us to make definitive predictions about the conformational dependence of $\nu(\text{SH})$ and $\nu(\text{CS})$ of the cysteine residue in proteins. In the case of the strong Raman $\nu(\text{SH})$ mode, the frequency shift due to hydrogen bonding is superimposed on a conformational variation, so separating these two effects must be taken into account. For the weaker Raman $\nu(\text{CS})$ mode, it is of course necessary that a correct identification of this band be made in the spectrum. If this is done, the calculated frequencies given here can provide a more specific assignment of conformation than was previously possible on the basis of simple model compound correlations.

This research was supported by NSF grants DMB-8816756 and DMR-9110353.

REFERENCES

1. Bare, G. H., Alben, J. O. & Bromberg, P. A. (1975) *Biochemistry* **14**, 1578–1583.
2. Moh, P. P., Fiamingo, F. G. & Alben, J. O. (1987) *Biochemistry* **26**, 6243–6249.
3. Yu, N.-T. & East, E. J. (1975) *J. Biol. Chem.* **250**, 2196–2202.
4. Chen, W., Nie, S., Kuck, J. F. R., Jr. & Yu, N.-T. (1991) *Biophys. J.* **60**, 447–455.
5. Pande, J., McDermott, M. J., Callender, R. H. & Spector, A. (1989) *Arch. Biochem. Biophys.* **269**, 250–255.
6. Byler, D. M., Susi, H. & Farrell, H. M., Jr. (1983) *Biopolymers* **22**, 2507–2511.
7. Thomas, G. J., Jr., Li, Y., Fuller, M. T. & King, J. (1982) *Biochemistry* **21**, 3866–3878.
8. Li, T., Chen, Z., Johnson, J. E. & Thomas, G. J., Jr. (1990) *Biochemistry* **29**, 5018–5026.
9. Smith, D., Devlin, J. P. & Scott, D. W. (1968) *J. Mol. Spectrosc.* **25**, 174–184.

10. Richter, W. & Schiel, D. (1984) *Chem. Phys. Lett.* **108**, 480-483.
11. Li, H. & Thomas, G. J., Jr. (1991) *J. Am. Chem. Soc.* **113**, 456-462.
12. Hayashi, M., Shiro, Y. & Murata, H. (1966) *Bull. Chem. Soc. Jpn.* **39**, 112-117.
13. Scott, D. W. & El-Sabban, M. Z. (1969) *J. Mol. Spectrosc.* **30**, 317-337.
14. Inagaki, F., Harada, I. & Shimanouchi, T. (1973) *J. Mol. Spectrosc.* **46**, 381-396.
15. Manocha, A. S., Fateley, W. G. & Shimanouchi, T. (1973) *J. Phys. Chem.* **77**, 1977-1981.
16. Durig, J. R., Bucy, W. E., Wurrey, C. J. & Carreira, L. A. (1975) *J. Phys. Chem.* **79**, 988-993.
17. Pennington, R. E., Scott, D. W., Finke, H. L., McCullough, J. P., Messerley, J. F., Hossenlopp, I. A. & Waddington, G. (1956) *J. Am. Chem. Soc.* **78**, 3266-3272.
18. Ozaki, Y., Sugeta, H. & Miyazawa, T. (1975) *Chem. Lett.* 713-716.
19. Sugeta, H., Go, A. & Miyazawa, T. (1972) *Chem. Lett.* 83-86.
20. Sugeta, H., Go, A. & Miyazawa, T. (1973) *Bull. Chem. Soc. Jpn.* **46**, 3407-3411.
21. Nogami, N., Sugeta, H. & Miyazawa, T. (1975) *Chem. Lett.* 147-150.
22. Sugeta, H. (1975) *Spectrochim. Acta* **31A**, 1729-1737.
23. Nogami, N., Sugeta, H. & Miyazawa, T. (1975) *Bull. Chem. Soc. Jpn.* **48**, 2417-2420.
24. Qian, W., Zhao, W. & Krimm, S. (1991) *J. Mol. Struct.* **250**, 89-102.
25. Qian, W. & Krimm, S. (1992) *Biopolymers*, **32**, 321-326.
26. Qian, W. & Krimm, S. (1992) *Biopolymers*, **32**, 1025-1033.
27. Qian, W. & Krimm, S., to be published.
28. Susi, H., Byler, D. M. & Gerasimowicz, V. (1983) *J. Mol. Struct.* **102**, 63-79.
29. Zhao, W., Bandekar, J. & Krimm, S. (1988) *J. Am. Chem. Soc.* **110**, 6891-6892.
30. Zhao, W. & Krimm, S. (1990) *J. Mol. Struct.* **224**, 7-20.
31. Zhao, W., Bandekar, J. & Krimm, S. (1990) *J. Mol. Struct.* **238**, 43-54.
32. Zuppiroli, G., Perchard, C., Baron, M. H. & de Loze, C. (1980) *J. Mol. Struct.* **69**, 1-16.
33. Mirkin, N. G. & Krimm, S. (1991) *J. Am. Chem. Soc.* **113**, 9742-9747.
34. Krimm, S. & Bandekar, J. (1986) *Adv. Protein Chem.* **38**, 181-364.
35. Fogarasi, G. & Pulay, P. (1985) *Vibrational Spectra and Structure*, Vol. 14, Durig, J. R., Ed., Elsevier Amsterdam, Oxford, New York, Tokyo pp. 125-219.
36. Snyder, R. G. & Schachtschneider, J. H. (1965) *Spectrochim. Acta* **21**, 169-195.
37. Görbitz, C. H. (1987) *Acta Chem. Scand.* **B41**, 362-366.
38. Kerr, K. A., Ashmore, J. P. & Koetzle, T. F. (1975) *Acta Cryst. B* **31**, 2022-2026.
39. Harding, M. M. & Long, H. A. (1968) *Acta Cryst. B* **24**, 1096-1102.

Received February 25, 1992

Accepted April 30, 1992