Vibrational Studies of the Disulfide Group in Proteins. Part V. Correlation of SS Stretch Frequencies with the CCSS Dihedral Angle in Known Protein Disulfide Bridges

WEILI QIAN and SAMUEL KRIMM

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

SYNOPSIS

Normal mode calculations have been done on 92 disulfide bridges in 25 known protein structures in order to correlate the SS stretch frequency with the CCSS dihedral angle. It is possible to classify the frequencies into four major categories, which provide a more detailed classification scheme than previously proposed from dialkyl disulfide correlations.

INTRODUCTION

Early alkyl disulfide studies[1] had established that the frequency of the Raman-active SS stretch mode, ν(SS), is correlated with the CCSS dihedral angle, χ^2: ν(SS) was found near 510, 525, or 540 cm^{-1} when the conformation had two, one, or no H atoms, respectively, *trans* to distal S atoms across the CS bond.

While such a relation may provide a rough spectrum–structure correlation, it cannot capture the subtleties of S–S bridge conformations in proteins, such as the dependence of ν(SS) on $\chi^2 \neq \pm 60°, 180°$, as well as possible changes due to variations in the NCCS dihedral angle, χ^1, and in the ϕ and ψ of the adjacent peptide groups. Such effects can only be revealed by reliable normal mode calculations on the relevant structures.

We have obtained a conformation-dependent force field for the disulfide group by scaling ab initio force constants to experimental frequencies of alkyl disulfides.[2-4] When combined with an empirical force field for the peptide group,[5] calculated ν(SS) frequencies of SS bridges in seven known peptide and protein structures are found to be in very good agreement with observed Raman bands associated with these modes.[6] This combined force field thus provides a sound basis for studying the structure–spectrum relationship referred to above.[1]

As a guide to a more extensive vibrational study of the χ^2, χ^1, ϕ, ψ, conformational space of the disulfide group,[7] we have chosen to examine first the spectral characteristics of known protein S–S bridges. These structures are an indication of the kinds of conformations that may be common in proteins, and a classification of their ν(SS) frequencies would therefore be useful in establishing preliminary correlations as well as in providing the basis for further spectral studies of these proteins.

A recent survey has been made of the geometric characteristics of 72 S–S bridges in 22 proteins,[8] and these were found to fall into eight families based on C^α–C^α and C^β–C^β distances and the chirality of the bridge (the CSSC dihedral angle, χ^3, is usually near 90°, with about an equal number of right-handed and left-handed bridges). We have calculated the ν(SS) of 92 S–S bridges in 25 proteins, whose structures were taken from the Protein Data Bank[9] (55 were the same as in Ref. 8), and although the above classification[8] is not relevant for vibrational frequencies, we find that the ν(SS) fall into reasonably well-defined classes, in several respects significantly different from those given earlier.[1]

CALCULATIONS

As in our previous study,[6] normal mode calculations were performed on the unit $[SCH_2CH(CONHC)(NHCOC)]_2$. Values for all 9 dihedral angles—viz., $\phi_1, \psi_1, \chi^1_1, \chi^2_1, \chi^3, \chi^2_2, \chi^1_2, \phi_2, \psi_2$—were

Biopolymers, Vol. 32, 321–326 (1992)

CCC 0006-3525/92/040321-06$04.00

Table I Calculated SS Stretch Frequencies (in cm^{-1}) of Disulfide Bridges in Proteins

Type	$\chi_1^2\chi_2^{2a}$	$\chi_1^1\chi_2^{1b}$	Chir[c]	v (SS)[d]	Protein[e]	Bridge[f]
1	G'G	CN	R	503 (91)	1BP2	61–91
	G'G	CN	R	505 (91)	2OVO	24–56
	G'G	CN	R	504 (95)	1SN3	12–65
2	GB'	CN	L	505 (88)	9PAP	22–63
	GB'	NC	R	504 (91)	9PAP	56–95
	GB'	CC	R	503 (92)	1TPP	22–157
	GB'	NC	R	507 (93)	3INS	A7–B7
	GB'	NC	R	505 (90)	3INS	C7–D7
3	G'B	CH	R	506 (82)	1ZL1	65–81
	G'B	CC	R	504 (85)	5RSA	65–72
	G'B	CH	R	512 (78)	2SGA	191–220
	G'B	CC	R	505 (73)	3RP2	A168–A182
4	B'B	CC	R	503 (90)	2APP	249–283
5	B'B'	CC	R	506 (98)	1TPP	136–201
	B'B'	CC	R	506 (97)	2ALP	137–159
6	GG	CC	L	507 (87)	5RSA	40–95
	GG	NC	R	506 (88)	1TPP	168–182
	GG	CC	L	509 (87)	2AZA	A3–A26
	GG	CC	L	509 (95)	2AZA	B3–B26
	GG	CC	L	508 (97)	3EBX	3–24
	GG	CC	L	507 (92)	1NXB	17–41
	GG	CC	L	508 (93)	3INS	A20–B19
	GG	CC	L	507 (71)	2OVO	8–38
	GG	CC	L	507 (87)	1SN3	25–46
	GG	NN	R	508 (93)	1SN3	29–48
7	GB	CN	L	508 (88)	1LZ1	30–116
	GB	CC	L	506 (84)	5RSA	26–84
	GB	CC	L	508 (97)	1CRN	3–40
	GB	CN	L	509 (93)	1CRN	16–26
	GB	CC	L	506 (94)	5PTI	5–55
	GB	CC	L	510 (96)	3EBX	17–41
	GB	CC	L	510 (97)	1NXB	3–24
	GB	CC	L	511 (82)	1BP2	27–123
	GB	CN	L	508 (92)	1BP2	44–105
	GB	CN	L	508 (90)	1BP2	51–98
	GB	CN	L	509 (95)	2OVO	16–35
	GB	NH	R	510 (84)	1SN3	16–41
	GB	NC	R	510 (90)	3RP2	B168–B182
8	SB'	NH	R	510 (85)	3FAB	L213–H220
9	B'S'	CC	R	510 (96)	1ACX	34–43
	B'S'	CC	R	510 (91)	3RP2	A136–A201
	B'S'	CC	R	509 (92)	4CHA	A136–A201
	B'S'	CC	R	509 (95)	4CHA	B136–B201
	B'S'	CC	R	511 (96)	2CGA	A136–A201
	B'S'	CC	R	511 (97)	2CGA	B136–B201
	B'S'	CC	R	509 (95)	1CRN	4–32
	B'S'	NC	L	507 (88)	3GRS	58–63
10	BB	HH	R	510 (87)	3EBX	55–60
	BB	HH	R	519 (75)	1NXB	55–60
	BB	NN	R	505 (87)	1BP2	11–77
	BB	CN	L	508 (86)	1BP2	84–96
	BB	CN	L	513 (89)	5PTI	30–51
11	GA	CC	L	510 (86)	3INS	C20–D19
	GA	CC	L	511 (92)	1LZ1	6–128
12	SS'	CH	R	507 (66)	5PTI	14–38
13	GS	CC	L	514 (91)	5RSA	58–110
14	G'D	CC	R	516 (77)	1BP2	29–45
	G'D	CC	R	513 (63)	5CPA	138–161
	G'D	CH	R	517 (78)	1TPP	128–232

Table I (*Continued*)

Type	$\chi_1^2\chi_2^{2a}$	$\chi_1^1\chi_2^{1b}$	Chir[c]	v (SS)[d]	Protein[e]	Bridge[f]
15	G'D'	CH	R	519 (54)	3INS	C6–C11
	G'D'	CC	R	512 (80)	3RP2	B136–B201
16	BD	CC	L	517 (84)	1TPP	42–58
	BD	CC	L	518 (68)	2ALP	42–58
	BD	CC	L	519 (69)	3RP2	A42–A58
	BD	CC	L	519 (77)	3RP2	B42–B58
	BD	CC	L	519 (84)	4CHA	A42–A58
	BD	CC	L	519 (85)	4CHA	B42–B58
	BD	CC	L	518 (84)	2CGA	A42–A58
	BD	CC	L	517 (84)	2CGA	B42–B58
17	B'T	CH	R	529 (25)	2CGA	A191–A220
				518 (51)		
18	AT	CH	L	526 (65)	9PAP	153–200
19	G'T	CH	R	524 (55)	3INS	A6–A11
				468 (22)		
20	GT	NC	R	528 (81)	4CHA	A191–A220
	GT	NC	R	524 (83)	1LZ1	77–95
	GT	NC	R	521 (74)	1TPP	191–220
	GT	NC	R	527 (75)	3EBX	43–54
	GT	CC	L	527 (70)	1ACX	83–88
	GT	NC	R	529 (80)	4CHA	B191–B220
	GT	CC	L	521 (79)	2CGA	B191–B220
21	D'T	NC	L	537 (26)	4CHA	A168–A182
				526 (49)		
	D'T	NC	L	539 (24)	4CHA	B168–B182
				526 (47)		
22	SD	NH	R	557 (21)	3FAB	L22–L87
				526 (56)		
	SD	CC	L	535 (29)	2SGA	42–58
				516 (47)		
	SD	NH	R	556 (33)	2RHE	22–89
				515 (32)		
	SD	HC	L	547 (42)	1REI	A23–A88
				521 (31)		
	SD	NH	R	552 (39)	1REI	B23–B88
				518 (35)		
23	DT	NN	L	544 (74)	3FAB	H144–H200
24	BT	NC	R	547 (27)	1NXB	43–54
				525 (40)		
				506 (26)		
25	TT	NN	L	551 (68)	3FAB	L136–L195
				507 (15)		
	TT	NN	L	549 (65)	3FAB	H22–H95
				521 (16)		
	TT	NC	L	541 (23)	2CGA	A168–A182
				529 (49)		
	TT	NC	L	540 (33)	2CGA	B168–B182
				529 (39)		

[a] $C^{\alpha}C^{\beta}SS$ dihedral angles (based on right-handed χ^3; see text): C (0°), A (30°), G (60°), B (90°), S (120°), D (150°), T (180°); prime: negative angle.

[b] $NC^{\alpha}C^{\beta}S$ dihedral angles: C—C *trans* to S ($\chi^1 = -60°$); N—N *trans* to S ($\chi^1 = 180°$); H—H *trans* to S ($\chi^1 = 60$).

[c] Chirality of S—S bridge: R—right handed ($\chi^3 > 0$); L—left-handed ($\chi^3 < 0$).

[d] Numbers in parentheses: potential energy distribution in SS stretch.

[e] Protein Data Bank Code: 1ACX, actinoxanthin; 1BP2, phospholipase A2; 1CRN, crambin; 1LZ1, lysozyme; 1NXB, neurotoxin B; 1REI, Bence–Jones protein (REI variable domain); 1SN3, scorpion neurotoxin (variant 3); 1TPP, trypsin complex with APPA; 2ALP, α-lytic protease; 2APP, acid proteinase (penicillium J); 2AZA, azurin; 2CGA, chymotrypsinogen A; 2OVO, ovomucoid third domain; 2RHE, Bence–Jones protein (λ, variable domain); 2SGA, proteinase A; 3EBX, erabutoxin B; 3FAB, immunoglobulin FAB*; 3GRS, glutathione reductase; 3INS, insulin (porcine); 3RP2, rat mast cell proteinase II; 4CHA, α-chymotrypsin; 5CPA, carboxypeptidase A; 5RSA, ribonuclease A; 5PTI, trypsin inhibitor; 9PAP, papain.

[f] Cystine residues connected by S–S bridge.

taken from the bridge structures in the Protein Data Bank.[9] Bond lengths and angles for the peptide groups were taken as standard values,[5] and the comparable quantities for the S–S bridge were those of diethyl disulfide.[4,6]

The force field consisted of two parts, as in our earlier study[6]: an empirical force field for the peptide group[5] and an ab initio conformation-dependent force field for the S–S bridge part.[4,6] In the latter, force constants were chosen in the following categories of the χ^2 dihedral angle[6]: C(0°), A(30°), G(60°), B(90°), S(120°), D(150°), T(180°), D′(−150°), S′(−120°), B′(−90°), G′(−60°), and A′(−30°).

RESULTS AND DISCUSSION

The ν(SS) frequencies resulting from full normal mode calculations on our model structure for each of the 92 S–S bridges are given in Table I. The conformations are grouped according to χ_1^2 and χ_2^2 (χ^3, which is usually designated by G or G′, has been left out, so GB′ would be GGB′ in a $\chi_1^2\chi^3\chi_2^2$ designation[6]). For convenience, all conformations are listed in terms of a right-handed S–S bridge [i.e., G′G′B = GB′(L)], although of course the calculations were done for the actual structures,[6] whose chiralities are given in the table. We also give the χ_1^1 and χ_2^1 values for each bridge, in terms of the atoms *trans* to S across the $C^\alpha C^\beta$ bonds. The potential energy distribution (PED) of each ν(SS) mode is given in parentheses after the frequency.

The frequency ranges obtained for each of the 25 observed types are presented on a conformational map in Table II. In a few cases we find that SS stretch contributes to more than one mode of a given bridge. Following our earlier experience,[6] we have included (as potentially observable) only those of such modes that have a PED for SS stretch larger than about half that of the maximum for the bridge. The conformational types are arranged in Table II in rough order of decreasing frequency ranges.

As can be seen from these tables, many conformational types are not represented. In particular, no conformations are found with χ^2 = C or A′, nor is S′S′ observed. These correspond to 7 of the 9 highest energy bridge structures,[3] and of the other two, SS and SS′, only one is found. Nor are structures with very short C^α–C^α distances and intermediate energies,[3] G′G′ and S′G′, observed (although small variants of these, B′B′ and S′B′, are found). Of the remaining 5 structures of intermediate energies[3]—GS′, SG′, TS′, GS, and ST—only one (GS) is found, but there are a number of observed cases in which one of the angles varies by ~ 30°. All of the low-energy conformations[3]—GG′, TG′, GT, GG, and TT—are observed, as well as some small variants of these. Thus, while low-energy S–S bridge conformations predominate in proteins, higher energy structures do occur, obviously com-

Table II Conformational Map of Calculated SS Stretch Frequency Ranges (in cm^{-1}) of Disulfide Bridges in Proteins

	χ_1^{2a}											
χ_2^{2a}	T	D	D′	C	A	A′	S	S′	B	G	B′	G′
T	529 540–51	544	526		526				525	521–529	518	524
D							515–26 535–56		517–519			513–517
D′												512–519
C												
A										510–511		
A′												
S								507		514	510	
S′											507–511	
B									505–519	506–511	503	504–506
G										506–509	503–507	503–505
B′											506	
G′												

[a] $C^\alpha C^\beta$SS dihedral angles (based on right-handed χ^3; see text): C (0°), A (30°), G (60°), B (90°), S (120°), D (150°), T (180°); prime: negative angle.

pensated by structural requirements in other parts of the protein. This means that when a protein structure changes (for whatever reasons), it can be expected that higher energy S–S bridges may relax to lower energy local conformations. Such structural changes may be usefully monitored if we have a detailed understanding of the ν(SS)–conformation relationship.

It is also worth noting that the frequency range for an observed conformational type is generally fairly small, indicating that there may either be a small dependence on parameters other than χ^2 (which is not suggested by more general conformational results[7]) or that some correlations with χ^1 (and perhaps χ^1 with ϕ,ψ) may exist. In fact, we see that of 23 GG and GB structures, 19 are of left-handed chirality and of these 14 have $\chi_1^1\chi_2^1$ = CC with the other 5 being CN; of 8 B'S' structures, 7 are right-handed with $\chi_1^1\chi_2^1$ = CC; and of 8 BD structures, all are left-handed with $\chi_1^1\chi_2^1$ = CC. Some of these preferences have already been noted.[8,10]

An initial view of the tables shows that, of the canonical structures, GG and GT fit the earlier classification[1] while TT conformations do not: for $\chi_1^1\chi_2^1$ = NN the observable band is near 550 cm^{-1} while for NC the predominant mode is at 529 cm^{-1} with another mode near 540 cm^{-1} being potentially observable. Deviations in dihedral angle of $\sim$ 30° are easily tolerable for the GG type (and in one case, B'S', up to 60°) without seriously affecting ν(SS), which means that this region ($\sim$510 cm^{-1}) may not be a sensitive determinant of conformation. For structures of the GT type, the situation is different with respect to deviations in dihedral angles: while some ν(SS) fall within the expected range ($\sim$ 525 cm^{-1}), others are clearly outside (G'D) or exhibit splittings (SD) that place them in a different category. For the TT type, one variant structure (D'T) has a frequency (526 cm^{-1}) that places it in a category to which it does not belong. Thus, we see that caution is clearly needed in applying the proposed classification[1] to the determination of S–S bridge conformation from the ν(SS) frequency.

If this classification[1] cannot be depended on in detail, what can we say about the structural implications of observed ν(SS) bands? A histogram of the frequency distribution in Table I shows that the ν(SS) of each conformation are mostly localized in a relatively small frequency region (including each component of the split modes) and that the distribution has four regions in which the frequencies fall: 503–512, 512–521, 521–529, and 535–556 cm^{-1}. This leads us to propose the classification scheme given in Table III, which can serve as a reasonable starting point for categorizing conformation from spectrum.

As noted above, the 503–512-cm^{-1} region contains many conformations related to GG by small ($\sim$ 30°) variations in χ_1^2 and/or χ_2^2. We have nevertheless subdivided this class into 1a (503–505 cm^{-1}) and 1b (506–512 cm^{-1}) since the predominant number of examples (10 of 13) of the 1a conformations fall in this limited low-frequency range. In all, 53 of 55 examples are encompassed in the entire range. The exception is BB, which contributes also at 513 and 519 cm^{-1} [the latter possibly because of a split mode at 491 (23) cm^{-1}, perhaps because of unusual χ_1^1 = 48.6°, χ_2^1 = 89.3°], and in general seems to have broadly spread frequencies. It is interesting that some conformations in class 1b span limited frequency ranges (9 of 10 examples of GG fall in the 507–509-cm^{-1} range, and 10 of 12 examples of SB', B'S', and GA fall in the 509–511-cm^{-1} range), while in one case (B'B') both examples fall at the same frequency (506 cm^{-1}).

All of the examples of class 2 conformations fall in the 512–521-cm^{-1} range, including the low-frequency components of the SD structure. Similarly, all of the examples of the class 3 conformations fall in the 521–529-cm^{-1} range, including (where appropriate) the low-frequency components of BT and

Table III SS Stretch Frequency Classes of Disulfide Bridges in Proteins

Class	Frequency Range (cm^{-1})	Conformations[a]
1a	503–505	G'G, GB', G'B, B'B
1b	506–512	B'B', GG, GB, SB', B'S', BB, GA, SS'
2	512–521	GS, G'D, G'D', BD, B'T, (SD)
3	521–529	AT, G'T, GT, D'T, SD, (BT), (TT)
4	535–556	(SD), DT, (BT), (TT)

[a] Letters in parentheses: Conformation gives split v (SS) modes.

TT (BT may also contribute an observable band at 506 cm^{-1}).

We place in class 4 the high-frequency bands, which are primarily of those conformations that can give rise to split ν(SS) modes. Such splittings and the relative PEDs seem to be sensitive to both χ^1 and ϕ,ψ[7], as can be seen from Table I. In one case, DT, no splitting occurs, perhaps because of the unusual $\chi_1^1 = 126.3°$, $\chi_2^1 = 109.3°$ in this case. As noted above, the frequencies of this class do not fit into the simple classification proposed earlier.[1]

CONCLUSIONS

On the basis of satisfactory agreement between observed Raman bands and ν(SS) modes calculated for known S–S bridges,[6] we feel that the present calculations of such modes for 92 S–S bridges in known protein structures[9] provides an accurate representation of the relationship between the ν(SS) frequency and the CCSS dihedral angles of the bridge.

We see that this relationship is more complex than previously proposed,[1] but that a classification scheme is nevertheless possible. Although the correlations somewhat complicate the spectrum–conformation deductions, they do provide deeper insights into the conclusions that can be drawn. This could be valuable in interpreting the detailed conformational changes associated with frequency shifts in ν(SS) modes that result from changes in the structure of a protein.

This research was supported by NSF grants DMB-8816756 and DMR-8806975.

REFERENCES

1. Sugeta, H., Go, A. & Miyazawa, T. (1972) *Chem. Lett.* 83.
2. Zhao, W., Bandekar, J. & Krimm, S. (1988) *J. Am. Chem. Soc.* **110,** 6891–6892.
3. Zhao, W. & Krimm, S. (1990) *J. Mol. Struct.* **224,** 7–20.
4. Zhao, W., Bandekar, J. & Krimm, S. (1990) *J. Mol. Struct.* **238,** 43–54.
5. Krimm, S. & Bandekar, J. (1986) *Adv. Protein Chem.* **38,** 181–364.
6. Qian, W., Zhao, W. & Krimm, S. (1991) *J. Mol. Struct.* **250,** 89–102.
7. Qian, W. & Krimm, S., *Biopolymers,* in press.
8. Srinivasan, N., Sowdhamini, R., Ramakrishnan, C. & Balaram, P. (1990) *Int. J. Peptide Protein Res.* **36,** 147–155.
9. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542.
10. Richardson, J. S. (1981) *Adv. Protein Chem.* **34,** 167–239.

Received June 10, 1991
Accepted August 6, 1991