Allosteric Disulfide Bonds[†]

Bryan Schmidt,[‡] Lorraine Ho,[‡] and Philip J. Hogg*,^{‡,§}

Centre for Vascular Research, University of New South Wales and Department of Haematology, Prince of Wales Hospital, Sydney, New South Wales 2052, Australia, and Children's Cancer Institute Australia for Medical Research, Randwick, New South Wales 2031, Australia

Received February 13, 2006; Revised Manuscript Received April 9, 2006

ABSTRACT: Disulfide bonds have been generally considered to be either structural or catalytic. Structural bonds stabilize a protein, while catalytic bonds mediate thiol—disulfide interchange reactions in substrate proteins. There is emerging evidence for a third type of disulfide bond that can control protein function by triggering a conformational change when it breaks and/or forms. These bonds can be thought of as allosteric disulfides. To better define the properties of allosteric disulfides, we have analyzed the geometry and dihedral strain of 6874 unique disulfide bonds in 2776 X-ray structures. A total of 20 types of disulfide bonds were identified in the dataset based on the sign of the five χ angles that make up the bond. The known allosteric disulfides were all contained in 1 of the 20 groups, the -RHStaple bonds. This bond group has a high mean potential energy and narrow energy distribution, which is consistent with a functional role. We suggest that the -RHStaple configuration is a hallmark of allosteric disulfides. About 1 in 15 of all structurally determined disulfides is a potential allosteric bond.

A significant fraction of the proteins made in nature contain disulfide bonds. About 10% of the proteins made by mammalian cells, for instance, contain disulfide bonds. These bonds have been mostly thought of as structural motifs that have been added during evolution to help stabilize the tertiary or quaternary structure of proteins. The bonds assemble during maturation of the protein and in most cases appear to remain unchanged for the life of the protein. Certain disulfide bonds, however, have evolved to control how proteins function by switching between bonded and unbonded states (I, 2).

The best characterized functional disulfides are those at the active sites of thiol—disulfide oxidoreductases (3). These dithiols/disulfides undergo a cycle of oxidation/reduction reactions with thiols or disulfides in a protein substrate, resulting in net formation, reduction, or isomerization of disulfide bond(s) in the substrate. It is becoming apparent, however, that some disulfide bonds regulate protein function in a nonenzymatic way by triggering changes in the intraor intermolecular structure of proteins (1). We refer to these types of bonds as allosteric disulfides, to distinguish them from the catalytic disulfides.

To define the properties of allosteric disulfides and to estimate their incidence, we have characterized the geometry and associated potential energy of 6874 unique disulfide bonds. A total of 20 bond groups were identified on the basis of their geometry, and all of the known and suspected

allosteric disulfides were contained in one of the groups, the high-energy —RHStaples.

EXPERIMENTAL PROCEDURES

There were 33,019 disulfide bonds in 6165 X-ray structures in the PDB database as of September 1, 2005. The database was refined to include only one example of each disulfide. The highest resolution structure was chosen when there were two or more examples of a disulfide in the PDB. This refined database contained 6874 disulfide bonds in 2776 X-ray structures. A script written in C shell was used to automate the execution of a program written in python to parse and process all of the PDB files in this database.

The dihedral strain energy (DSE) of each disulfide was predicted from the magnitude of the five χ angles (Figure 1A) using the following empirical formula (4, 5):

DSE (kJ mol⁻¹) = 8.37(1 + cos
$$3\chi_1$$
) +
8.37(1 + cos $3\chi_1'$) + 4.18(1 + cos $3\chi_2$) +
4.18(1 + cos $3\chi_2'$) + 14.64(1 + cos $2\chi_3$) +
2.51(1 + cos $3\chi_3$) (1)

This relationship only considers the dihedral angles. Other factors such as bond lengths and van der Waals contacts are not part of the calculation. Nevertheless, this relationship has been shown to provide useful semiquantitative insights into the amount of strain in a disulfide bond (6-9).

The combination of χ angles that give the highest and lowest dihedral strain energies was determined by partial differentiation of eq 1 with respect to the five χ angles and using the Hessian matrix of f to determine the nature of the critical points. Minimum and maximum dihedral strain energies of 2.0 and 84.5 kJ mol⁻¹, respectively, are predicted. These values correspond to $\chi_1, \chi_2, \chi_3, \chi_2'$, and χ_1' torsional

[†] This work was supported by the National Health and Medical Research Council of Australia, the Australian Research Council, and the New South Wales Cancer Council and Health Department. Datasets will be made available upon request.

^{*} To whom correspondence should be addressed. Telephone: 61-2-9385-1004. Fax: 61-2-9385-1389. E-mail: p.hogg@unsw.edu.au.

[‡] University of New South Wales and Department of Haematology.

[§] Children's Cancer Institute Australia for Medical Research.

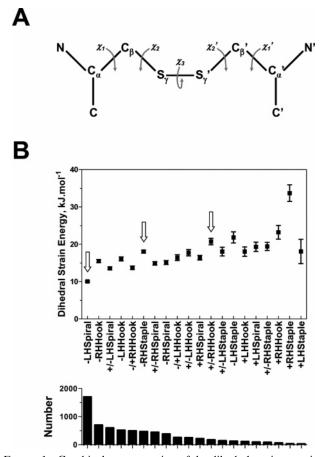
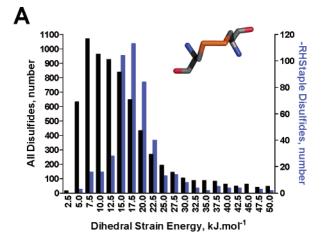


FIGURE 1: Graphical representation of the dihedral strain energies of the 20 types of disulfides. (A) Five χ angles of the disulfide bond. (B) Mean and 95% confidence intervals of the dihedral strain energies of the 20 types of disulfide bonds. Refer to Table 1 for statistical analysis. The arrows indicate the major disulfide types that are discussed in the text.

angles of 60° , 60° , $+/-83^{\circ}$, 60° , and 60° (minimum) and 0° , 0° , 0° , 0° , and 0° (maximum). For reference, the disulfide with the lowest strain energy in the database is the Cys203—Cys259 bond of the H-2K^b MHC class Ia molecule (PDB ID 1N59; 2.1 kJ mol⁻¹) and the disulfide with the highest strain energy is the Cys37—Cys54 bond of the 3-oxoacylacyl carrier protein reductase from *Thermotoga maritima* (PDB ID 1O5I; 72.4 kJ mol⁻¹).

RESULTS AND DISCUSSION

Disulfide bond geometry is described in terms of the five χ angles (Figure 1A). By convention, there are three basic types of disulfide bonds based on the sign of the χ_2 , χ_3 , and χ_2 torsional angles. These are the spirals, hooks, or staples. The disulfides are further characterized as right- or lefthanded according to whether the χ_3 angle is positive or negative, respectively. Strictly speaking, the spiral, hook, or staple definition should only apply to disulfides where the χ_2 angle is between 30° and 90° (gauche+) or -30° and -90° (gauche-) (10, 11). A χ_2 angle between 150° and 180° or -150° and -180° is considered to be trans+ or trans-, respectively. There are many examples, however, of disulfides that do not satisfy these gauche limits of the χ_2 angle. To avoid complicating the categories of disulfides unnecessarily, we chose to define the χ_2 angle as simply being either positive or negative. In the final analysis, this was not found to be a limitation (see below).



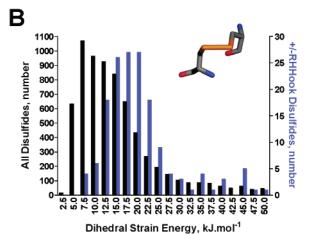


FIGURE 2: Dihedral strain energies of all disulfides compared to —RHStaple and +/—RHHook disulfides. The insets are an example of a —RHStaple or +/—RHHook disulfide bond. The inset in A is the Cys429—Cys453—RHStaple bond of botulinum neurotoxin A (PDB ID 3BTA), and the inset in B is the Cys30—Cys33 +/—RHHook bond of thioredoxin (PDB ID 1R26). The structures look down upon the S—S bond, which is shown in the horizontal position. The structures were generated using PyMol (28).

The six possible disulfide bond configurations based on the sign of the χ_2 , χ_3 , and χ_2' angles are all represented in protein structures (10, 11). We included the χ_1 and χ_1' torsional angles in the current analysis to further refine the disulfide bond structures. This expanded the number of possible configurations from 6 to 20. The standard definitions have been amended to reflect the sign of the χ_1 and χ_1' torsional angles (Table 1). For example, a disulfide is a minus right-handed staple (-RHStaple) if the χ_1 , χ_2 , χ_3 , χ_2' , and χ_1' angles are -,-,+,-, and -, respectively. The disulfides are treated as symmetrical. For example, a disulfide is a +/-RHStaple if the χ_1 , χ_2 , χ_3 , χ_2' , χ_1' angles are +,-,+, -,-, or -,-,+,-, +.

The geometry of 6874 unique disulfide bonds in 2776 X-ray structures was characterized from the five χ angles, and the potential energy of each bond was estimated using eq 1. All 20 configurations are represented in the database, with 17 of the bond types having an incidence of >1% (Figure 1B and Table 1). Analysis of the dataset when entries were restricted to disulfides in structures with resolutions of ≤ 3 , ≤ 2.5 , or ≤ 2 Å resulted in the same general conclusions about the bond groups (data not shown).

The -LHSpiral configuration is the most prevalent (24.8%) and has the lowest DSE (10.1 kJ mol⁻¹) and a

Table 1: Classification of Disulfide Bonds Based on Geometry^a

$\chi_1, \chi_2, \chi_3, {\chi_2}', {\chi_1}'$	number	incidence (%)	DSE (kJ mol ⁻¹)	Cα-Cα' (Å)	designation
-,-,-,-	1703	24.77	10.1 (9.7-10.4)	5.76 (5.74-5.78)	-LHSpiral
-,+,+,-,-	703	10.23	15.5 (14.7–16.3)	5.27 (5.23-5.31)	-RHHook
+, -, -, -, -	599	8.71	13.6 (12.9-14.3)	6.08 (6.04-6.11)	+/-LHSpiral
-,-,-,+,-	514	7.48	16.1 (15.1–17.1)	5.60 (5.54-5.66)	-LHHook
+,+,+,-,-	495	7.20	13.7 (12.8-14.7)	5.17 (5.12-5.21)	-/+RHHook
-,-,+,-,-	469	6.82	18.1 (17.3–18.8)	4.32 (4.26-4.37)	-RHStaple
+,+,+,+,-	443	6.44	14.9 (14.0-15.8)	6.06 (6.02-6.10)	+/-RHSpiral
-,+,+,-	384	5.59	15.1 (14.0-16.2)	5.93 (5.88-5.97)	-RHSpiral
+,+,-,-,-	254	3.70	16.4 (14.9-18.0)	5.43 (5.33-5.52)	-/+LHHook
+, -, -, +, -	249	3.62	17.8 (16.1-19.4)	6.13 (6.06-6.20)	+/-LHHook
+,+,+,+,+	215	3.13	16.4 (15.2-17.6)	6.29 (6.24-6.35)	+RHSpiral
+, -, +, +, -	171	2.49	20.8 (19.0-22.5)	5.77 (5.68-5.87)	+/-RHHook
+,+,-,+,-	142	2.07	18.1 (15.7-20.4)	5.39 (5.22-5.56)	+/-LHStaple
-,+,-,+,-	124	1.80	21.9 (18.9-24.8)	5.54 (5.36-5.72)	-LHStaple
+, -, -, +, +	111	1.61	18.1 (15.6-20.5)	6.00 (5.87-6.13)	+LHHook
+, -, -, -, +	90	1.31	19.3 (16.9-21.8)	6.38 (6.30-6.47)	+LHSpiral
+, -, +, -, -	84	1.22	19.4 (17.3-21.6)	5.25 (5.07-5.42)	+/-RHStaple
+,+,+,-,+	63	0.92	23.2 (19.6-26.9)	6.00 (5.83-6.17)	+RHHook
+, -, +, -, +	32	0.47	33.7 (29.2-38.3)	5.92 (5.53-6.32)	+RHStaple
+,+,-,+,+	29	0.42	18.1 (11.4-24.8)	5.81 (5.54-6.09)	+LHStaple
all disulfides	6874		14.8 (14.5-15.0)	5.63 (5.61-5.65)	

^a A total of 6874 unique disulfide bonds in 2776 X-ray structures were separated into 20 groups according to the sign of χ₁, χ₂, χ₃, χ₂', and χ₁' angles. The DSE and distance between the two α-carbon atoms were calculated for each disulfide, and the mean and 95% confidence intervals are shown for each group. The -RHStaple and +/-RHHook bonds contain the allosteric and catalytic disulfides, respectively.

narrow distribution of energies (95% CI, 9.7–10.4 kJ mol⁻¹) (Table 1). The -LHSpiral, therefore, can be considered the primary structural disulfide. The other 19 disulfide types have mean strain energies ranging from 13.6 kJ mol⁻¹ (95% CI, $12.9-14.3 \text{ kJ mol}^{-1}$) for the \pm LHSpiral group to 33.7 kJ mol^{-1} (95% CI, 29.2–38.3 kJ mol^{-1}) for the +RHStaple group.

The known or suspected allosteric disulfides are all contained in the -RHStaple group (Table 2). This group also contains two catalytic disulfides, the 207–212 disulfide in PapD (12) and the 103–109 disulfide in DsbD (13). Most of the catalytic bonds are found in the +/-RHHook group (Table 2). The mean DSE for the -RHStaple and +/-RHHook disulfides are 18.1 kJ mol⁻¹ (95% CI, 17.3-18.8 $kJ \text{ mol}^{-1}$) and 20.8 $kJ \text{ mol}^{-1}$ (95% CI, 19.0–22.5 $kJ \text{ mol}^{-1}$), respectively, compared to 14.8 kJ mol⁻¹ (95% CI, 14.5-15.0 kJ mol⁻¹) for all disulfides (Table 1). The high potential energy of these bonds is consistent with their functional role. High-energy disulfide bonds are more easily cleaved than bonds with lower stored energy (6, 7). The distribution of the strain energies of the -RHStaple and +/-RHHook bonds is narrow and unimodal (Figure 2), which further supports a particular role for these disulfides in controlling protein function.

The -LHStaple (21.9 kJ mol⁻¹; 95% CI, 18.9-24.8 kJ mol⁻¹), +RHHook (23.2 kJ mol⁻¹; 95% CI, 19.6–26.9 kJ mol^{-1}), and +RHStaple (33.7 kJ mol^{-1} ; 95% CI, 29.2–38.3 kJ mol⁻¹) disulfides have mean strain energies even higher than the -RHStaple (allosteric) and +/-RHHook (catalytic) bonds (Figure 1B and Table 1). The number of disulfides in these three groups is relatively small (<2% incidence), however, and the distribution of their strain energies in terms of 95% confidence intervals is broader than that of the -RHStaple and +/-RHHook bonds. Visual inspection of the distribution of the strain energies of the -LHStaple, +RHHook, and +RHStaple bonds, as described for the -RHStaple and +/-RHHook bonds (Figure 2), confirmed a broad and bi- or trimodal distribution (data not shown).

Table 2: Known or Suspected Allosteric and Catalytic Disulfide Bonds

Description	protein	disulfide-bond residues	PDB ID ^a				
botulinum neurotoxin A botulinum neurotoxin B botulinum neurotoxin B CD4 130-159 1CDY DsbD 103-109 1JPE gelsolin 188-201 1KCQ HIV gp120 126-196, 1G9M 296-331, 385-418 PapD 207-212 1NOL semaphorin 3A 132-141 1Q47 serum amyloid P tissue factor 186-209 2HFT +/-RHHook Ccmg copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDy 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus rotin disulfide isomerase from yeast ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 11-14 1H75 T1pA 72-75 1JFU Trp14 43-46 1WOU	F						
Dotulinum neurotoxin B			3BTA				
CD4 130-159 1CDY DsbD 103-109 1JPE gelsolin 188-201 1KCQ HIV gp120 126-196, 1G9M 296-331, 385-418 136-196, 1G9M PapD 207-212 1NOL semaphorin 3A 132-141 1Q47 serum amyloid P 36-95 1SAC tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E							
DsbD 103-109 gelsolin 1BE gelsolin HIV gp120 126-196, 1G9M 296-331, 385-418 296-331, 385-418 PapD 207-212 1NOL semaphorin 3A 132-141 1Q47 serum amyloid P 36-95 1SAC tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP 1QUP DsbA 30-33 1FVK 18EJ DsbDy 40-43 1UC7 105B DsbDy 40-43 1UC7 1V58 Erv2p 54-57 1JRA 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC 62-65 1OQC MerP 14-17 1OSD 11-14 1H75 Oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast ResA 73-76 1ST9 1ST9 Rp19 66-69 1SEN thioredoxin 130-33 1R26 thioredoxin 44-49 1F9M thioredoxin 2 32-35 1THX 1THX TlpA 72-75 1JFU 1TP14 43-46 1WOU							
HIV gp120 HIV gp120 126-196, 296-331, 385-418 PapD 207-212 1NOL semaphorin 3A 132-141 1Q47 serum amyloid P 36-95 1SAC tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDy 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin They are the sulf of the sulf or th	DsbD	103-109					
HIV gp120 126-196, 296-331, 385-418 PapD	gelsolin	188-201	1KCO				
PapD 207-212 1NOL		126 - 196.					
PapD semaphorin 3A 207-212 semaphorin 3A 132-141 serum anyloid P 1047 serum anyloid P serum amyloid P tissue factor 186-295 serum anyloid P 186-209 serum anyloid P +/-RHHook Ccmg 92-95 serum anyloid P copper chaperone for SOD 17-20 serum anyloid P DsbA 30-33 serum anyloid P DsbC 98-101 serum anyloid P DsbDy 40-43 serum anyloid P DsbG 109-112 serum anyloid P Erv2p 54-57 serum anyloid P FAD-dependent sulfhydryl oxidase 62-65 serum anyloid P MerP 14-17 serum anyloid P Nrdh redoxin 11-14 serum a	CI CI	296-331,					
semaphorin 3A serum amyloid P tissue factor +/-RHHook Ccmg copper chaperone for SOD DsbA DsbC DsbC DsbDy A FAD-dependent sulfhydryl oxidase Frv2p FAD-dependent sulfhydryl oxidase FAD-dependent sul		385-418					
serum amyloid P 36-95 1SAC tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDy 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	PapD	207-212	1N0L				
tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	semaphorin 3A	132 - 141	1Q47				
tissue factor 186-209 2HFT +/-RHHook Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	serum amyloid P	36-95	1SAC				
Ccmg 92-95 1KNG copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU		186 - 209	2HFT				
copper chaperone for SOD 17-20 IQUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDy 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	+/-RHHook						
copper chaperone for SOD 17-20 1QUP DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	Ccmg	92-95	1KNG				
DsbA 30-33 1FVK DsbC 98-101 1EEJ DsbDy 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	2	17 - 20	1QUP				
DsbDγ 40-43 1UC7 DsbG 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU		30-33					
DsbG' 109-112 1V58 Erv2p 54-57 1JRA FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	DsbC	98-101	1EEJ				
Erv2p 54–57 1JRA FAD-dependent sulfhydryl oxidase 62–65 1OQC MerP 14–17 1OSD Nrdh redoxin 11–14 1H75 oxidoreductase from A. pyrococcus 146–149 1A8L protein disulfide isomerase from yeast 61–64 2B5E ResA 73–76 1ST9 Rp19 66–69 1SEN thioredoxin 30–33 1R26 thioredoxin F 46–49 1F9M thioredoxin 2 32–35 1THX TlpA 72–75 1JFU Trp14 43–46 1WOU	$DsbD\gamma$	40-43	1UC7				
FAD-dependent sulfhydryl oxidase 62-65 1OQC MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	DsbG	109-112	1V58				
MerP 14-17 1OSD Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	Erv2p	54-57	1JRA				
Nrdh redoxin 11-14 1H75 oxidoreductase from A. pyrococcus 146-149 1A8L protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	FAD-dependent sulfhydryl oxidase	62-65	1OQC				
oxidoreductase from A. pyrococcus 146–149 1A8L protein disulfide isomerase from yeast 61–64 2B5E ResA 73–76 1ST9 Rp19 66–69 1SEN thioredoxin 30–33 1R26 thioredoxin F 46–49 1F9M thioredoxin 2 32–35 1THX TlpA 72–75 1JFU Trp14 43–46 1WOU		14 - 17	1OSD				
protein disulfide isomerase from yeast 61-64 2B5E ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	Nrdh redoxin	11 - 14	1H75				
ResA 73-76 1ST9 Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	oxidoreductase from A. pyrococcus	146 - 149	1A8L				
Rp19 66-69 1SEN thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	protein disulfide isomerase from yeast	61-64	2B5E				
thioredoxin 30-33 1R26 thioredoxin F 46-49 1F9M thioredoxin 2 32-35 1THX TlpA 72-75 1JFU Trp14 43-46 1WOU	ResA		1ST9				
thioredoxin F thioredoxin 2 32–35 1THX TlpA 72–75 1JFU Trp14 43–46 1WOU							
thioredoxin 2 32–35 1THX TlpA 72–75 1JFU Trp14 43–46 1WOU							
TlpA 72-75 1JFU Trp14 43-46 1WOU			1F9M				
Trp14 43-46 1WOU	thioredoxin 2		1THX				
	TlpA						
tryparedoxin I 40–43 1QK8							
	tryparedoxin I	40-43	1QK8				

^a Represents the highest resolution X-ray structure of this protein in the PDB.

These factors make firm conclusions as to the potential significance of these bond types not possible at present. They warrant further inspection, however, as the database of disulfide structures expands and new allosteric disulfides are identified.

We also looked for patterns of χ angles in the -RHStaple (allosteric) and +/-RHHook (catalytic) bonds. It was possible that the disulfides within these groups could be refined further into subsets with a particular geometry. The five χ angles and the strain values for each disulfide were normalized and then displayed graphically on a principle components analysis plot. Principle components analysis is a clustering technique used to group similar data values together based on values that occur more frequently than others. We found no obvious subsets of disulfides in the -RHStaple and +/-RHHook bonds or in any of the other 18 disulfide groups. This analysis implies that the orientation of a single dihedral angle is not enough to significantly alter the overall strain energy. Rather, it is the aggregated effect of all of the dihedral angles that determines the strain energy.

The -RHStaple group is populated mostly by disulfide bonds that link adjacent strands in the same antiparallel β sheet or what have been called cross-strand bonds (1, 2). These are unusual bonds because they occur in a secondary structure that is already bound together by noncovalent interactions (14). The distance between the α -carbon atoms of the Cys residues is often short because of their position in the antiparallel β sheet. The -RHStaple bonds have a mean $C\alpha-C\alpha'$ distance of 4.32 Å (95% CI, 4.26–4.37 Å), compared to a mean of 5.63 Å (95% CI, 5.61–5.65 Å) for all disulfides (Table 1). The strain associated with these bonds arises from the torsional energy of the linkage as well as the deformation energy in the sheet itself. The adjacent strands need to pucker to accommodate the disulfide bond. The bonds are often found straddling a β -hairpin structure.

The -RHStaple disulfides in the botulinum neurotoxins (2, 15), the *Escherichia coli* chaperone PapD (12), the oxidoreductase DsbD (13), and the immune receptor CD4 (16, 17) are redox-active and control protein function. Cleavage of one or more of the -RHStaple disulfides in the HIV envelope glycoprotein gp120 also appears to be an important aspect of the HIV infection process (2, 18, 19).

Gelsolin is an abundant cytosolic protein and is also secreted and found in plasma. It is an actin-binding and severing protein that is the primary activator of the gel—sol transition of cells (20). The -RHStaple disulfide is in the second domain, which is responsible for F-actin binding and has been shown to independently possess weak severing activity. The bond is oxidized in the plasma form of the protein and reduced in the cellular form and has been suggested to function in transmitting regulatory information from the fourth and sixth domains of the enzyme to the first (functional) domain (21).

Serum amyloid P is an acute-phase plasma protein and is found in all amyloid deposits (22). It binds to its partners in a calcium-dependent fashion, and its role in amyloid formation is to stop fibril degradation. It interacts with inflammatory and complement factors and possesses DNA and chromatin-binding capabilities. It has been proposed that its normal physiological role is the binding of DNA from apoptotic and necrotic cells to prevent an autoimmune response (23). There is no reference to the —RHStaple bond in this protein being important for its function. However, there is evidence for its susceptibility to cleavage in crystal structures. Of the five monomers in the crystal structure (PDB ID 1SAC), the bond is formed in only one of the monomers and is cleaved in the other four. In the crystal

structure of serum amyloid P dimer and a bivalent ligand of the cyclic pyruvate of glycerol (PDB ID 2A3Y), the bond is oxidized in all five monomers.

Tissue factor is a transmembrane glycoprotein that initiates coagulation. Binding of blood coagulation factor VIIa to tissue factor activates the serine proteinase zymogens factors IX and X by limited proteolysis, leading to the formation of thrombin and ultimately a fibrin meshwork that stabilizes the platelet thrombus (24). Tissue factor on the plasma membrane of cells is mostly in a cryptic configuration, which rapidly transforms into an active configuration in response to certain stimuli (25). It appears that the —RHStaple bond in tissue factor is reduced in the cryptic form of the protein and that activation involves the formation of the disulfide (unpublished observations).

The \pm /-RHHook disulfides are dominated by the catalytic residues of the oxidoreductases (Table 2). These disulfides cycle between reduced and oxidized configurations in coordination with a dithiol or disulfide in a protein substrate (3, 26). The outcome can be reduction, formation, or isomerization of a disulfide in the protein substrate depending upon the nature of the protein substrate, the redox conditions, and the presence of other thiols and disulfides. The reducing equivalents for the cycling come from cellular glutathione or NADPH. The bonds most often occur in a CysXXCys motif and have a redox potential in the range from \pm 120 to \pm 270 mV.

The high potential energy of +/-RHHook disulfides is likely a factor in determining the redox activity of these bonds. Another factor is the electrostatic environment of the disulfide. For instance, Asp26 serves as an acid/base in the oxidation/reduction reactions catalyzed by the Cys32-Cys35+/-RHHook of *E. coli* thioredoxin (27). Specifically, Asp26 protonates or deprotonates the thiol of Cys35 during substrate oxidation or reduction, respectively. A similar role is likely played by the analogous aspartate (or glutamate) residue in other oxidoreductases.

The observation that the approximately half dozen allosteric disulfides that have been identified are all —RHStaple bonds suggests that this configuration may be a hallmark of these disulfides. About 1 in 15 of the 6874 disulfides analyzed in this study is a —RHStaple bond compared to 1 in 40 for +/—RHHook bonds, which implies that allosteric disulfides may be more prevalent than the catalytic bonds. With the rapid advances in the structural determination of proteins, the new classification of the allosteric disulfide bond should aid in the functional description of proteins and how that function is regulated.

REFERENCES

- Hogg, P. J. (2003) Disulfide bonds as switches for protein function, Trends Biochem. Sci. 28, 210-214.
- Wouters, M. A., Lau, K. K., and Hogg, P. J. (2004) Cross-strand disulphides in cell entry proteins: Poised to act, *Bioessays* 26, 73-79.
- 3. Nakamura, H. (2005) Thioredoxin and its related molecules: Update 2005, *Antioxid. Redox Signal.* 7, 823–828.
- Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. J., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *J. Am. Chem. Soc.* 106, 765–784.
- Katz, B. A., and Kossiakoff, A. (1986) The crystallographically determined structures of atypical strained disulfides engineered into subtilisin, *J. Biol. Chem.* 261, 15480–15485.

- Wells, J. A., and Powers, D. B. (1986) In vivo formation and stability of engineered disulfide bonds in subtilisin, *J. Biol. Chem.* 261, 6564–6570.
- Kuwajima, K., Ikeguchi, M., Sugawara, T., Hiraoka, Y., and Sugai, S. (1990) Kinetics of disulfide bond reduction in α-lactalbumin by dithiothreitol and molecular basis of superreactivity of the Cys6-Cys120 disulfide bond, *Biochemistry* 29, 8240-8249.
- Wetzel, K., Perry, L. J., Baase, W. A., and Becktel, W. J. (1988) Disulfide bonds and thermal stability in T4 lysozyme, *Proc. Natl. Acad. Sci. U.S.A.* 85, 401–405.
- Pjura, P. E., Matsumura, M., Wozniak, J. A., and Matthews, B. W. (1990) Structure of a thermostable disulfide-bridge mutant of phage T4 lysozyme shows that an engineered cross-link in a flexible region does not increase the rigidity of the folded protein, *Biochemistry* 29, 2592–2598.
- Richardson, J., and Richardson, D. (1989) Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York.
- 11. Harrison, P. M., and Sternberg, M. J. (1996) The disulphide β -cross: From cystine geometry and clustering to classification of small disulphide-rich protein folds, *J. Mol. Biol.* 264, 603–623.
- Sauer, F. G., Pinkner, J. S., Waksman, G., and Hultgren, S. J. (2002) Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation, *Cell* 111, 543-551.
- 13. Haebel, P. W., Goldstone, D., Katzen, F., Beckwith, J., and Metcalf, P. (2002) The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: Crystal structure of the DsbC–DsbDα complex, *EMBO J. 21*, 4774–4784.
- 14. Wouters, M. A., and Curmi, P. M. (1995) An analysis of side chain interactions and pair correlations within antiparallel β -sheets: The differences between backbone hydrogen-bonded and non-hydrogen-bonded residue pairs, *Proteins* 22, 119–131.
- Erdal, E., Bartels, F., Binscheck, T., Erdmann, G., Frevert, J., Kistner, A., Weller, U., Wever, J., and Bigalke, H. (1995) Processing of tetanus and botulinum A neurotoxins in isolated chromaffin cells, *Naunyn-Schmiedeberg's Arch. Pharmacol.* 351, 67-78
- Matthias, L. J., Yam, P. T., Jiang, X. M., Vandegraaff, N., Li, P., Poumbourios, P., Donoghue, N., and Hogg, P. J. (2002) Disulfide exchange in domain 2 of CD4 is required for entry of HIV-1, *Nat. Immunol.* 3, 727–732.
- Maekawa, A., Schmidt, B., Fazekas de St. Groth, B., Sanejouand, Y.-H., and Hogg, P. J. (2006) Evidence for a domain-swapped

- CD4 dimer as the co-receptor for binding to class II major histocompatibility complex, *J. Immunol.*, in press.
- Gallina, A., Hanley, T. M., Mandel, R., Trahey, M., Broder, C. C., Viglianti, G. A., and Ryser, H. J. (2002) Inhibitors of protein—disulfide isomerase prevent cleavage of disulfide bonds in receptor-bound glycoprotein 120 and prevent HIV-1 entry, *J. Biol. Chem.* 277, 50579-50588.
- Barbouche, R., Miquelis, R., Jones, I. M., and Fenouillet, E. (2003) Protein—disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion, *J. Biol. Chem.* 278, 3131— 3136.
- Silacci, P., Mazzolai, L., Gauci, C., Stergiopulos, N., Yin, H. L., and Hayoz, D. (2004) Gelsolin superfamily proteins: Key regulators of cellular functions, *Cell Mol. Life Sci.* 61, 2614– 2623
- 21. Allen, P. G. (1997) Functional consequences of disulfide bond formation in gelsolin, *FEBS Lett.* 401, 89–94.
- Pepys, M. B., Dyck, R. F., de Beer, F. C., Skinner, M., and Cohen, A. S. (1979) Binding of serum amyloid P-component (SAP) by amyloid fibrils, *Clin. Exp. Immunol.* 38, 284–293.
- Bickerstaff, M. C., Botto, M., Hutchinson, W. L., Herbert, J., Tennent, G. A., Bybee, A., Mitchell, D. A., Cook, H. T., Butler, P. J., Walport, M. J., and Pepys, M. B. (1999) Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity, *Nat. Med.* 5, 694–697.
- Versteeg, H. H., and Ruf, W. (2006) Emerging insights in tissue factor-dependent signaling events, *Semin. Thromb. Hemostasis* 32, 24–32.
- Le, D. T., Rapaport, S. I., and Rao, L. V. (1992) Relations between factor VIIa binding and expression of factor VIIa/tissue factor catalytic activity on cell surfaces, *J. Biol. Chem.* 267, 15447– 15454.
- Holmgren, A. (1989) Thioredoxin and glutaredoxin systems, *J. Biol. Chem.* 264, 13963–13966.
- Chivers, P. T., and Raines, R. T. (1997) General acid/base catalysis in the active site of *Escherichia coli* thioredoxin, *Biochemistry* 36, 15810–15816.
- DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA.

BI0603064