Functional site profiling and electrostatic analysis of cysteines modifiable to cysteine sulfenic acid

FREDDIE R. SALSBUURY JR.,¹ STACY T. KNUTSON,^{1,2} LESLIE B. POOLE,³
AND JACQUELYN S. FETROW^{1,2}

¹Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27109, USA

²Department of Computer Science, Wake Forest University, Winston-Salem, North Carolina 27109, USA

³Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, USA

(RECEIVED July 1, 2007; FINAL REVISION October 30, 2007; ACCEPTED October 31, 2007)

Abstract

Cysteine sulfenic acid (Cys-SOH), a reversible modification, is a catalytic intermediate at enzyme active sites, a sensor for oxidative stress, a regulator of some transcription factors, and a redox-signaling intermediate. This post-translational modification is not random: specific features near the cysteine control its reactivity. To identify features responsible for the propensity of cysteines to be modified to sulfenic acid, a list of 47 proteins (containing 49 known Cys-SOH sites) was compiled. Modifiable cysteines are found in proteins from most structural classes and many functional classes, but have no propensity for any one type of protein secondary structure. To identify features affecting cysteine reactivity, these sites were analyzed using both functional site profiling and electrostatic analysis. Overall, the solvent exposure of modifiable cysteines is not different from the average cysteine. The combined sequence, structure, and electrostatic approaches reveal mechanistic determinants not obvious from overall sequence comparison, including: (1) pK_as of some modifiable cysteines are affected by backbone features only; (2) charged residues are underrepresented in the structure near modifiable sites; (3) threonine and other polar residues can exert a large influence on the cysteine pK_a; and (4) hydrogen bonding patterns are suggested to be important. This compilation of Cys-SOH modification sites and their features provides a quantitative assessment of previous observations and a basis for further analysis and prediction of these sites. Agreement with known experimental data indicates the utility of this combined approach for identifying mechanistic determinants at protein functional sites.

Keywords: functional site profile; redox signaling; cysteine sulfenic acid; cysteine reactivity; mechanistic determinants; post-translational modification; oxidative modification

Supplemental material: see www.proteinscience.org

Protein post-translational modifications are well known to play important biological roles by rapidly modifying the structure and function of proteins. The most common and

well-known example is the involvement of protein phosphorylation in signal transduction. Analysis of phosphorylation sites has led to a better understanding of kinase substrate specificity (Brinkworth et al. 2002; Kobe et al. 2005), methods for site prediction (Koenig and Grabe 2004; Huang et al. 2005; Plewczynski et al. 2005; Xue et al. 2005), and a combined experimental/computational approach that has led to a better understanding of the yeast phosphoproteome (Brinkworth et al. 2006; Molina et al. 2007).

The reversible oxidation of cysteine side chains to cysteine sulfenic acid (Cys-SOH) has been recognized as

Reprint requests to: Jacquelyn S. Fetrow, 100 Olin Physical Laboratory, 7507 Reynolda Station, Wake Forest University, Winston-Salem, NC 27019-7507, USA; e-mail: fetrowjs@wfu.edu; fax: (336) 758-6142.

Abbreviations: Prx, peroxiredoxin; Msr, methionine sulfoxide reductase; Cys-SOH, cysteine sulfenic acid; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; PTP, protein tyrosine phosphatase.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.073096508>.

a post-translational modification for several decades (Allison 1976; Poole and Claiborne 1989); however, the broader significance of its biological roles has been emerging only in the last decade (Claiborne et al. 1999). Cys-SOH plays roles at enzyme catalytic sites, senses oxidative and nitrosative stress, and regulates some transcriptional regulators (for review, see Poole et al. 2004). For example, cysteine is a catalytic residue in protein tyrosine phosphatases (PTPs): Cys-SOH modification at that active site residue is responsible for reversible inhibition of some PTPs (Denu and Tanner 1998; Cho et al. 2004; Tonks 2005). The transcriptional regulators OxyR and OhrR are also reversibly regulated through Cys-SOH formation (Zheng et al. 1998; Fuangthong and Helmann 2002; Kim et al. 2002; Panmanee et al. 2006). Cys-SOH can also be hyperoxidized to sulfinic and sulfonic acid (Cys-SO₂H and Cys-SO₃H, respectively), modifications that may play roles in signaling, aging, or disease processes (for review, see Berlett and Stadtman 1997; Jacob et al. 2004; Cross and Templeton 2006).

The biological relevance of Cys-SOH post-translational modification suggests that this modification is not random, but rather facilitated by specific features at or near the functional site that influence the modification and its stability. Previous observation of protein structure has suggested that the microenvironment that stabilizes Cys-SOH is characterized by three features: (1) lack of solvent accessibility to the modified cysteine; (2) lack of nearby reduced cysteines; and (3) local hydrogen-bonding residues that stabilize the sulfenate form (Claiborne et al. 1993, 1999). Until now, however, these observations have been general and qualitative. Given their significance, we aimed to better characterize known Cys-SOH sites in proteins.

Functional site profiling is a method that allows the analysis and comparison of sequence and structure features at a functional site, as outlined in Figure 1 (Cammer et al. 2003). We describe here the functional site profiling of Cys-SOH modification sites in proteins of known structure, the characterization of the sequence and structural features near the modifiable cysteine, and clustering of the sites by common sequence features.

Functional site profiling, with a sequence-based scoring function, has previously been applied to enzyme active sites (Cammer et al. 2003; Baxter et al. 2004; Huff et al. 2005). Such sites are more definitively related than are Cys-SOH modification sites; consequently, identification of features other than just sequence and structure is likely necessary to identify relationships between these sites. The accepted mechanism for Cys-SOH generation by hydrogen peroxide-mediated oxidation involves initial cysteine deprotonation (Denu and Tanner 1998; Wood et al. 2003b), thus suggesting a lowered sulfhydryl pK_a

Figure 1. An example of construction of functional site signatures and a profile. (A) The three-dimensional structure of 1vhrA, a dual-specificity phosphatase, is shown as a ribbon. The location of the modifiable (and, in this protein, the active site) cysteine residue is indicated as a light-gray sphere. The structural segments that contain residues located within 10 Å of the key cysteine residue are shown in colors, each color representing a different segment. (B) The segments are extracted from the global structure. (C) The sequence fragments corresponding to the structural segments are combined, from N to C terminus, into a single contiguous sequence that is called the functional site signature. The color of each residue is indicative of the segment in which it was originally located in the structure, as shown in A and B. (D) Related signatures can be aligned to create a functional site profile, as shown for three example proteins. The modified cysteine is underlined. A scoring function for a given profile rewards both identities and similarities, but penalizes gaps. Functional site profile scores greater than 0.25 indicate a significant relationship between the signatures in the profile for a closely related family (Cammer et al. 2003); however, some proteins with scores slightly less than 0.25 are known to be related (Baxter et al. 2004). Protein structures were prepared in VMD v1.8.3 (Humphrey et al. 1996).

for the modifiable cysteine. This observation indicates that analysis of the cysteine pK_a and its electrostatic environment should provide important clues to its modifiability, an idea similar to that suggested for enzyme active sites (Ondrechen et al. 2001; Jones et al. 2003). Thus, in this work, we apply the commonly used semi-macroscopic electrostatic methods or simple continuum models (Bashford and Gerwert 1992; Antosiewicz et al. 1994; Sham et al. 1997) to the modifiable cysteine sites in proteins and compare the electrostatic features with the sequence features identified by profiling.

Our long-term goals are to develop methods to predict Cys-SOH modification sites in protein sequences and to create a database of modification sites, and potential sites, for use by biological researchers. Toward this long-term goal, we have compiled a list of proteins known to contain Cys-SOH modification sites and report that list here. We characterize site features using a combination of functional site profiling and electrostatic analysis, which allows us to more quantitatively cluster Cys-SOH sites, compare features between the sites, and identify similarities and differences related to cysteine reactivity.

Results

Overview of structure and function of proteins that contain Cys-SOH sites

The modifiable protein set contains 47 proteins with 49 distinct Cys-SOH sites (Table 1). These proteins either exhibit Cys-SOH in the crystal structure (27 proteins) or are known from biochemical studies to generate a Cys-SOH during exposure to oxidants (20 proteins). Each site, therefore, is located in an environment in which the cysteine can be modified by some mechanism to sulfenic acid. A total of 27 of the 47 proteins contains a Cys-SOH that is sufficiently stable to allow observation in the crystal structure. Most of the proteins in the modifiable data set exhibit little sequence identity to one another. Sequence identities range from insignificant (<10% identity) to 99% (1kyg and 1n8j are the same protein sequence, with a single mutation difference). Most proteins are unrelated, with ~99% (1116 of 1128) of all global pairwise sequence identities <30%.

As expected, cysteine oxidation to sulfenic acid is not limited to proteins of specific biochemical function or structure. The biochemical functions include metabolic enzymes (e.g., GAPDH, malate synthase), DNA-binding and transcriptional regulatory proteins (e.g., papilloma-virus E2 protein, NF- κ B p50/p65), inhibitors (e.g., α -1 antitrypsin), oxygen carriers (e.g., hemoglobin), and proteins involved in the immune system (e.g., H-2 Class I histocompatibility antigen). A few are involved in cellular redox regulation and/or signaling (e.g., peroxiredoxin, glutathione reductase, thioredoxin), but most are not. Comparison of SCOP classifications (Murzin et al. 1995) for these proteins indicates that they are not limited to a single structural class: All structural classes and some multidomain proteins are represented. Additional analyses (see Supplemental material and Supplemental Fig. S1) indicate that modifiable cysteines have no propensity for any particular secondary structure; thus, while the helix macrodipole might contribute to some modification sites, it is not the only mechanism by which the cysteine reactivity is modified.

Modifiable cysteines might be more reactive due to greater solvent exposure; however, to stabilize the reactive sulfenic acid intermediate, one might expect the cysteine to be more buried. To determine whether modifiable cysteines were more or less accessible than average, solvent accessibility was calculated for some of the side chains. Conformational change occurs in some of these proteins when the cysteine is modified, thus we did not calculate the solvent accessibility for cysteines that were actually modified in the crystal structure. The average accessibility for 15 modifiable (but not modified) cysteines was $94.7 \text{ \AA}^2$, while the side-chain surface area for the average cysteine

was $90.7 \text{ \AA}^2$ (with average defined as all free cysteines in the modifiable protein set). The $4 \text{ \AA}^2$ difference is too small to be physically relevant in determining reactivity, though the calculation should be repeated when more structures of modifiable (but not modified) sites are known.

Features of the functional site signatures and profiles for modifiable cysteine sites

The modifiable cysteine was selected as the “key residue,” and signatures were identified for each cysteine modification site (procedure shown in Fig. 1) (Cammer et al. 2003). These signatures were aligned (see Materials and Methods) to create a functional site profile for these sites (Fig. 2). The signatures are highly diverse, a characteristic different from profiles previously created for enzyme active sites (Cammer et al. 2003; Baxter et al. 2004). In fact, the sequence-based profile score is -0.46 , an insignificant score (Cammer et al. 2003). This result is expected given the diversity of proteins (Table 1) and local structures (Supplemental Fig. S1) in which this post-translational modification occurs.

Frequencies for residues within the signature containing the modifiable cysteine were calculated (Table 2). One observation is immediately surprising: Three of the four charged residues, Asp, Glu, and Lys, are found in active site signatures significantly less often than would be expected by chance. The frequency of occurrence of Arg in the signatures is about the same as its occurrence in the entire modifiable protein set. If charged residues were a key feature in determining cysteine reactivity, one would expect these residues to be overrepresented. Residues that are overrepresented in the signatures include the hydrogen-bonding residues—particularly Thr, Ser, and His—suggesting that hydrogen bonding might play an important role in cysteine reactivity.

Frequencies for specific residues directly adjacent to the modifiable cysteine in the sequence were also calculated (Table 2). Given the caveat of small numbers, three residues are overrepresented in this analysis with frequencies of adjacent residues greater than 2.7 in each case: His and Met N-terminal and Trp C-terminal to the modifiable cysteine. His was observed as an N-terminal adjacent residue seven times in mostly unrelated signatures (Fig. 3) and, thus, its overrepresentation in this position is statistically significant, suggesting His-Cys might be a predictive motif. Adjacent Met and Trp residues were only observed two times each, an insufficient number of observations to draw conclusions. Seven residues were never observed in adjacent positions: Glu, Arg, Trp, Phe, Pro, Cys, and Ile N-terminal and Cys and Asn C-terminal to the modifiable cysteine (Table 2). Because of the small numbers, drawing conclusions may

Figure 2. All signatures for each Cys-modifiable site, organized and aligned to show the functional site profile. For each protein in the modifiable protein set (Table 1), the functional site signature was extracted using the process shown in Figure 1. The signatures were aligned using ClustalW (Thompson et al. 1994; Higgins et al. 1996), and then the functional site profile score was calculated for each pair of signatures. The signatures are organized by these pairwise scores, with the dendrogram to the *right* indicating the relationships based on these pairwise scores (note that the branch length in the dendrogram is meaningless and does not indicate a distance). The modifiable cysteine is shown as a white “C” on a black background and the structural fragment that contains the modifiable cysteine is highlighted in light blue. Yellow and red shading indicates strongly and weakly interacting residues, respectively. (As described in the Materials and Methods, strong interactors, yellow, are identified when the interaction energy [measured in pK units] is >1.0 pK units, and weak interactors, red, when the interaction energy is between 0.5 and 1.0 pK units.) The identity of the strong interactors is given in Table 1. Clusters of signatures discussed in the text are indicated by gray shading; these are also shaded in Table 1.

be premature, but in one case the difference in occurrence between the directly N- and C-terminal residues is statistically significantly different: Pro occurs nine times on the C-terminal side, but never on the N-terminal side. These observations, and the data provided in Table 2, provide initial clues into potentially important features for the prediction of modifiable cysteines.

Despite the profile diversity, groups of signatures within the profile appear to be related. If those similarities are significant, they would indicate common features necessary for cysteine modification within these clusters. A first step in identifying such common features is to

explore the relationship between the overall protein sequence and the modifiable cysteine site. A second step is to explore other methods, such as inclusion of biophysical parameters, which would be useful in identifying similarities beyond simple sequence comparison of the signatures. The first step will be discussed in the next paragraphs, and the second will be explored by inclusion of electrostatics in the profile analysis, as described in subsequent sections.

To explore the relationship between the overall sequences and the signatures, we compared the pairwise sequence identity between signatures with the pairwise

Table 1. Protein structures from the *pdh* with cysteines modifiable to sulfenic acid^a and strongly interacting residues

PDB ^b and chain ^c	Reference ^d	Protein name	Cys-SOH site ^e	Cys pK _a	Cys pK _a (-Thr, if Thr has an effect)	Strongly coupled residues (>1 pK Unit)
1QXH_A	(Baker and Poole 2003; Choi et al. 2003)	Thiol peroxidase	61	7.5		C95
2BOP_A	(Hegde et al. 1992; McBride et al. 1992)	Papillomavirus-2 E2 DNA-binding domain	340	6.0		
1SK4_A	(Guan et al. 2004)	Peptidoglycan recognition protein I-α	300x	7.6		
1Q79_A	(Martin et al. 2004)	Polynucleotide adenyltransferase	160x	8.3		
1JZ7_A	(Iuvers et al. 2001)	β-galactosidase	247x			
1G55A (1) ^f	(Dong et al. 2001)	DNA cytosine methyltransferase DNMT2	140x	8.5		
1VKX_B (p50)	(Chen et al. 1998a; Pineda-Molina et al. 2001)	NF-κB p50/p65 heterodimer transcription factor	359	6.7		
1KYG_A ^g	(Poole and Ellis 2002; Wood et al. 2002)	Alkyl hydroperoxide reductase AhpC	46	7.0	7.3	
1QQ2_A	(Hirotzu et al. 1999)	2-Cys Peroxiredoxin, Hbp23 (Prx I)	52	7.2		
1PKZ_A ^d		Glutathione-S- transferase A1	112x	7.45		K120
1A02_J	(Abate et al. 1990; Chen et al. 1998b)	Complex of DNA binding domains of Nfat, Fos and Jun bound to DNA	279	4.5		
1A02_F	(Abate et al. 1990; Chen et al. 1998b)	Complex of DNA binding domains of Nfat, Fos and Jun bound to DNA	154	5.5		
1FZJ_A	(Chen et al. 1998b)	H-2 Class I histocompatibility antigen, K-B	121x	8.2		
1VHQ_A	(Badger et al. 2005)	Enhancing lycopene biosynthesis protein 2	138x	9.7		
1QVZ_A	(Graillie et al. 2004)	YDR533C protein (function unknown)	138x	8.9	9.6	E30 H139
1D8C_A	(Howard et al. 2000; Anstrom et al. 2003)	Malate synthase G	617	8.5	9.0	S274 D631
2AHJ_A	(Nagashima et al. 1998)	Nitrile hydratase (lyase)	114x	7.9		C109 R167 T162
1IRE_A	(Miyanaga et al. 2001)	Nitrile hydratase	113x			
1VHR_A	(Denu and Tanner 1998)	Human Vh1-related dual specificity phosphatase	124	4.3		S131 R130 S129
1UB7_A ^d		3-Oxoacyl-[acyl-carrier protein] synthase	110x	5.2		H246 S278 S305 E138
1OET_A	(van Montfort et al. 2003)	PTP1B	215x	0.8		S222 R221
1I9T_A	(Changela et al. 2001)	RNA triphosphatase domain of mRNA capping enzyme	126x	0.2	0.5	T133
1N8J_A	(Poole and Ellis 2002; Wood et al. 2003a)	Alkyl hydroperoxide reductase AhpC	46	5.9	7.0	T43 Y38 R119 E49
1QMV_A ^b	(Schroder et al. 2000)	Thioredoxin peroxidase B (Prx II)	51x ^h	7.0	7.5	T48 Y43 R127
1E2Y_A	(Montemartini et al. 1999; Alphey et al. 2000)	Tryparedoxin peroxidase	52	7.5	8.6	R128
1J0X_O	(Cowan-Jacob et al. 2003)	Muscle glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	149x	4.2	4.9	Y311 H176 T150 S238
1HD2_A	(Declercq et al. 2001)	Peroxiredoxin V	47	5.9	6.6	R127 T44
1PRX_A	(Choi et al. 1998)	hOrf6 peroxidase (Prx VI)	47x ^e	4.4	6.2	H39 E50 T48 R132 S72 T44
1PRX_B	(Choi et al. 1998)	hOrf6 peroxidase (Prx VI)	47x ^e	6.4	7.5	H39 E50 T48 R132 S72
9PAP_	(Kamphuis et al. 1984)	Papain	25	6.3		H159 S176
1MEM_A	(McGrath et al. 1997; Percival et al. 1999)	Cathepsin K	25	7.2		H159 S176
1FVA_A	(Boschi-Muller et al. 2000; Lowther et al. 2000)	Methionine sulfoxide reductase (msr)	72	10.4		D150 Y217 E115
1O4C_A	(Lange et al. 2003)	Proto-oncogene tyrosine-protein kinase SRC	44x	4.7		R34 S46 S36
1GSN_	(Becker et al. 1998)	Glutathione reductase	63x	6.0	6.4	T339 C58
1EKF_A	(Yennawar et al. 2001; Conway et al. 2003; Conway et al. 2004)	Branched chain amino acid aminotransferase	315	6.8	7.0	C318
1L1D_A	(Lowther et al. 2002; Olry et al. 2002)	C-terminal methionine sulfoxide reductase domain (msrB)	495	7.6		C440

(continued)

Table 1. Continued

PDB ^b and chain ^c	Reference ^d	Protein name	Cys-SOH site ^e	Cys pK _a	Cys pK _a (-Thr, if Thr has an effect)	Strongly coupled residues (> 1 pK Unit)
1MJB_A	(Yan et al. 2002)	ESA1 protein (acetyltransferase domain)	304x	5.5		K261 S255 S291 Y289
1ATU_	(Ryu et al. 1996; Griffiths et al. 2002)	α-1-Antitrypsin	232	6.6		
1RQ4_B	(Chan et al. 2004)	Hemoglobin β chain	93x			
1I69_A	(Zheng et al. 1998; Choi et al. 2001)	Reduced form of OxyR	199	7.9		T129
1LG7_A	(Gaudier et al. 2002)	VSV matrix protein	135x	7.35		
1QWL_A	(Lesniak et al. 2003; Rehse et al. 2004)	OsmC, hydrogen peroxide reductase	59	5.4	5.8	C125 R39* E48* H56
1N2F_A	(Lesniak et al. 2002; Oliveira et al. 2006)	OhrR transcriptional repressor	60	7.2		C124
1EQ2_A	(Deacon et al. 2000)	ADP-L-glycero-D-mannoheptose 6-epimerase	78x	6.4		Y88
1G55_A(2) ^f	(Dong et al. 2001)	DNA cytosine methyltransferase DNMT2	287x	7.35		
1FNJ_A	(Kast et al. 2000)	Chorismate mutase	75x	6.9		
1N5U_A	(Wardell et al. 2002; Carballal et al. 2003)	Serum albumin complexed with heme	34	7.7		
1HKU_A	(Nardini et al. 2003)	C-terminal binding protein 3 (CTBP3)	27x	9.0		E30
1JOA_	(Yeh et al. 1996)	NADH peroxidase	42x	6.7		
1CXP_C	(Fiedler et al. 2000)	Myeloperoxidase isoform C	150x	5.7		

^aProteins are listed in the same order in which they are identified by clustering their functional site signatures. Gray-shaded groups indicate clusters identified in Figure 2 and discussed in the text.

^bPDB structure code, followed by the chain letter, if multiple chains are found in the file.

^cThe letter indicates the representative chain for all identical chains that contain the modifiable or reactive Cys.

^dReferences that describe the crystal structure (if one was published) are provided; additional references provide experimental evidence for the existence of the Cys-SOH site. No reference indicates that there was no literature reference for either the crystal structure or experimental evidence of Cys-SOH, but the moiety was observed in the crystal structure.

^eThe number indicates the residue number of the Cys that is modified to sulfenic acid; “x” indicates that the Cys-SOH (or SO₂H) modification is stable enough to be observed in the crystal structure at the site.

^fThe protein 1G55 is modified at two residues: 140 and 287, indicated as 1G55(1) and 1G55(2), respectively.

^gThe 1kyg file has been replaced by 1yep (Parsonage et al. 2005).

^hThe 1qmv protein contained a sulfenic acid, not sulfenic acid, modification in the crystal structure.

(*) Residues on the other monomer in the dimer.

Table 2. Frequencies of residue occurrence in signatures and adjacent to the modifiable Cys

Residue	Frequency in signatures	N-term residue frequency ^a	C-term residue frequency ^a
Cys	4.12	— ^b	— ^b
Trp	1.38	0	3.3
Thr	1.30	0.71	1.32 ^a
Gly	1.28	1.12 ^a	0.60
His	1.28	2.78 ^a	1.06
Ser	1.21	1.24 ^a	1.27
Tyr	1.18	1.27	1.59
Phe	1.18	0	1.16
Met	1.17	3.18	1.06
Pro	1.13	0	1.43 ^a
Ala	1.09	1.22 ^a	1.27
Val	1.08	1.38 ^a	0.40
Arg	0.95	0	1.16
Leu	0.76	0.23	0.58
Gln	0.72	1.27	0.40
Ile	0.68	0	0.79
Asn	0.63	0.35	0
Asp	0.59	0.64	0.64
Glu	0.59	0	0.91
Lys	0.49	1.19	1.06

^aFrequency of occurrence of residues directly N- or C-terminal to the modifiable cysteine; ^anext to the number indicates that there were five or more occurrences in the signature data set. There are 13,377 residues in the modifiable set of proteins, 962 residues in the functional site signatures, and 162 nondisulfide bonded and nonmodifiable cysteines in the 47 proteins.

^bNo examples of CC were observed in the entire modifiable data set of proteins.

sequence identity between complete sequences. Exploration of this relationship will indicate whether the fragments proximal to the modifiable cysteines are more or less similar than the overall sequences. This comparison shows that the signature sequence identity is largely uncorrelated with the overall sequence identity (data not shown). There are three possible explanations for lack of correlation between global sequence identity and identity between signatures in the same protein: (1) the signatures are trivially short, so sequence comparison between signatures is not significant; (2) structural polymorphism (differences in conformation) at the cysteine modification sites create different signatures; and (3) the actual mechanism for cysteine modification is more or less similar than the overall identity between the two proteins would suggest. Prior knowledge of protein structure and function allows us to identify examples of each explanation, as discussed in the next three paragraphs.

The small size of some signatures results in a lack of information with which to create an accurate alignment; consequently, two signatures might appear more or less related than they actually are. An example includes signatures [1kyg, 1vkx] and [9pap, 1mem]: The 1kyg signature is 36% identical to signatures for 9pap and

1mem, while the 1vkx signature is 30% identical to the 9pap and 1mem signatures. The 1kyg and 1vkx signatures are only 10 and 11 residues in length, respectively (Fig. 2). Sequence identity of 30%–36% between such short fragments is not significant. Conclusions about mechanistic similarity for cysteine modification at such sites would be unwarranted.

Signatures are identified by structural criteria (Fig. 1); consequently, conformational differences (or structural polymorphism) at the modifiable cysteine site can make the signatures appear less related than they might actually be. Several examples of this effect are observed in the modifiable protein set. The most obvious example is 1n8j and 1kyg, both the bacterial peroxiredoxin, AhpC. These structures differ by a single residue change at the modifiable cysteine, and thus exhibit 99% global sequence identity; however, the two proteins exhibit different conformations near the modifiable cysteine site (red arrow and red/blue side chains, Fig. 3A). The structural polymorphism exists because the structures represent analogs of different conformations seen during the enzymatic reaction: 1kyg contains a disulfide bond between the modifiable cysteine and another mechanistically important cysteine in another subunit, while 1n8j is a mutant with a serine replacing the modifiable cysteine and, thus, no disulfide bond (Wood et al. 2003a). The structural polymorphism produces different functional site signatures with 81% pairwise sequence identity (Fig. 3B). Even though one of the “key” residues is an engineered serine rather than a cysteine and these residues are not aligned in superposition of the structures (Fig. 3A, cf. red and blue side chains), these residues are properly aligned in the profile (Fig. 3B). In addition, both signatures contain a threonine that is important for cysteine reactivity (discussed subsequently) and this threonine is aligned in the profile (Fig. 3B). This observation indicates that a profile can identify and align mechanistically important residues, even in cases where structural changes are observed; however, if conformational differences are observed in identical proteins, the resulting signatures will appear less similar than they actually are.

Similarity in the cysteine deprotonation mechanism, or lack thereof, is the third and most interesting explanation for observing or not observing a correlation between the sequence identities of full protein and signature. In this case, two signatures would exhibit more (or less) similarity than the overall sequence comparison would suggest, indicating that mechanisms by which the reactive cysteine is modified to Cys-SOH are similar (or are not similar). Such observations would potentially allow identification of common features of the modification mechanism that would not be obvious from the overall sequence comparison. The proteins 1hd2 and 1lpx provide an example. Overall sequence identity between these two proteins is 8%, an insignificant value that

Figure 3. An example of structural polymorphism, one cause for differences between functional site signatures of otherwise related proteins. (A) Proteins 1n8jA (cyan ribbon) and 1kygA (yellow ribbon) are both alkyl hydroperoxide reductase AhpC proteins. These two proteins are identical in sequence and largely similar in structure, with the exception of the protein C terminus and the region around the Cys-SOH site (areas indicated by arrows). The modifiable cysteine residue in 1n8jA is shown as blue ball-and-stick (in this case the Cys46 has been mutated to a Ser), while in 1kygA is shown in red (in this case, disulfide bonded to the resolving cysteine, though that subunit is not shown in this figure). The weakly interacting Thr residue common to both signatures is shown in purple (1kygA) and pink (1n8jA). (B) Alignment of the functional site signatures of 1n8j (top) and 1kyg (bottom) shows the different sequence fragments identified as the signatures for these proteins. Although the proteins are mostly identical, the signatures are different due to the structural polymorphisms shown in A. Strong interacters are highlighted in yellow and weak interacters are highlighted in red. The modified cysteine is shown as a red character. Protein structures were prepared in VMD v1.8.3 (Humphrey et al. 1996).

would suggest no relationship; however, the sequence identity and profile score between the signatures are both significant at 41% and 0.38, respectively. The sequence similarity at the modifiable cysteine site is apparent from comparison of the signatures (Fig. 2, light-gray shading) and structures (Fig. 4). This observation makes sense: These proteins are peroxiredoxins (Prxs) V and VI, which exhibit a common mechanism of cysteine modification (Poole 2007). (Their electrostatic similarity extends and supports this observation and is discussed subsequently.) While this result is not new, the ability to identify previously known mechanistic similarities by comparison of functional site signatures demonstrates the method's utility and its potential for identifying novel similarities.

Functional site profiling focuses on sequence similarities. Identification of mechanistic similarities at diverse sites, such as the Cys-SOH modification sites, requires additional information. As part of the mechanism of Cys-SOH formation through reaction with substrates such as hydrogen peroxide, the cysteine is likely to be deprotonated (Wood et al. 2003b); thus, decreased cysteine pK_a facilitates the reaction. The pK_a , and residues that cause

its shift, are thus ideal biophysical features with which to further characterize these sites. As a first approach to including additional features in profiles, we explore the pK_a of the modifiable cysteine and the residues that cause its pK_a to be shifted using newly validated cysteine parameters and methods (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.).

Electrostatic characterization of Cys-SOH modification sites

A pK_a was calculated for the reduced cysteine at each Cys-SOH modification site in each structure (Table 1). As expected, the calculated pK_a distribution for modifiable cysteines is shifted compared with the distribution for a control set from 8.14 for the control set to 6.9 for the modifiable set (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.). This result is consistent with the molecular function of these sites, in which the first step in the typical oxidation mechanism is cysteine deprotonation (Wood et al. 2003b).

All but eight modifiable cysteines exhibit pK_a s lower than the mean of a control protein set (8.14) and nearly half, 20 cysteines, exhibit pK_a s shifted more than 1.5 below the mean. The eight with increased pK_a s ranging from 8.2 to 10.4 are: 1fzj (H-2 class-1 histocompatibility antigen K-B), 1q79 (polynucleotide adenylyltransferase), 1g55 (DNA cytosine methyl transferase), 1qvz (gene product of YDR533C, function unknown), 1hku (C-terminal binding protein 3), 1d8c (malate synthase G), 1vhq (enhancing lycopene biosynthesis protein 2), and 1fva (methionine sulfoxide reductase, Msr) (Table 1). Only the last two exhibit pK_a s greater than one standard deviation from the mean for a control cysteine data set. There are two potential explanations for increased cysteine pK_a s. First, protein conformational changes and local flexible loops modulate the pK_a of the modifiable cysteine, an aspect not captured using static structures. Second, Cys-SOH generation could be a result of modification in the synchrotron and, thus, would likely occur by a mechanism (e.g., hydroxyl radical attack) (Xu and Chance 2005) different from a typical biological oxidation mechanism (e.g., oxidation by hydrogen peroxide) (Wood et al. 2003b). Seven of these eight protein structures were solved at synchrotrons, increasing the possibility of oxidative modification during data collection. Flexibility is likely the explanation for the eighth protein, Msr, as a nearby loop is found in multiple conformations in different crystal structures.

Again, we can use prior knowledge of protein structure and function to illustrate the utility of electrostatic analysis to analyze modifiable cysteine microenvironments. The modifiable cysteines in the protein tyrosine phosphatase 1B (1oet) and RNA triphosphatase domain of

Figure 4. Comparison of the functional sites of 1prx, 1hd2, and 1j0x shows similarities in the structural location of key functional residues, despite the lack of overall sequence similarity. The backbone structure for 1prx, Prx VI (A), 1hd2, Prx V (B), and 1j0x, GAPDH (C) are shown as white ribbons; segments that comprise the functional site signature are colored cyan. (The signatures themselves are aligned in Fig. 2, with 1prx and 1hd2 shaded in light gray and 1j0x shaded in dark gray.) Consistent with the coloring in Figure 2, side chains of strongly interacting residues are colored yellow, and weakly interacting residues are colored red. The modifiable Cys at these functional sites are shown as blue van der Waals side chains (Cys 47, Cys 47, and Cys 149, respectively). Strongly interacting residues Arg 132 and Thr 44 (1prx) and Arg 127 and Thr 44 (1hd2) are shown as yellow side chains; these align nicely in both the signature and the structure of the two proteins. Although Tyr 317 (red side chain) of 1j0x does not align with Arg 132 and Arg 127 in the signatures, it does share a similar position in structure. The strong interactor Glu 50 (1prx, yellow) and aligned weak interactors His 51 (1hd2, red) and Cys 153 (1j0x, red) are also located in structurally similar positions. The strong interactors Ser 72 (1prx, yellow) and His 176 (1j0x, yellow), and weak interactor Cys 72 (1hd2, red) align within the sequence signature (Fig. 2), are not in exactly the same position, but are all found in a β strand adjacent to the active site. Overall, the proteins exhibit low sequence identity (1prx and 1hd2: 8%; 1prx and 1j0x: 7%; and 1hd2 and 1j0x: 11%), but the structural similarity between the proteins, particularly at the functional site identified in the signature comparison in Figure 2, can be observed in the structures.

mRNA capping enzyme (1i9t) exhibit calculated pK_a s of the modifiable cysteines that are less than one (Table 1). While this pK_a is nonbiological (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.), the result indicates the extremely low probability of cysteine protonation in a physiological environment, consistent with the mechanism of Cys-SOH formation. The intrinsic pK_a (not including titratable residues) of both phosphatase cysteine microenvironments is low, 6.1 and 5.1 for 1oet and 1i9t, respectively. The intrinsic pK_a is a result of significant electrostatic interactions with nontitrating groups, including dipoles generated by the partial charges of the backbone atoms in an adjacent loop. Interactions with titrating residues lower the pK_a even more: 5.3 and 4.9 pH units for 1oet and 1i9t, respectively (Table 1). We identify Arg 221 and Ser 222 as significant interacting residues in 1oet and Thr 133 in 1i9t (Table 1). The identification of Arg 221 is in agreement with previous work, which compared three phosphatases and determined that the only ionizable side chain significantly influencing the cysteine pK_a was this conserved arginine (Peters et al. 1998); however, the previous work did not determine the interaction of Ser or Thr with the active site Cys by treating them as titratable in the calculations. This is a second indication of the importance of serine and threonine and is consistent with the residue frequencies (Table 2).

Qualitatively, the electrostatic calculations on the modifiable cysteine sites are consistent with the expected: an overall decrease in the modifiable cysteine pK_a . While the calcu-

lations cannot exactly pinpoint the cysteine pK_a , they can help us to identify residues important to cysteine reactivity.

Effect of threonine residues on the pK_a of Cys-SOH modification sites

The calculations were first performed without considering Thr as an ionizable residue, a standard assumption given its high pK_a (given in standard tables as 15.0). Residue frequency calculations indicate that Thr is over-represented in the signatures compared with its overall presence in these proteins (Table 2), suggesting the importance of hydrogen bonding. To quantify the interaction and effect of Thr on the modifiable cysteine, electrostatics were recalculated including Thr as an ionizable residue.

The inclusion of Thr significantly lowers the calculated pK_a s for the modifiable cysteine in three proteins and lowers it somewhat for 10 more (Table 1). Altogether, the Thr is important for maintaining the modifiable cysteine pK_a in six Prx crystal structures. This observation suggests that this Thr, which is essentially invariant across all classes of Prxs (Hofmann et al. 2002), is of considerable importance in determining the electrostatic properties of these functional sites. In fact, mutations at this position indicate the importance of the Thr hydroxyl group, as replacement by serine or valine yields an active or inactive enzyme, respectively, in studies of a *Leishmania donovani* Prx (Flohe et al. 2002). Again, this post hoc

“prediction” supported by experimental evidence indicates the utility of the method for analysis of functional site features.

The pK_a s of seven other proteins are also affected by inclusion of Thr: 1qvz (product of gene YDR533C, an unknown protein from *Saccharomyces cerevisiae*), 1j0x (glyceraldehyde-3-phosphate dehydrogenase), 1d8c (malate synthase G), 1qwi (OsmC hydrogen peroxide reductase), 1gsn (glutathione reductase), 1ekf (branched chain amino transferase), and 1i9t (RNA triphosphatase). There is no obvious commonality between these proteins, except that all contain sites of Cys-SOH modification. Thr, its hydroxyl group, and hydrogen-bonding capabilities are thus predicted to play an important role in these protein functions and cysteine deprotonation.

Comparison of functional site profiling and electrostatic analysis

To accomplish our long-term goals of developing methods to predict Cys-SOH modification sites, we must more fully identify the sequence, structure, and electrostatic features that are required for propensity toward cysteine modification. Toward this end, we combined functional site profiling, which is sequence based, with information about electrostatics. First, the complete functional site profile (Fig. 2) was clustered based on pairwise active site profile scores (Cammer et al. 2003). Second, as part of the electrostatic calculations, residues that affect the pK_a of the modifiable cysteine were also identified (see Materials and Methods). Residues were divided into two classes: strong interacters, where the interaction energy is >1.0 pK units; and weak interacters, where the interaction energy is between 0.5 and 1.0 pK units. Strong interacters are listed in Table 1 and colored yellow in Figure 2. Weak interacters are indicated in red in the functional site profile (Fig. 2). We then compared sequence-based similarities with similarities in the location of interacting residues.

Of the 210 interacting residues identified by the electrostatic calculations, 169 are found in the functional site signatures. Thus, as expected, most (80%) of the residues that shift the modifiable cysteine pK_a are located within 10 Å of that cysteine. However, the converse is not true—not all titratable residues within 10 Å of the cysteine influence its pK_a . A total of 166 titratable residues are found in the functional site signatures that are not interacting residues, compared with 169 titratable residues that are; thus, only about half of all titratable residues within 10 Å of the modifiable cysteine are important for its pK_a shift. Proximity of side chain centers of mass of these residues is therefore not the only parameter affecting the cysteine pK_a . We observed a number of causes for this noninteraction, including hydrogen-bonding residues pointing away from the mod-

ifiable cysteine and other protein atoms located between the titratable residue and the modified cysteine.

Comparison of the interacting residues with the functional site signatures profile reveals several commonalities. Recall that Prxs V and VI (1hd2 and 1prx, respectively) exhibit an overall sequence identity of only 8%, but sequence identity and significant profile scores between the functional site signatures of 41% and 0.38, respectively, indicating that the cysteine modification sites are similar. The electrostatic analysis identifies four aligned interacting residues: Thr; Arg; His/Glu; and Cys/Thr (Table 1 and Fig. 2, light-gray shading). The mechanism for shifting pK_a thus appears to be shared among these proteins. The similarity in the structure of their functional sites is apparent (Fig. 4). The identification of these potential common mechanistic determinants was revealed by comparison of the functional site signatures and the pK_a analysis, and not by overall sequence analysis (which is only 8%). The result is consistent with known mechanisms of these Prxs (Choi et al. 1998; Declercq et al. 2001).

Other Prxs provide further examples, including 1n8j (AhpC), 1qmv (thioredoxin peroxidase B, also known as PrxII), and 1e2y (tryparedoxin peroxidase) (Fig. 2, light-gray shading). Positions of two of the four interacting residues are conserved in all of these proteins: the Thr located three residues to the N terminus of the modifiable cysteine, and the His/Glu located several residues toward the C terminus from the modifiable cysteine. Other interacting residues are common only to a subset of Prxs, e.g., 1n8j (AhpC) and 1qmv (Prx II) contain six conserved interacting residue positions (Fig. 2), suggesting a common mechanism for cysteine deprotonation in these two Prxs.

Interestingly, GAPDH (1j0x), which is not a Prx, is found amidst this Prx cluster and exhibits several common interacting residue positions (cf. functional site signatures, shaded dark and medium gray, Fig. 2). The most obvious similarities exist within the helix containing the modifiable cysteine (Fig. 4). All three proteins have an interacting residue approximately one helical turn from the cysteine on the C-terminal side (Glu 50 in 1prx, His 51 in 1hd2, and Cys 153 in 1j0x). Also, 1prx Thr 48 shares a similar position in the structure to 1j0x Thr 150. Two other notable structurally similar interacting residues are observed (Fig. 4). First, all three proteins contain an interacting residue on a β strand directly behind the cysteine: Ser 72 of 1prx, Cys 72 of 1hd2, and His 176 of 1j0x. Second, residue Tyr 317 in 1j0x is located near the cysteine in a similar manner as Arg 132 in 1prx and Arg 127 in 1hd2, a residue thought to be important in stabilizing the deprotonated cysteine in these Prxs (Wood et al. 2003b; Copley et al. 2004). These observations suggest that these common features might play similar roles in cysteine deprotonation in these proteins.

Discussion

Functional site profiling has previously been applied to enzyme active sites (Cammer et al. 2003; Baxter et al. 2004), sites that are closely related compared with post-translational modification sites. Analysis of post-translational modification sites is a more difficult problem because of the diversity of the modifiable sites. Here, we have applied functional site profiling to Cys-SOH post-translational modification sites. As expected, the individual signatures are diverse and the known data set is currently small. Attempts to align the entire profile (Fig. 2) illustrate the difficulties in using a sequence-only based method to align these diverse signatures. Pairwise analysis of the functional site profile scores can be used to identify potentially related subgroups, but sequence-only methods are not powerful enough to elucidate similarities at these sites. Additional features must be included in the scoring function to better identify similarities between the signatures.

Previous observation of protein structure has suggested that the microenvironment that stabilizes Cys-SOH is characterized by three features: (1) lack of solvent accessibility to the modified cysteine; (2) lack of nearby reduced cysteines; and (3) local hydrogen-bonding residues that stabilize the sulfenate form (Claiborne et al. 1993, 1999). Others have suggested the helix dipole (Hol et al. 1978; Hol 1985; Iqbalsyah et al. 2006), interaction with histidine (Polgar 1974; Lo Bello et al. 1993), or ion pair formation (Griffiths et al. 2002) as important features in cysteine reactivity. Integration of profiling with electrostatic analysis across the entire modifiable protein set allows us to comment generally on these observations and to identify possible mechanistic determinants for cysteine deprotonation. First, solvent accessibility does not appear to be a key feature of these sites. Second, lack of other nearby cysteines does not appear important for reactivity. If the modifiable cysteine is excluded from the counts, the frequency of occurrence of cysteines in signatures is 1.2, just slightly above its average occurrence in this set of proteins overall. (The absence of proximal cysteines could still account for Cys-SOH stability, which is not addressed by the current studies.) Third, titratable, polar, but not necessarily charged, residues are important. Polar residues are overrepresented, but three of four charged residues are significantly underrepresented in the signatures (Table 2). In addition, the cysteine pK_a can be decreased without nearby charged or titratable residues (for example, 1fnjA and 1qq2A signatures in Fig. 2). Subtle polar interactions, influenced by local backbone conformation and local hydrogen-bonding patterns, thus appear to be important in decreasing the cysteine pK_a . Fourth, histidines, in particular, may be a key feature of many modifiable sites. They are overrepre-

sented in both functional site signatures generally and N-terminally adjacent to the modifiable cysteine specifically. Fifth, we have observed important interactions with Thr that shift pK_a s in some of these proteins (Table 1). Thr is overrepresented in these signatures overall (Table 2). We suggest that Thr can play an important role in cysteine modification and should be included in future calculations, not because Thr can ionize in physiological environments, but because the calculation can identify and quantify important interactions between the Thr hydroxyl and pK_a s of other residues.

Conclusions

This study provides a first general analysis of known Cys-SOH modification sites in proteins. From the structure database and experimental evidence, we have identified a set of proteins containing modifiable cysteines. We have reported here a detailed sequence, structural, and electrostatic analysis of these sites, and have calculated the cysteine pK_a s. With this study as a baseline, we can now begin to include and compare sequence and electrostatic features of other cysteine modifications such as those being identified by specific chemical probes (Moos et al. 2003; Poole et al. 2005; Dennehy et al. 2006; Greco et al. 2006), with an aim toward understanding how specificity of cysteine modification is determined. The features identified in this study will aid in prediction of protein sequences that contain modifiable cysteine sites, an important step in the identification of “candidate proteins” for the design of experiments characterizing redox signaling pathways. Analysis of the signatures and the interactions with the modifiable cysteines will aid researchers in understanding the modification mechanisms important in redox signaling and disease.

Materials and Methods

Protein data sets

The modifiable protein set consists of proteins containing one or more cysteines that are known to be modifiable to Cys-SOH: a total of 47 proteins and 49 modifiable sites (Table 1). All of these proteins are of known structure and were taken from PDB release Jan 2005 (Berman et al. 2002). Some are proteins for which Cys-SOH is generated as part of their biological function and/or there is significant biochemical evidence supporting Cys-SOH formation; others actually exhibit a Cys-SOH in their structure. Additional details about this data set can be found in the Supplemental material.

Protein sequence and structure analysis

Calculation of secondary structure, cysteine side-chain solvent accessibility, and amino acid frequency calculations are described in the Supplemental material.

Identification of functional site signatures and profiles

Functional site signatures and profiles were created and clustered for modifiable cysteine sites in the modifiable protein set using the procedures previously described (Cammer et al. 2003; Fetrow 2006) and outlined in Figure 1. Additional details can be found in the Supplemental material.

Calculation of cysteine pK_a s

pK_a s were calculated using the MEAD multiflex package (Bashford 1997), essentially as previously described (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.). Ser, Asp, Arg, Glu, His, Cys, Lys, Tyr, and Thr residues were considered titratable. Parameters for the cysteine were taken as previously validated (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.). An *intrinsic* pK_a was determined first for each titratable residue, which includes the effects on the electrostatic free energies due to solvent accessibility of the titratable site and to its interactions with the partial charges of the backbone and nontitrating residues. The *full* pK_a , which includes the effects of titratable groups, is then determined, either with Monte Carlo (Bashford 1997) or the reduced-site titration method (Bashford and Gerwert 1992; Antosiewicz et al. 1994; Bashford 1997; van Vlijmen et al. 1998). Model parameters for Thr were nonexistent and were developed as recently described for cysteines (F.R. Salsbury Jr., L.B. Poole, and J.S. Fetrow, in prep.). Details of the cysteine and threonine charge models can be found in Supplemental Table S1. Additional methodological details can be found in the Supplemental material.

Interaction between modifiable Cys residues and other titratable residues

The initial step in identification of residues affecting the modifiable Cys pK_a is construction of a fully protonated protein reference state. The pK_a calculations were performed using MEAD (Bashford 1997). To calculate interaction energies, the electrostatics were calculated with each titratable residue (one at a time) in the protonated and deprotonated forms. These site-site interaction energies are used to identify interacting residues or interactors—those side chains that are interacting with the residue of interest and participating in its pK_a shift. Residues interacting with the modifiable Cys are identified as strong interactors when the interaction energy (measured in pK units) is >1.0 pK units, and weak interactors when the interaction energy is between 0.5 and 1.0 pK units.

Acknowledgments

We thank Mick Knaggs and Todd Lowther for helpful discussions. We acknowledge support from both the NIH (R21 CA112145) to L.B.P. and the NSF (MCB-0517343) to J.S.F. These calculations were performed on Wake Forest University's DEAC cluster (<http://www.deac.wfu.edu>) including a SUR grant from IBM for storage hardware, and the support of the Wake Forest IS department is gratefully acknowledged.

References

Abate, C., Patel, L., Rauscher III, F.J., and Curran, T. 1990. Redox regulation of Fos and Jun DNA-binding activity *in vitro*. *Science* **249**: 1157–1161.

- Allison, W.S. 1976. Formation and reactions of sulfenic acids in proteins. *Acc. Chem. Res.* **9**: 293–299.
- Alphey, M.S., Bond, C.S., Tetaud, E., Fairlamb, A.H., and Hunter, W.N. 2000. The structure of reduced trypanothione peroxidase reveals a decamer and insight into reactivity of 2Cys-peroxiredoxins. *J. Mol. Biol.* **300**: 903–916.
- Anstrom, D.M., Kallio, K., and Remington, S.J. 2003. Structure of the *Escherichia coli* malate synthase G:pyruvate:acetyl-coenzyme A abortive ternary complex at 1.95 Å resolution. *Protein Sci.* **12**: 1822–1832.
- Antosiewicz, J., McCammon, J.A., and Gilson, M.K. 1994. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**: 415–436.
- Badger, J., Sauder, J.M., Adams, J.M., Antonysamy, S., Bain, K., Bergsied, M.G., Buchanan, S.G., Buchanan, M.D., Batiyenko, Y., Christopher, J.A., et al. 2005. Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins* **60**: 787–796.
- Baker, L.M. and Poole, L.B. 2003. Catalytic mechanism of thiol peroxidase from *Escherichia coli*. Sulfenic acid formation and overoxidation of essential CYS61. *J. Biol. Chem.* **278**: 9203–9211.
- Bashford, D. 1997. An object-oriented programming suite for electrostatic effects in biological molecules. In *Lecture notes in computer science* (eds. Y. Ishikawa et al.), pp. 233–240. Springer, Berlin, Germany.
- Bashford, D. and Gerwert, K. 1992. Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* **224**: 473–486.
- Baxter, S.M., Rosenblum, J.S., Knutson, S., Nelson, M.R., Montimurro, J.S., Di Gennaro, J.A., Speir, J.A., Burbaum, J.J., and Fetrow, J.S. 2004. Synergistic computational and experimental proteomics approaches for more accurate detection of active serine hydrolases in yeast. *Mol. Cell. Proteomics* **3**: 209–225.
- Becker, K., Savvides, S.N., Keese, M., Schirmer, R.H., and Karplus, P.A. 1998. Enzyme inactivation through sulfhydryl oxidation by physiologic NO-carriers. *Nat. Struct. Biol.* **5**: 267–271.
- Berlett, B.S. and Stadtman, E.R. 1997. Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.* **272**: 20313–20316.
- Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., et al. 2002. The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 899–907.
- Boschi-Muller, S., Azza, S., Sanglier-Cianferani, S., Talfournier, F., Van Dorsselaar, A., and Branlant, G. 2000. A sulfenic acid enzyme intermediate is involved in the catalytic mechanism of peptide methionine sulfoxide reductase from *Escherichia coli*. *J. Biol. Chem.* **275**: 35908–35913.
- Brinkworth, R.I., Horne, J., and Kobe, B. 2002. A computational analysis of substrate binding strength by phosphorylase kinase and protein kinase A. *J. Mol. Recognit.* **15**: 104–111.
- Brinkworth, R.I., Munn, A.L., and Kobe, B. 2006. Protein kinases associated with the yeast phosphoproteome. *BMC Bioinformatics* **7**: 47. doi: 10.1186/1471-2105-7-47.
- Cammer, S.A., Hoffman, B.T., Speir, J.A., Canady, M.A., Nelson, M.R., Knutson, S., Gallina, M., Baxter, S.M., and Fetrow, J.S. 2003. Structure-based active site profiles for genome analysis and functional family subclassification. *J. Mol. Biol.* **334**: 387–401.
- Carballal, S., Radi, R., Kirk, M.C., Barnes, S., Freeman, B.A., and Alvarez, B. 2003. Sulfenic acid formation in human serum albumin by hydrogen peroxide and peroxynitrite. *Biochemistry* **42**: 9906–9914.
- Chan, N.L., Kavanagh, J.S., Rogers, P.H., and Arnone, A. 2004. Crystallographic analysis of the interaction of nitric oxide with quaternary-T human hemoglobin. *Biochemistry* **43**: 118–132.
- Changela, A., Ho, C.K., Martins, A., Shuman, S., and Mondragon, A. 2001. Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J.* **20**: 2575–2586.
- Chen, F.E., Huang, D.B., Chen, Y.Q., and Ghosh, G. 1998a. Crystal structure of p50/p65 heterodimer of transcription factor NF- κ B bound to DNA. *Nature* **391**: 410–413.
- Chen, L., Glover, J.N., Hogan, P.G., Rao, A., and Harrison, S.C. 1998b. Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature* **392**: 42–48.
- Cho, S.H., Lee, C.H., Ahn, Y., Kim, H., Ahn, C.Y., Yang, K.S., and Lee, S.R. 2004. Redox regulation of PTEN and protein tyrosine phosphatases in H₂O₂-mediated cell signaling. *FEBS Lett.* **560**: 7–13.
- Choi, H.J., Kang, S.W., Yang, C.H., Rhee, S.G., and Ryu, S.E. 1998. Crystal structure of a novel human peroxidase enzyme at 2.0 Å resolution. *Nat. Struct. Biol.* **5**: 400–406.
- Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G., and Ryu, S. 2001. Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**: 103–113.
- Choi, J., Choi, S., Cha, M.K., Kim, I.H., and Shin, W. 2003. Crystal structure of *Escherichia coli* thiol peroxidase in the oxidized state: Insights into

- intramolecular disulfide formation and substrate binding in atypical 2-Cys peroxiredoxins. *J. Biol. Chem.* **278**: 49478–49486.
- Claiborne, A., Miller, H., Parsonage, D., and Ross, R.P. 1993. Protein-sulfenic acid stabilization and function in enzyme catalysis and gene regulation. *FASEB J.* **7**: 1483–1490.
- Claiborne, A., Yeh, J.I., Mallett, T.C., Luba, J., Crane 3rd, E.J., Charrier, V., and Parsonage, D. 1999. Protein-sulfenic acids: Diverse roles for an unlikely player in enzyme catalysis and redox regulation. *Biochemistry* **38**: 15407–15416.
- Conway, M.E., Yennawar, N., Wallin, R., Poole, L.B., and Hutson, S.M. 2003. Human mitochondrial branched chain aminotransferase: Structural basis for substrate specificity and role of redox active cysteines. *Biochim. Biophys. Acta* **1647**: 61–65.
- Conway, M.E., Poole, L.B., and Hutson, S.M. 2004. Roles for cysteine residues in the regulatory CXXC motif of human mitochondrial branched chain aminotransferase enzyme. *Biochemistry* **43**: 7356–7364.
- Copley, S.D., Novak, W.R., and Babbitt, P.C. 2004. Divergence of function in the thioredoxin fold suprafamily: Evidence for evolution of peroxiredoxins from a thioredoxin-like ancestor. *Biochemistry* **43**: 13981–13995.
- Cowan-Jacob, S.W., Kaufmann, M., Anselmo, A.N., Stark, W., and Grutter, M.G. 2003. Structure of rabbit-muscle glyceraldehyde-3-phosphate dehydrogenase. *Acta Crystallogr. D Biol. Crystallogr.* **59**: 2218–2227.
- Cross, J.V. and Templeton, D.J. 2006. Regulation of signal transduction through protein cysteine oxidation. *Antioxid. Redox Signal.* **8**: 1819–1827.
- Deacon, A.M., Ni, Y.S., Coleman Jr., W.G., and Ealick, S.E. 2000. The crystal structure of ADP-L-glycero-D-mannoheptose 6-epimerase: Catalysis with a twist. *Structure* **8**: 453–462.
- Declercq, J.P., Evrard, C., Clippe, A., Stricht, D.V., Bernard, A., and Knoops, B. 2001. Crystal structure of human peroxiredoxin 5, a novel type of mammalian peroxiredoxin at 1.5 Å resolution. *J. Mol. Biol.* **311**: 751–759.
- Dennehy, M.K., Richards, K.A., Wernke, G.R., Shyr, Y., and Liebler, D.C. 2006. Cytosolic and nuclear protein targets of thiol-reactive electrophiles. *Chem. Res. Toxicol.* **19**: 20–29.
- Denu, J.M. and Tanner, K.G. 1998. Specific and reversible inactivation of protein tyrosine phosphatases by hydrogen peroxide: Evidence for a sulfenic acid intermediate and implications for redox regulation. *Biochemistry* **37**: 5633–5642.
- Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H., and Cheng, X. 2001. Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. *Nucleic Acids Res.* **29**: 439–448.
- Fetrow, J.S. 2006. Active site profiling to identify protein functional sites in sequences and structures using the Deacon active site profiler (DASP) In *Current Protocols in Bioinformatics*, Unit 8.10. Wiley, Hoboken, NJ.
- Fiedler, T.J., Davey, C.A., and Fenna, R.E. 2000. X-ray crystal structure and characterization of halide-binding sites of human myeloperoxidase at 1.8 Å resolution. *J. Biol. Chem.* **275**: 11964–11971.
- Flohe, L., Budde, H., Bruns, K., Castro, H., Clos, J., Hofmann, B., Kansal-Kalavar, S., Krumme, D., Menge, U., Plank-Schumacher, K., et al. 2002. Tryparedoxin peroxidase of *Leishmania donovani*: Molecular cloning, heterologous expression, specificity, and catalytic mechanism. *Arch. Biochem. Biophys.* **397**: 324–335.
- Fuangthong, M. and Helmann, J.D. 2002. The OhrR repressor senses organic hydroperoxides by reversible formation of a cysteine-sulfenic acid derivative. *Proc. Natl. Acad. Sci.* **99**: 6690–6695.
- Gaudier, M., Gaudin, Y., and Knossow, M. 2002. Crystal structure of vesicular stomatitis virus matrix protein. *EMBO J.* **21**: 2886–2892.
- Graille, M., Quevillon-Cheruel, S., Leulliot, N., Zhou, C.Z., de la Sierra Gallay, I.L., Jacquamet, L., Ferrer, J.L., Liger, D., Poupon, A., Janin, J., et al. 2004. Crystal structure of the YDR533c *S. cerevisiae* protein, a class II member of the Hsp31 family. *Structure* **12**: 839–847.
- Greco, T.M., Hodara, R., Parastatidis, I., Heijnen, H.F., Dennehy, M.K., Liebler, D.C., and Ischiropoulos, H. 2006. Identification of S-nitrosylation motifs by site-specific mapping of the S-nitrosocysteine proteome in human vascular smooth muscle cells. *Proc. Natl. Acad. Sci.* **103**: 7420–7425.
- Griffiths, S.W., King, J., and Cooney, C.L. 2002. The reactivity and oxidation pathway of cysteine 232 in recombinant human α 1-antitrypsin. *J. Biol. Chem.* **277**: 25486–25492.
- Guan, R., Malchiodi, E.L., Wang, Q., Schuck, P., and Mariuzza, R.A. 2004. Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein α . *J. Biol. Chem.* **279**: 31873–31882.
- Hegde, R.S., Grossman, S.R., Laimins, L.A., and Sigler, P.B. 1992. Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* **359**: 505–512.
- Higgins, D.G., Thompson, J.D., and Gibson, T.J. 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**: 383–402.
- Hirotsu, S., Abe, Y., Okada, K., Nagahara, N., Hori, H., Nishino, T., and Hakoshima, T. 1999. Crystal structure of a multifunctional 2-Cys peroxiredoxin heme-binding protein 23 kDa/proliferation-associated gene product. *Proc. Natl. Acad. Sci.* **96**: 12333–12338.
- Hofmann, B., Hecht, H.-J., and Flohé, L. 2002. Peroxiredoxins. *Biol. Chem.* **383**: 347–364.
- Hol, W.G.J. 1985. The role of the α -helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.* **45**: 149–195.
- Hol, W.G., van Duijnen, P.T., and Berendsen, H.J. 1978. The α -helix dipole and the properties of proteins. *Nature* **273**: 443–446.
- Howard, B.R., Endrizzi, J.A., and Remington, S.J. 2000. Crystal structure of *Escherichia coli* malate synthase G complexed with magnesium and glyoxylate at 2.0 Å resolution: Mechanistic implications. *Biochemistry* **39**: 3156–3168.
- Huang, H.D., Lee, T.Y., Tzeng, S.W., and Horng, J.T. 2005. KinasePhos: A web tool for identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res.* **33**: W226–W229. doi: 10.1093/nar/gki471.
- Huff, R.G., Bayram, E., Tan, H., Knutson, S.T., Knaggs, M.H., Richon, A.B., Santago, P., and Fetrow, J.S. 2005. Chemical and structural diversity in cyclooxygenase protein active sites. *Chem. and Biodivers.* **2**: 1533–1552.
- Humphrey, W., Dalke, A., and Schulten, K. 1996. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**: 27–38.
- Iqbalsyah, T.M., Moutevelis, E., Warwicker, J., Errington, N., and Doig, A.J. 2006. The CXXC motif at the N terminus of an α -helical peptide. *Protein Sci.* **15**: 1945–1950.
- Jacob, C., Holme, A.L., and Fry, F.H. 2004. The sulfinic acid switch in proteins. *Org. Biomol. Chem.* **2**: 1953–1956.
- Jones, S., Shanahan, H.P., Berman, H.M., and Thornton, J.M. 2003. Using electrostatic potentials to predict DNA-binding sites on DNA-binding proteins. *Nucleic Acids Res.* **31**: 7189–7198.
- Juers, D.H., Heightman, T.D., Vasella, A., McCarter, J.D., Mackenzie, L., Withers, S.G., and Matthews, B.W. 2001. A structural view of the action of *Escherichia coli* (lacZ) β -galactosidase. *Biochemistry* **40**: 14781–14794.
- Kamphuis, I.G., Kalk, K.H., Swarte, M.B., and Drenth, J. 1984. Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.* **179**: 233–256.
- Kast, P., Grisostomi, C., Chen, I.A., Li, S., Krenkel, U., Xue, Y., and Hilvert, D. 2000. A strategically positioned cation is crucial for efficient catalysis by chorismate mutase. *J. Biol. Chem.* **275**: 36832–36838.
- Kim, S.O., Merchant, K., Nudelman, R., Beyer, W.F., Keng, T., DeAngelo, J., Hausladen, A., and Stamler, J.S. 2002. OxyR: A molecular code for redox-related signaling. *Cell* **109**: 383–396.
- Kobe, B., Kampmann, T., Forwood, J.K., Listwan, P., and Brinkworth, R.I. 2005. Substrate specificity of protein kinases and computational prediction of substrates. *Biochim. Biophys. Acta* **1754**: 200–209.
- Koenig, M. and Grabe, N. 2004. Highly specific prediction of phosphorylation sites in proteins. *Bioinformatics* **20**: 3620–3627.
- Lange, G., Lesuisse, D., Deprez, P., Schoot, B., Loenze, P., Benard, D., Marquette, J.P., Broto, P., Sarubbi, E., and Mandine, E. 2003. Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of (pp60)Src are identical to those for high affinity binding of full length inhibitors. *J. Med. Chem.* **46**: 5184–5195.
- Lesniak, J., Barton, W.A., and Nikolov, D.B. 2002. Structural and functional characterization of Ohr, an organic hydroperoxide resistance protein from *Pseudomonas aeruginosa*. *EMBO J.* **21**: 6649–6659.
- Lesniak, J., Barton, W.A., and Nikolov, D.B. 2003. Structural and functional features of the *Escherichia coli* hydroperoxide resistance protein OsmC. *Protein Sci.* **12**: 2838–2843.
- Lo Bello, M., Parker, M.W., Desideri, A., Polticelli, F., Falconi, M., Del Boccio, G., Pennelli, A., Federici, G., and Ricci, G. 1993. Peculiar spectroscopic and kinetic properties of Cys-47 in human placental glutathione transferase. Evidence for an atypical thiolate ion pair near the active site. *J. Biol. Chem.* **268**: 19033–19038.
- Lowther, W.T., Brot, N., Weissbach, H., and Matthews, B.W. 2000. Structure and mechanism of peptide methionine sulfoxide reductase, an “anti-oxidation” enzyme. *Biochemistry* **39**: 13307–13312.
- Lowther, W.T., Weissbach, H., Etienne, F., Brot, N., and Matthews, B.W. 2002. The mirrored methionine sulfoxide reductases of *Neisseria gonorrhoeae* pilB. *Nat. Struct. Biol.* **9**: 348–352.
- Martin, G., Moglich, A., Keller, W., and Doubie, S. 2004. Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. *J. Mol. Biol.* **341**: 911–925.
- McBride, A.A., Klausner, R.D., and Howley, P.M. 1992. Conserved cysteine residue in the DNA-binding domain of the bovine papillomavirus type 1 E2

- protein confers redox regulation of the DNA-binding activity *in vitro*. *Proc. Natl. Acad. Sci.* **89**: 7531–7535.
- McGrath, M.E., Klaus, J.L., Barnes, M.G., and Bromme, D. 1997. Crystal structure of human cathepsin K complexed with a potent inhibitor. *Nat. Struct. Biol.* **4**: 105–109.
- Miyana, A., Fushinobu, S., Ito, K., and Wakagi, T. 2001. Crystal structure of cobalt-containing nitrile hydratase. *Biochem. Biophys. Res. Commun.* **288**: 1169–1174.
- Molina, H., Horn, D.M., Tang, N., Mathivanan, S., and Pandey, A. 2007. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci.* **104**: 2199–2204.
- Montemartini, M., Kalisz, H.M., Hecht, H.J., Steinert, P., and Flohe, L. 1999. Activation of active-site cysteine residues in the peroxiredoxin-type trypanoxin peroxidase of *Crithidia fasciculata*. *Eur. J. Biochem.* **264**: 516–524.
- Moos, P.J., Edes, K., Cassidy, P., Massuda, E., and Fitzpatrick, F.A. 2003. Electrophilic prostaglandins and lipid aldehydes repress redox-sensitive transcription factors p53 and hypoxia-inducible factor by impairing the selenoprotein thioredoxin reductase. *J. Biol. Chem.* **278**: 745–750.
- Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**: 536–540.
- Nagashima, S., Nakasako, M., Dohmae, N., Tsujimura, M., Takio, K., Odaka, M., Yohda, M., Kamiya, N., and Endo, I. 1998. Novel non-heme iron center of nitrile hydratase with a claw setting of oxygen atoms. *Nat. Struct. Biol.* **5**: 347–351.
- Nardini, M., Spano, S., Cericola, C., Pesce, A., Massaro, A., Millo, E., Luini, A., Corda, D., and Bolognesi, M. 2003. CtBP/BARS: A dual-function protein involved in transcription co-repression and Golgi membrane fission. *EMBO J.* **22**: 3122–3130.
- Oliveira, M.A., Guimaraes, B.G., Cussiol, J.R., Medrano, F.J., Gozzo, F.C., and Netto, L.E. 2006. Structural insights into enzyme-substrate interaction and characterization of enzymatic intermediates of organic hydroperoxide resistance protein from *Xylella fastidiosa*. *J. Mol. Biol.* **359**: 433–445.
- Olry, A., Boschi-Muller, S., Marraud, M., Sanglier-Cianferani, S., Van Dorsselaar, A., and Branlant, G. 2002. Characterization of the methionine sulfoxide reductase activities of PILB, a probable virulence factor from *Neisseria meningitidis*. *J. Biol. Chem.* **277**: 12016–12022.
- Ondrechen, M.J., Clifton, J.G., and Ringe, D. 2001. THEMATICS: A simple computational predictor of enzyme function from structure. *Proc. Natl. Acad. Sci.* **98**: 12473–12478.
- Panmanee, W., Vattanaviboon, P., Poole, L.B., and Mongkolsuk, S. 2006. Novel organic hydroperoxide-sensing and responding mechanisms for OhrR, a major bacterial sensor and regulator of organic hydroperoxide stress. *J. Bacteriol.* **188**: 1389–1395.
- Parsonage, D., Youngblood, D.S., Sarma, G.N., Wood, Z.A., Karplus, P.A., and Poole, L.B. 2005. Analysis of the link between enzymatic activity and oligomeric state in AhpC, a bacterial peroxiredoxin. *Biochemistry* **44**: 10583–10592.
- Percival, M.D., Ouellet, M., Campagnolo, C., Claveau, D., and Li, C. 1999. Inhibition of cathepsin K by nitric oxide donors: Evidence for the formation of mixed disulfides and a sulfenic acid. *Biochemistry* **38**: 13574–13583.
- Peters, G.H., Frimurer, T.M., and Olsen, O.H. 1998. Electrostatic evaluation of the signature motif (H/V)CX5R(S/T) in protein-tyrosine phosphatases. *Biochemistry* **37**: 5383–5393.
- Pineda-Molina, E., Klatt, P., Vázquez, J., Marina, A., García de Lacoba, M., Pérez-Sala, D., and Lamas, S. 2001. Glutathionylation of the p50 subunit of NF- κ B: A mechanism for redox-induced inhibition of DNA binding. *Biochemistry* **40**: 14134–14142.
- Plewczynski, D., Tkacz, A., Wyrwicz, L.S., and Rychlewski, L. 2005. AutoMotif server: Prediction of single residue post-translational modifications in proteins. *Bioinformatics* **21**: 2525–2527.
- Polgar, L. 1974. Mercaptide-imidazolium ion-pair: The reactive nucleophile in papain catalysis. *FEBS Lett.* **47**: 15–18.
- Poole, L.B. 2007. The catalytic mechanism of peroxiredoxins. In *Peroxiredoxin systems* (eds. L. Flohé and J.R. Harris), pp. 61–81. Springer, New York.
- Poole, L.B. and Claiborne, A. 1989. The non-flavin redox center of the streptococcal NADH peroxidase. II. Evidence for a stabilized cysteine-sulfenic acid. *J. Biol. Chem.* **264**: 12330–12338.
- Poole, L.B. and Ellis, H.R. 2002. Identification of cysteine sulfenic acid in AhpC of alkyl hydroperoxide reductase. *Methods Enzymol.* **348**: 122–136.
- Poole, L.B., Karplus, P.A., and Claiborne, A. 2004. Protein sulfenic acids in redox signaling. *Annu. Rev. Pharmacol. Toxicol.* **44**: 325–347.
- Poole, L.B., Zeng, B.B., Knaggs, S.A., Yakubu, M., and King, S.B. 2005. Synthesis of chemical probes to map sulfenic acid modifications on proteins. *Bioconjug. Chem.* **16**: 1624–1628.
- Rehse, P.H., Ohshima, N., Nodake, Y., and Tahirov, T.H. 2004. Crystallographic structure and biochemical analysis of the *Thermus thermophilus* osmotically inducible protein C. *J. Mol. Biol.* **338**: 959–968.
- Ryu, S.E., Choi, H.J., Kwon, K.S., Lee, K.N., and Yu, M.H. 1996. The native strains in the hydrophobic core and flexible reactive loop of a serine protease inhibitor: Crystal structure of an uncleaved α 1-antitrypsin at 2.7 Å. *Structure* **4**: 1181–1192.
- Schröder, E., Littlechild, J.A., Lebedev, A.A., Errington, N., Vagin, A.A., and Isupov, M.N. 2000. Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 Å resolution. *Structure* **8**: 605–615.
- Sham, Y.Y., Chu, Z.T., and Warshel, A. 1997. Consistent calculations of pK_as of ionizable residues in proteins: Semi-microscopic and microscopic approaches. *J. Phys. Chem. B* **101**: 4458–4472.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Tonks, N.K. 2005. Redox redux: Revisiting PTPs and the control of cell signaling. *Cell* **121**: 667–670.
- van Montfort, R.L., Congreve, M., Tisi, D., Carr, R., and Jhoti, H. 2003. Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature* **423**: 773–777.
- van Vlijmen, H.W., Schaefer, M., and Karplus, M. 1998. Improving the accuracy of protein pK_a calculations: Conformational averaging versus the average structure. *Proteins* **33**: 145–158.
- Wardell, M., Wang, Z., Ho, J.X., Robert, J., Ruker, F., Ruble, J., and Carter, D.C. 2002. The atomic structure of human methalbumin at 1.9 Å. *Biochem. Biophys. Res. Commun.* **291**: 813–819.
- Wood, Z.A., Poole, L.B., Hantgan, R.R., and Karplus, P.A. 2002. Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* **41**: 5493–5504.
- Wood, Z.A., Poole, L.B., and Karplus, P.A. 2003a. Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* **300**: 650–653.
- Wood, Z.A., Schröder, E., Harris, J.R., and Poole, L.B. 2003b. Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem. Sci.* **28**: 32–40.
- Xu, G. and Chance, M.R. 2005. Radiolytic modification of sulfur-containing amino acid residues in model peptides: Fundamental studies for protein footprinting. *Anal. Chem.* **77**: 2437–2449.
- Xue, Y., Zhou, F., Zhu, M., Ahmed, K., Chen, G., and Yao, X. 2005. GPS: A comprehensive web server for phosphorylation sites prediction. *Nucleic Acids Res.* **33**: W184–W187. doi: 10.1093/nar/gki393.
- Yan, Y., Harper, S., Speicher, D.W., and Marmorstein, R. 2002. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat. Struct. Biol.* **9**: 862–869.
- Yeh, J.I., Claiborne, A., and Hol, W.G. 1996. Structure of the native cysteine-sulfenic acid redox center of enterococcal NADH peroxidase refined at 2.8 Å resolution. *Biochemistry* **35**: 9951–9957.
- Yennawar, N., Dunbar, J., Conway, M., Hutson, S., and Farber, G. 2001. The structure of human mitochondrial branched-chain aminotransferase. *Acta Crystallogr. D Biol. Crystallogr.* **57**: 506–515.
- Zheng, M., Aslund, F., and Storz, G. 1998. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**: 1718–1721.