

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5284952>

# Cysteine pKa Depression by a Protonated Glutamic Acid in Human DJ-1 †‡

ARTICLE *in* BIOCHEMISTRY · AUGUST 2008

Impact Factor: 3.02 · DOI: 10.1021/bi800282d · Source: PubMed

---

CITATIONS

47

READS

52

## 6 AUTHORS, INCLUDING:



Benjamin C Remington

TNO

22 PUBLICATIONS 220 CITATIONS

[SEE PROFILE](#)



Sahar Hasim

University of Tennessee

6 PUBLICATIONS 102 CITATIONS

[SEE PROFILE](#)



Edwin Pozharski

University of Maryland, Baltimore

45 PUBLICATIONS 1,339 CITATIONS

[SEE PROFILE](#)



Mark A Wilson

University of Nebraska at Lincoln

39 PUBLICATIONS 2,365 CITATIONS

[SEE PROFILE](#)

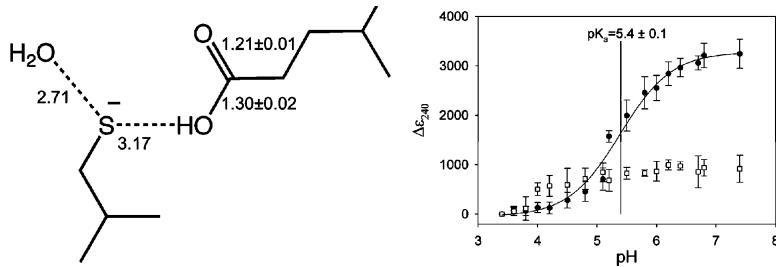
## Article

### Cysteine pK Depression by a Protonated Glutamic Acid in Human DJ-1<sup>1</sup>

Anna C. Witt, Mahadevan Lakshminarasimhan, Benjamin C. Remington, Sahar Hasim, Edwin Pozharski, and Mark A. Wilson

*Biochemistry*, 2008, 47 (28), 7430-7440 • DOI: 10.1021/bi800282d • Publication Date (Web): 21 June 2008

Downloaded from <http://pubs.acs.org> on November 17, 2008



### More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



ACS Publications

High quality. High impact.

Biochemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

## Cysteine p<sub>K</sub><sub>a</sub> Depression by a Protonated Glutamic Acid in Human DJ-1<sup>†,‡</sup>

Anna C. Witt,<sup>§</sup> Mahadevan Lakshminarasimhan,<sup>§</sup> Benjamin C. Remington,<sup>§</sup> Sahar Hasim,<sup>§</sup> Edwin Pozharski,<sup>||</sup> and Mark A. Wilson\*,<sup>§</sup>

Department of Biochemistry and Redox Biology Center, The University of Nebraska, Lincoln, Nebraska 68588-0664, and Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland 21201

Received February 18, 2008; Revised Manuscript Received May 29, 2008

**ABSTRACT:** Human DJ-1, a disease-associated protein that protects cells from oxidative stress, contains an oxidation-sensitive cysteine (C106) that is essential for its cytoprotective activity. The origin of C106 reactivity is obscure, due in part to the absence of an experimentally determined p<sub>K</sub><sub>a</sub> value for this residue. We have used atomic-resolution X-ray crystallography and UV spectroscopy to show that C106 has a depressed p<sub>K</sub><sub>a</sub> of 5.4 ± 0.1 and that the C106 thiolate accepts a hydrogen bond from a protonated glutamic acid side chain (E18). X-ray crystal structures and cysteine p<sub>K</sub><sub>a</sub> analysis of several site-directed substitutions at residue 18 demonstrate that the protonated carboxylic acid side chain of E18 is required for the maximal stabilization of the C106 thiolate. A nearby arginine residue (R48) participates in a guanidinium stacking interaction with R28 from the other monomer in the DJ-1 dimer and elevates the p<sub>K</sub><sub>a</sub> of C106 by binding an anion that electrostatically suppresses thiol ionization. Our results show that the ionizable residues (E18, R48, and R28) surrounding C106 affect its p<sub>K</sub><sub>a</sub> in a way that is contrary to expectations based on the typical ionization behavior of glutamic acid and arginine. Lastly, a search of the Protein Data Bank (PDB) produces several candidate hydrogen-bonded aspartic/glutamic acid–cysteine interactions, which we propose are particularly common in the DJ-1 superfamily.

Cysteine thiolates are potent nucleophiles that are used by many proteins for catalysis, metal binding, or to facilitate post-translational modification. However, the solution p<sub>K</sub><sub>a</sub> value of cysteine (8.3) is not within the optimal physiological pH range of most proteins. To render cysteine residues reactive, the thiol p<sub>K</sub><sub>a</sub> must be depressed by stabilization of the conjugate thiolate anion by the protein environment. Multiple types of stabilizing interactions can decrease cysteine p<sub>K</sub><sub>a</sub> values, including electrostatic complementarity with nearby cations (1, 2), the α-helix macropole (3), and hydrogen bonding to the thiol (4).

The Parkinsonism-associated protein DJ-1 contains an oxidation-sensitive cysteine residue (C106 in human DJ-1) of unknown p<sub>K</sub><sub>a</sub> that is required for the cytoprotective activity of the protein (5, 6). DJ-1 is a homodimeric protein of 189 amino acid monomers that is found both in the cytoplasm (7) and in the mitochondria (5, 8) and protects cells from

various types of oxidative stress (5, 9–18). Although C106 has been shown to be essential for the protective function of DJ-1 in cell culture (5, 16) and *Drosophila melanogaster* (6), the specific biochemical activity of the protein that requires this cysteine residue is uncertain. Identifying the structural determinants of C106 ionization (and hence reactivity) is an essential first step in understanding how this conserved cysteine residue contributes to the neuroprotective activity of DJ-1.

The reactive cysteine in human DJ-1 is one of the most highly conserved residues in the DJ-1 superfamily (19, 20). Crystal structures determined by several independent groups (21–25) show that C106 is located at a solvent-exposed sharp bend between a β-strand and an α-helix called the “nucleophile elbow” (26). The backbone φ and ψ angles for the nucleophile elbow cysteine fall in the unfavorable region of Ramachandran space for every structurally characterized member of the DJ-1 superfamily (27), possibly contributing to its reactivity. Furthermore, a computational study of structurally characterized proteins in the DJ-1 superfamily shows that functionally important residues with perturbed *in silico* pH titration profiles cluster around the reactive cysteine residue and can be used to classify distinct members of the superfamily (28). Of the residues in the environment of the conserved cysteine, a proximal glutamic acid (E18 in human DJ-1) is found in many members of the DJ-1 superfamily. The importance of this residue, however, remains poorly understood (27, 28).

To improve our understanding of the origin of cysteine reactivity in DJ-1, we have used a combination of structural and biochemical methods to show that C106 has a depressed p<sub>K</sub><sub>a</sub> value of 5.4 ± 0.1 and that the environment of C106

<sup>†</sup> This work was supported in part by a grant to M.A.W. from the American Parkinson’s Disease Association as well as a grant from the National Institutes of Health (P 20RR-17675). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. Use of BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under Grant R01707.

<sup>‡</sup> Refined model coordinates and experimental structure factors have been deposited in the Protein Data Bank as entries 2OR3 (orthorhombic wtDJ-1), 3CY6 (E18Q), 3CYF (E18N), 3CZ9 (E18L), and 3CZA (E18D).

\* To whom correspondence should be addressed: N164 Beadle Center, University of Nebraska, Lincoln, NE 68588-0664. E-mail: mwilson13@unl.edu. Phone: (402) 472-3626. Fax: (402) 472-4961.

<sup>§</sup> The University of Nebraska.

<sup>||</sup> University of Maryland School of Pharmacy.

contains multiple ionizable residues that exert countervailing influence on the  $pK_a$  of this residue. The conserved E18 residue has a protonated carboxylic acid side chain that enhances cysteine ionization by donating a hydrogen bond to the  $S\gamma$  atom of C106, while R28 and R48 interact through guanidinium stacking and bind an anion that suppresses C106 thiolate formation. The significance of the low  $pK_a$  value for C106 for the proposed biochemical activities and oxidative regulation of DJ-1 is discussed.

## EXPERIMENTAL PROCEDURES

**Protein Expression and Purification.** Three forms of recombinant DJ-1 were used in this study: untagged DJ-1, DJ-1 bearing a noncleavable C-terminal hexahistidine tag, and DJ-1 bearing an N-terminal thrombin-cleavable hexahistidine tag. DJ-1 with a noncleavable C-terminal hexahistidine tag was expressed in bacterial expression vector pET21a bearing the DJ-1 gene cloned between the NdeI and XhoI sites. This construct results in eight vector-derived amino acids (LEHHHHH) appended to the C-terminus of the expressed protein and was used for DJ-1 crystallization. For the cysteine  $pK_a$  experiments, DJ-1 with an N-terminal thrombin-cleavable hexahistidine tag was expressed in bacterial expression vector pET15b bearing the DJ-1 gene cloned between the NdeI and XhoI sites. After tag removal, this construct results in a recombinant protein with three vector-derived amino acids (GSH) at the N-terminus of the protein and is called “tag-cleaved” DJ-1 throughout the remainder of the paper. For all recombinant proteins, *Escherichia coli* BL21(DE3) cells (Novagen) containing the appropriate DJ-1 expression construct were grown in LB medium supplemented with 100  $\mu$ g/mL ampicillin at 37 °C with shaking, and protein expression was induced by the addition of 0.1 mM IPTG. The induced culture was incubated at 20 °C with shaking overnight and harvested by centrifugation the next day. Cell pellets were stored at –80 °C until they were needed.

Hexahistidine-tagged recombinant DJ-1 was purified by Ni<sup>2+</sup>-NTA chromatography using His-Select resin (Sigma) following the manufacturer’s instructions. For C-terminally His-tagged protein, the Ni<sup>2+</sup>-NTA purification was followed by dialysis against storage buffer [25 mM HEPES (pH 7.5), 100 mM KCl, and 2 mM DTT]<sup>1</sup> and passage over a Q anion exchange resin to remove contaminating nucleic acid. DJ-1 appears in the flow-through fraction, while contaminating nucleic acids remain bound to the Q column. For the N-terminally His-tagged DJ-1, the hexahistidine tag was removed by thrombin cleavage at 4 °C for 6–12 h, followed by sequential passage over His-Select resin to remove any protein that retained the tag, benzamidine-Sepharose resin to remove thrombin, and a final passage over a Q anion exchange resin to remove nucleic acid. Untagged DJ-1 was purified as previously described (29). For all preparations, purified DJ-1 runs as a single band with an apparent molecular mass of 25 kDa on overloaded Biosafe (Bio-Rad) Coomassie-stained SDS-PAGE. DJ-1 was concentrated to 20 mg/mL ( $\epsilon_{280} = 4000 \text{ M}^{-1} \text{ cm}^{-1}$ ) in storage buffer, frozen in liquid nitrogen, and stored at –80 °C.

<sup>1</sup> Abbreviations: ADP, anisotropic atomic displacement parameter; DTT, dithiothreitol; EGCG, epigallocatechin 3-gallate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; IPTG, isopropyl β-D-1-thiogalactopyranoside; PEG, polyethylene glycol; PDB, Protein Data Bank; rmsd, root-mean-square deviation; UV, ultraviolet; wt, wild-type.

**Crystallization, Data Collection, and Processing.** We grew crystals of untagged wtDJ-1 in space group  $P2_12_12_1$  with two molecules in the asymmetric unit (termed orthorhombic wt throughout the remainder of the paper) using the hanging drop vapor equilibration method by mixing 2  $\mu$ L of DJ-1 (20 mg/mL) supplemented with 2.5 mM epigallocatechin 3-gallate (EGCG) with 2  $\mu$ L of reservoir solution [28% PEG 4000, 200 mM ammonium sulfate, and 50 mM sodium acetate (pH 6.0)] and incubating the mixture at room temperature. Platelike crystals appeared in 2–5 days and were cryoprotected by serial transfer through increasing concentrations of ethylene glycol in the reservoir solution to a final concentration of 25% (v/v). The crystals were mounted in nylon loops and cryocooled by immersion into liquid nitrogen. EGCG facilitates crystal growth, although it does not appear in the final electron density maps.

We grew crystals of C-terminally His-tagged E18Q, E18N, E18L, and E18D DJ-1 using the hanging drop vapor equilibration method by mixing 2  $\mu$ L of DJ-1 (20 mg/mL) with 2  $\mu$ L of reservoir solution [1.4–1.6 M sodium citrate, 50 mM HEPES (pH 7.5–8.0), and 10 mM DTT] and incubating the mixture at room temperature for 3–5 days. Bipyramidal crystals in space group  $P3_121$  appeared in 1–5 days and were cryoprotected by serial transfer through solutions of sodium malonate (pH 7.0) and 5 mM DTT to a final concentration of 3.4 M sodium malonate (30). All DJ-1 crystals were mounted in a nylon loop and cryocooled by immersion into liquid nitrogen.

Diffraction data for orthorhombic wt, E18L, E18D, E18Q, and E18N DJ-1 were collected at the Advanced Photon Source (APS), BioCARS beamline 14 BM-C, using 0.9 Å incident X-rays and an ADSC Q315 detector. Single crystals maintained at 100 K were used for the collection of each data set, and the data for E18Q, E18L, and E18D DJ-1 were collected in two passes (a separate high-resolution and a low-resolution pass) with differing exposure times and detector distances to prevent overloaded pixels for intense low-resolution reflections. Because C106 is sensitive to radiation-induced damage, the beam intensity was attenuated and the crystal was exposed to X-rays for ≤10 s per 1° oscillation. Nevertheless, negative  $mF_o - DF_c$  difference electron density around the  $S\gamma$  atom of C106 indicates that some X-ray-induced damage has occurred in the E18D, E18L, and E18Q data sets. Diffraction data were integrated and scaled using HKL2000 (31), and final data statistics are provided in Table 1.

**Crystal Structure Refinement.** The structure of orthorhombic wtDJ-1 was determined by molecular replacement with human DJ-1 (PDB entry 1P5F) (25) as a search model in PHASER (32) as implemented in the CCP4 suite of programs (33). The structures of E18L, E18Q, E18D, and E18N DJ-1 in space group  $P3_121$  were refined using human DJ-1 (1P5F) as the starting model (25). Diffraction data that extend to atomic resolution ( $d_{\min} \leq 1.2 \text{ \AA}$ ) were collected for orthorhombic wt, E18L, and E18D DJ-1, while the data for E18Q and E18N DJ-1 were of lower resolution at 1.35 and 1.6 Å, respectively. SHELX-97 was used for the restrained least-squares refinement of the atomic-resolution structures against a weighted intensity-based residual target function (34). The refinements for E18Q and E18N DJ-1 were performed using REFMAC5 against an amplitude-based maximum likelihood target function (35). All refinements included both stereochemical and displacement parameter restraints and excluded

Table 1: Crystallographic Data and Model Statistics

	wild type	E18L	E18Q	E18D	E18N
Data Statistics					
wavelength (Å)	0.900	0.900	0.900	0.900	0.900
space group	$P2_12_12_1$	$P3_121$	$P3_121$	$P3_121$	$P3_121$
unit cell					
$a$ (Å)	43.91	74.70	74.78	74.83	75.03
$b$ (Å)	85.77	74.70	74.78	74.83	75.03
$c$ (Å)	98.83	74.88	75.44	74.78	75.05
resolution limits (Å)	23–1.20	25–1.15	30–1.35	24–1.20	30–1.60
no. of unique reflections	114729	85989	53400	75880	32697
completeness (%) <sup>a</sup>	97.8 (96.0)	100 (100)	98.6 (98.7)	99.9 (100)	99.8 (100)
mean redundancy <sup>a</sup>	5.3 (4.2)	10.9 (8.0)	7.1 (4.8)	11.1 (7.6)	11.2 (11.2)
$R_{\text{merge}}$ (%) <sup>a,b</sup>	6.5 (70.2)	6.8 (36.5)	7.3 (60.7)	6.9 (69.5)	7.0 (57.4)
$\langle I \rangle / \langle \sigma(I) \rangle^a$	24.3 (2.0)	38.3 (6.0)	25.2 (2.3)	35.5 (2.9)	33.1 (5.1)
Model Statistics					
no. of molecules in ASU	2	1	1	1	1
no. of modeled residues	374	187	186	186	186
no. of solvent molecules	583	239	193	221	182
no. of residues in dual conformations	24	13	6	10	7
heteroatoms	two sulfate	one PEG400	none	one malonate	none
$R_{\text{work}} (\%)^c$	13.1	11.5	15.3	13.1	15.8
$R_{\text{work}} (\%)$ for $F_o > 4\sigma(F_o)^{c,d}$	10.9	10.7	nd	11.6	nd
$R_{\text{free}} (\%)^e$	18.2	14.0	16.8	16.4	18.4
$R_{\text{free}} (\%)$ for $F_o > 4\sigma(F_o)^{d,e}$	15.4	13.0	nd	14.8	nd
$R_{\text{all}} (\%)^f$	13.4	11.5	15.3	13.1	16.0
$R_{\text{all}} (\%)$ for $F_o > 4\sigma(F_o)^d$	10.7	10.7	nd	12.9	nd
mean protein $B_{\text{eq}}$ (Å <sup>2</sup> )	13.5	15.0	16.4	15.9	17.4
mean protein anisotropy <sup>g</sup>	0.40	0.52	1.00	0.46	1.00
mean solvent $B_{\text{eq}}$ (Å <sup>2</sup> )	33.3	33.6	32.6	35.5	30.0
mean solvent anisotropy <sup>g</sup>	0.36	0.39	1.00	0.39	1.00
rmsd for angle lengths (Å)	0.013	0.015	0.013	0.014	0.007
rmsd for angle lengths (deg) <sup>d</sup>	0.031	0.031	nd	0.030	nd
rmsd for bond angles (deg) <sup>h</sup>	nd	nd	1.37	nd	1.06
rmsd for chiral volumes (Å <sup>3</sup> ) <sup>d</sup>	0.077	0.096	0.088	0.091	0.095

<sup>a</sup> Values in parentheses indicate the statistics in the highest-resolution shells. For the wild-type and E18D DJ-1 data sets, this range is 1.24–1.20 Å. For the E18L DJ-1 data set, the highest-resolution shell is from 1.19 to 1.15 Å. For E18Q DJ-1, it is from 1.40 to 1.35 Å, and for E18N DJ-1, it is from 1.66 to 1.60 Å. <sup>b</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_{hkl}| - \langle I_{hkl} \rangle / \sum_{hkl} \sum_i |I_{hkl}|$ , where  $i$  is the  $i$ th observation of a reflection with indices  $h$ ,  $k$ ,  $l$  and angle brackets indicate the average over all  $i$  observations. <sup>c</sup>  $R_{\text{work}} = \sum_{hkl} |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{cal}}| / \sum_{hkl} F_{hkl}^{\text{obs}}$ , where  $F_{hkl}^{\text{cal}}$  is the calculated structure factor amplitude with index  $h$ ,  $k$ ,  $l$ . <sup>d</sup> A statistic that is reported for structures refined in SHELXL; nd means not determined. <sup>e</sup>  $R_{\text{free}}$  is calculated as  $R_{\text{work}}$ , where the  $F_{hkl}^{\text{obs}}$  values are taken from a test set comprising 5% of the data that were excluded from the refinement. <sup>f</sup>  $R_{\text{all}}$  is calculated as  $R_{\text{work}}$ , where the  $F_{hkl}^{\text{obs}}$  values include all measured data (including the  $R_{\text{free}}$  test set). <sup>g</sup> Anisotropy is calculated as the ratio of the smallest to the largest eigenvalues of the refined ADP tensor and is calculated using PARVATI. <sup>h</sup> A statistic that is reported for structures refined in REFMAC5; nd means not determined.

a test set of 5% of randomly chosen reflections that were sequestered and used for the calculation of the  $R_{\text{free}}$  value (36). Bulk solvent models were used to allow inclusion of all measured data (excluding the test set) in all refinements, and anisotropic scaling was used in the REFMAC5 refinements.

Manual adjustments to the models were guided by inspection of  $2mF_o - DF_c$  and  $mF_o - DF_c$  electron density maps in COOT (37), followed by additional cycles of refinement. For orthorhombic wt, E18L, and E18D DJ-1, anisotropic atomic displacement parameters (ADPs) were refined, resulting in an approximately 5% decrease in both  $R$  and  $R_{\text{free}}$  compared to the isotropic displacement parameter model. Due to the lower resolution of the E18Q and E18N DJ-1 data sets, only the isotropic displacement parameter models were used for these structures. For all refinements, hydrogen atoms were introduced in the riding positions on all atoms for which there is an unambiguous assignment of hydrogen geometry.

Final model quality was assessed with PROCHECK (38), MolProbity (39), and the validation tools in COOT (37). In every DJ-1 structure, the backbone torsion angles for C106 were in the unfavorable region of Ramachandran space. This cysteine residue is always found in marginal or unfavorable regions in the Ramachandran plots of DJ-1 superfamily

proteins (40). The refined ADP models for orthorhombic wt, E18L, and E18D DJ-1 were analyzed using the PARVATI server (41). Final model statistics are listed in Table 1.

**Atomic-Resolution Bond Length Analysis.** The protonation state of E18 was determined using an analysis of the changes in the bond length between heavier atoms that occurs upon protonation of one of the atoms (42, 43). To allow determination of the bond length for E18 that was minimally biased by crystallographic restraints, the refined 1.2 Å resolution structure of orthorhombic wtDJ-1 was subjected to 10 additional cycles of conjugate gradient least-squares refinement in SHELX-97 (34) after removal of all geometric (DFIX, DANG, and CHIV) restraints for E18. The resulting model was then subjected to a single cycle of blocked (BLOC 1), unrestrained full matrix least-squares refinement and inversion of the least-squares matrix to determine the estimated standard uncertainties (esu) of the bond lengths in E18. We note that the final model deposited in the PDB (entry 2OR3) was refined with bond length and angle restraints; the unrestrained refinement was used only to determine the protonation state of E18 during analysis of the structure.

**Spectrophotometric Cysteine  $pK_a$  Determination.** Samples of tag-cleaved DJ-1 in storage buffer (25 mM HEPES, 100 mM KCl, and 2 mM DTT) were centrifugally desalinated using

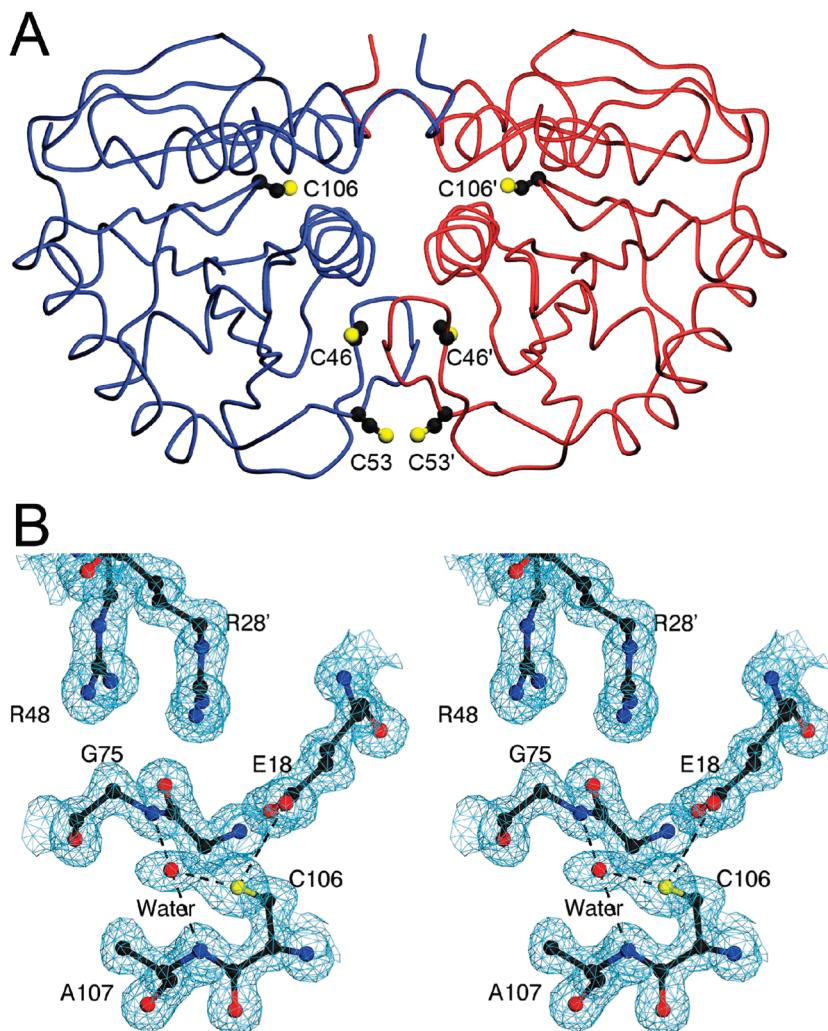


FIGURE 1: Cysteine residues in human DJ-1. In panel A, the DJ-1 dimer is represented with one monomer colored blue and the other red. The three cysteine residues in DJ-1 are shown, and symmetry-related residues are labeled with a prime. Both C106 and C53 are solvent-exposed, while C46 is buried. In panel B, the residues composing the environment of C106 are shown in stereoview with  $2mF_o - DF_c$  electron density contoured at  $1\sigma$  (blue). The thiol of C106 makes two direct hydrogen bonds with surrounding atoms (dashed): one to E18 and the other to an ordered water molecule. The ordered water is bound in the G75/A107 "amide pocket" and is stabilized by two hydrogen bonds with the peptide backbone. This bound water molecule is not always seen in crystal structures of DJ-1 and may initiate the oxidation of C106. This figure was created with POVscript+ (61).

PD-6 resin (Bio-Rad) that had been equilibrated with freshly degassed water. Solutions with pH values in the range of 3.3–7.4 were prepared using an equimolar sodium citrate and sodium phosphate double buffer system adjusted to the desired pH by addition of small volumes (0.5–1  $\mu$ L) of 500 mM NaOH. The double buffer was used at a concentration of 10 mM for each buffer component, and the pH of each freshly prepared buffer used in the experiment was measured at its working concentration using a micro pH electrode after the addition of 20–30  $\mu$ M DJ-1. Titrations of wtDJ-1 in 10 mM double buffer supplemented with 100 mM NaCl showed no effect of increased ionic strength on the  $pK_a$  of C106 in wtDJ-1 (Supporting Information, Figure S1).

Ionization of the cysteine thiol was monitored by absorption of the thiolate anion at 240 nm (44, 45) using a Cary 50 UV-visible spectrometer (Varian) and a 1 cm path length quartz cuvette. Freshly desalted DJ-1 in water was diluted 10–15-fold (to a final concentration of 20–30  $\mu$ M DJ-1) into 10 mM double buffer, and the absorption of the sample at 240 and 280 nm was measured after correction for the absorption of the buffer alone. The extinction coefficient at

240 nm ( $\epsilon_{240}$ ) was calculated using the ratio of absorbance at 280 and 240 nm with the following equation:

$$\epsilon_{240} = \epsilon_{280} \frac{A_{240}}{A_{280}} \quad (1)$$

where  $A_{240}/A_{280}$  is the ratio of the absorbance of the protein at 240 and 280 nm,  $\epsilon_{280}$  is the known extinction coefficient of DJ-1 at 280 nm ( $4000 \text{ M}^{-1} \text{ cm}^{-1}$ ), and  $\epsilon_{240}$  is the extinction coefficient at 240 nm. Using the  $A_{240}/A_{280}$  ratio normalizes all measurements to the amount of protein present in the cuvette. We note that this method assumes that the  $\epsilon_{280}$  value for DJ-1 is not pH-dependent, which is justified by the behavior of C106S DJ-1 (Figure 2A).

The titrations for tag-cleaved wild-type and C106S DJ-1 were performed in triplicate, and average  $\epsilon_{240}$  values with associated standard deviations from three trials for each protein sample were calculated. The data were plotted as a function of pH using SigmaPlot (Systat Software), and the  $pK_a$  was determined by fitting a version of the Henderson–Hasselbalch equation to the data:

$$\varepsilon_{240}^O(\text{pH}) = \varepsilon_{240}^{\text{SH}} + \frac{\varepsilon_{240}^{S^-} - \varepsilon_{240}^{\text{SH}}}{1 + 10^{pK_a - \text{pH}}} \quad (2)$$

where  $\varepsilon_{240}^O(\text{pH})$  is the observed extinction coefficient of the protein at 240 nm as a function of pH,  $\varepsilon_{240}^{\text{SH}}$  is the extinction coefficient of the protein at 240 nm when the cysteine residue is in the thiol form,  $\varepsilon_{240}^{S^-}$  is the extinction coefficient of the protein at 240 nm when the cysteine residue is in the thiolate form, and pH is the measured pH of the buffer. Nonlinear regression was performed using SigmaPlot, and the reported errors for the  $pK_a$  values are those obtained from the fit procedure.

**Search of the PDB for Aspartic/Glutamic Acid–Cysteine Interactions.** A list of all nonredundant X-ray crystal structures determined at a resolution of 3.0 Å or better was obtained by searching the April 15, 2008, release of the PDB using the RCSB search tool (46). Nonredundant structures were defined as those that are <90% identical in sequence with other deposited structures, and the resulting list (11765 entries) included only the highest-resolution example of a set of similar proteins. We searched this set of structures for all instances in which one of the carboxylate oxygen atoms from aspartic ( $O\delta 1$  and  $O\delta 2$ ) or glutamic ( $O\epsilon 1$  and  $O\epsilon 2$ ) acid occurred within 4.0 Å of the  $S\gamma$  atom of a cysteine residue using a custom-written search routine. Additional geometric selection criteria were applied to select for candidate hydrogen bonds between D/E and C by including the hydrogen bond angle and a “virtual” dihedral angle that includes the hydrogen bond. The hydrogen bond angle was defined as the angle spanned by the  $C(\delta/\gamma)$ ,  $O(\delta/\epsilon;1/2)$ , and  $S\gamma$  atoms, where the  $C(\delta/\gamma)$  and  $O(\delta/\epsilon;1/2)$  atoms belong to the carboxylic acid group of glutamic or aspartic acid. The virtual dihedral angle was defined as the dihedral spanned by the  $O(\delta/\epsilon;1/2)$ ,  $C(\delta/\gamma)$ ,  $O(\delta/\epsilon;2/1)$ , and  $S\gamma$  atoms, where the first three atoms belong to the carboxylic acid group of aspartic or glutamic acid. The target hydrogen bond angle was  $109.5 \pm 15^\circ$ , and the target dihedral angle values were  $0 \pm 20^\circ$  and  $180 \pm 20^\circ$ , corresponding to the cis and trans orientations, respectively, of this dihedral. For each set of values, the latter number indicates the accepted deviation from the target. By convention, clockwise rotation is a positive displacement for the dihedral angle.

## RESULTS

**C106 Has a Depressed  $pK_a$ .** Human DJ-1 contains three cysteine residues (C46, C53, and C106), two of which (C106 and C53) are solvent-exposed and thus would be more likely to form reactive thiols (Figure 1A). The most oxidation sensitive cysteine in human DJ-1 is C106 (5, 47), and its environment is rich in ionizable residues that could perturb the  $pK_a$  of this functionally essential residue (Figure 1B). The  $pK_a$  of C106 was monitored by measuring the increased absorbance of UV light at 240 nm resulting from formation of the thiolate anion (44, 45). The pH dependence of the molar extinction coefficient of DJ-1 at 240 nm ( $\varepsilon_{240}$ ) displays a transition with a  $pK_a$  of  $5.4 \pm 0.1$  and a total increase of  $3500 \text{ M}^{-1} \text{ cm}^{-1}$  (Figure 2A), consistent with the reported extinction coefficient of  $4000 \text{ M}^{-1} \text{ cm}^{-1}$  for a single thiolate at 240 nm (44). This increase in  $\varepsilon_{240}$  is due solely to ionization of C106, as demonstrated by the absence of a pH-dependent transition for C106S DJ-1 (Figure 2A). The  $pK_a$

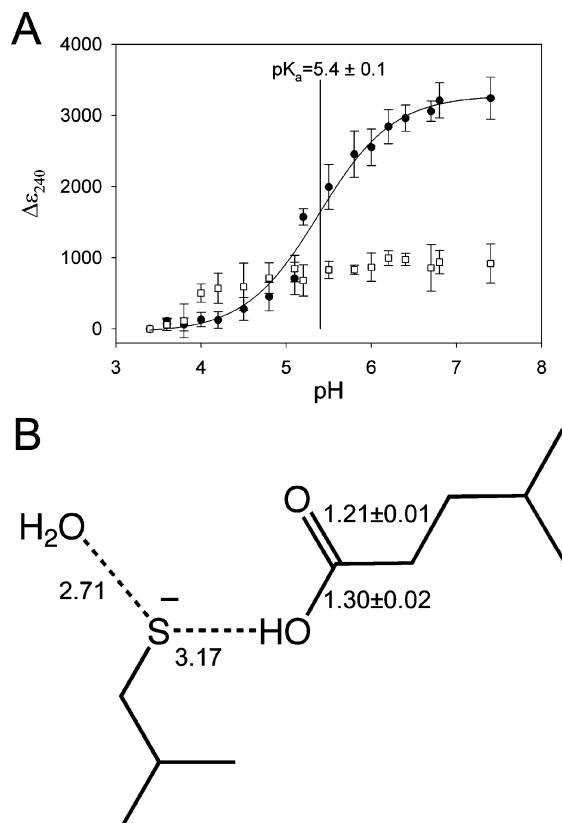


FIGURE 2: C106 has a depressed  $pK_a$  and accepts a hydrogen bond from a protonated E18. In panel A, the change in the extinction coefficient at 240 nm ( $\Delta\varepsilon_{240}$ ) is measured as a function of pH to monitor formation of the cysteine thiolate anion. The  $\Delta\varepsilon_{240}$  value is calculated by subtraction of the  $\varepsilon_{240}$  measured at the lowest pH from all other measured  $\varepsilon_{240}$  values. Wild-type DJ-1 (●) exhibits a  $\Delta\varepsilon_{240}$  transition corresponding to a  $pK_a$  of  $5.4 \pm 0.1$ , where mean data and associated standard deviations for triplicate measurements are shown. The curve shows the nonlinear regression fit of eq 2 to the measured data. C106S DJ-1 (□) shows no transition, indicating that ionization of C106 is solely responsible for the observed transition in wtDJ-1. In panel B, unrestrained bond length refinement of a 1.2 Å resolution crystal structure of wtDJ-1 indicates that E18 is constitutively protonated and donates a hydrogen bond to C106. The refined bond lengths and estimated standard uncertainties for the carbon–oxygen bonds of E18 are shown, and hydrogen bonds are represented with dashed lines. Distances are given in angstroms, and the values for molecule B in the DJ-1 dimer are shown. Panel B was created with ChemDraw (CambridgeSoft).

of C106 in unchanged when measured in 10 mM double buffer supplemented with 100 mM NaCl (Supporting Information, Figure S1), demonstrating that the  $pK_a$  of C106 is  $5.4 \pm 0.1$  in solutions at physiologically relevant ionic strengths.

The depression of the  $pK_a$  value of C106 from the solution value of 8.3 to the measured value of 5.4 corresponds to a  $-4.0 \text{ kcal mol}^{-1}$  stabilization of the thiolate ion by the local environment of C106 in DJ-1, calculated as follows:

$$\Delta\Delta G_{\text{ion}}^0 = \Delta G^0(\text{protein}) - \Delta G^0(\text{solution}) \\ = RT \ln[10^{pK_a(\text{protein}) - pK_a(\text{solution})}] \quad (3)$$

where  $\Delta\Delta G_{\text{ion}}^0$  is the difference in the Gibbs free energy of ionization of the cysteine thiol in the protein environment compared with solution and is a measure of the degree of thiolate stabilization by the protein environment. In eq 3,  $\Delta G^0(\text{protein})$  is the free energy of ionization of the cysteine residue in the protein environment,  $\Delta G^0(\text{solution})$  is the free

energy of ionization of the cysteine residue in solution,  $pK_a(\text{protein})$  is the  $pK_a$  of the thiol in the protein environment,  $pK_a(\text{solution})$  is the  $pK_a$  of the thiol in solution,  $R$  is the gas constant, and  $T$  is the temperature (kelvin). In the environment of C106, the carboxylic acid side chain of E18 is the only residue that is oriented properly to hydrogen bond to the thiol of C106 (Figure 1B). Both of these residues are very highly conserved in the DJ-1 superfamily (19, 20), suggesting that the E18–C106 interaction is likely to be functionally important.

We note that an ordered water molecule bound in the G75/A107 amide pocket is apparent in this new crystal structure of DJ-1 and is within sub-van der Waals distance (2.5 Å in molecule A, 2.7 Å in molecule B) of the  $S\gamma$  atom of C106 (Figures 1B and 5). This sulfur–oxygen distance is too great to correspond to a cysteine sulfenic acid but too short to be a normal hydrogen bond between these atoms. We propose that this density is most consistent with a mixture of species, including the bound water (or other oxygen-containing species) and a minor contribution from C106-sulfenic acid. Therefore, we speculate that this may represent a partially oxidized form of C106 that contains a mixture of stable intermediate species. This is consistent with a previously proposed oxidation sequence for the reactive cysteine residue in the *E. coli* DJ-1 homologue YajL (40).

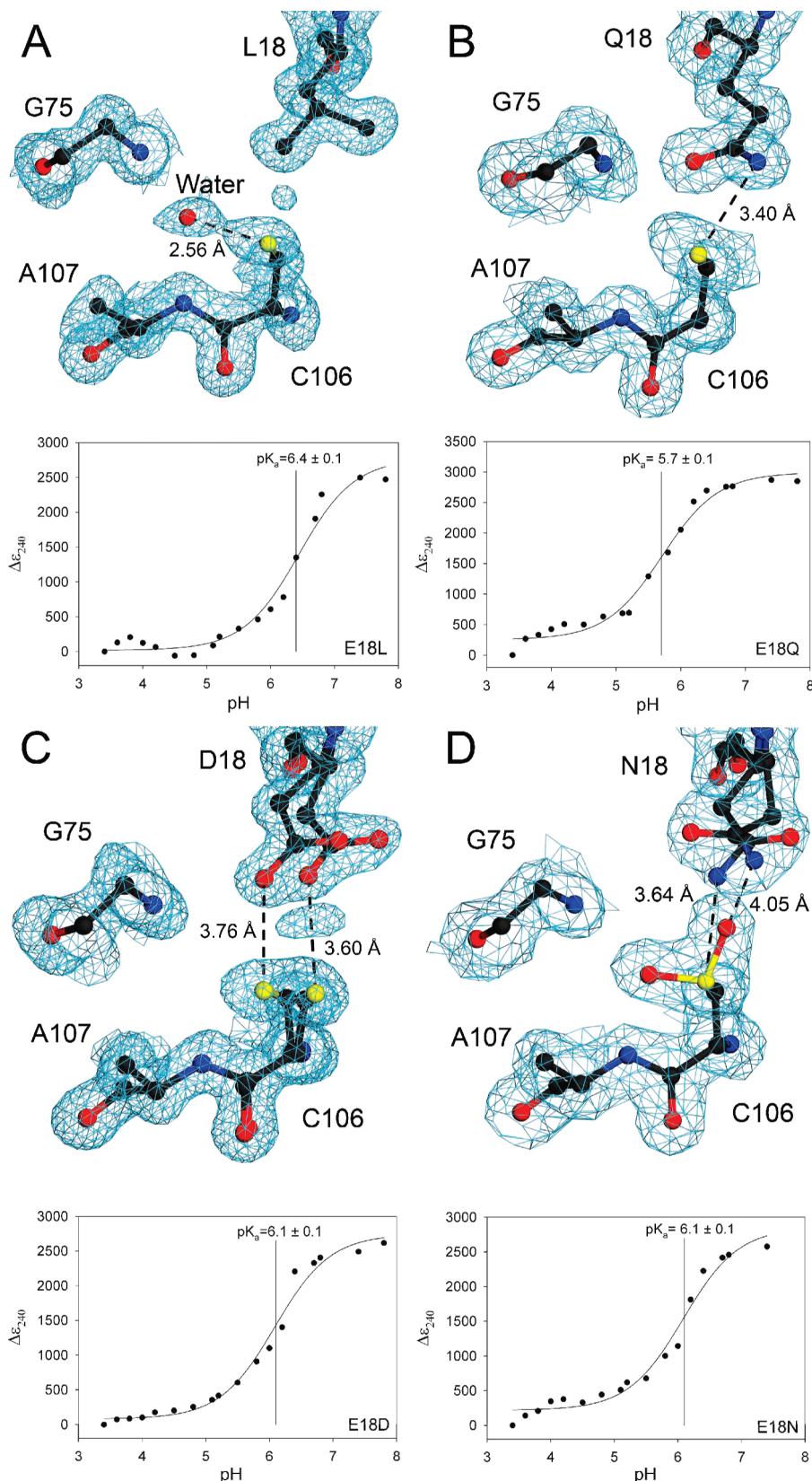
**E18 Is Protonated and Donates a Hydrogen Bond to C106.** Formation of a cysteine thiolate anion is electrostatically disfavored by nearby negative charges. It is therefore surprising that C106 has both a depressed  $pK_a$  and makes a 3.2 Å hydrogen bond with E18 (Figure 1B). Bond length analysis of a new crystal structure of orthorhombic wtDJ-1 refined at 1.2 Å resolution shows that E18 is protonated (and thus uncharged) and donates a hydrogen bond to C106. The orthorhombic crystal form of DJ-1 has both molecules of the physiological DJ-1 dimer in the asymmetric unit, providing two independent views of E18. Unrestrained refinement of E18 bond lengths in SHELX-97 (see Experimental Procedures) yields  $C\delta$ – $O\epsilon 1$  bond lengths of 1.241 Å (esu = 0.021 Å) for molecule A and 1.213 Å (esu = 0.012 Å) for molecule B. The  $C\delta$ – $O\epsilon 2$  bond lengths are 1.296 Å (esu = 0.021 Å) for molecule A and 1.304 Å (esu = 0.020 Å) for molecule B, demonstrating that the  $O\epsilon 2$  atom of E18 is protonated (Figure 2B). These bond lengths compare well with the bond lengths for a fully protonated (uncharged) carboxylic acid, which are 1.21 Å for a carbon–oxygen double bond and 1.30 Å for a carbon–oxygen single bond from the Engh and Huber parameter set that is widely used in macromolecular refinement (48). The crystals were grown at pH 7.5, suggesting that E18 is constitutively protonated at physiological pH.

**Donation of a Hydrogen Bond from E18 Strongly Influences the C106  $pK_a$ .** The influence of hydrogen bond donation by E18 on the  $pK_a$  of C106 was determined using site-directed mutagenesis and spectrophotometric  $pK_a$  determination. The E18–C106 hydrogen bond was eliminated by creation of an E18L mutation, which elevates the  $pK_a$  of C106 to  $6.4 \pm 0.1$  (Table 2). The elevation of the  $pK_a$  of C106 by 1.0 unit in E18L DJ-1 corresponds to a loss of 1.4 kcal/mol of stabilization energy for the C106 thiolate compared to wtDJ-1 [ $1.4 \text{ kcal/mol} = \Delta\Delta G_{\text{ion}}^0(\text{E18L}) - \Delta\Delta G_{\text{ion}}^0(\text{wt})$ ; see eq 3 and Table 2], approximately one-third of the total stabilization of C106 ionization by DJ-1 relative

to free cysteine in solution. The 1.15 Å resolution crystal structure of E18L DJ-1 shows that no major differences in the overall structure of DJ-1 or the local conformation of C106 occur in response to the E18L substitution (Figure 3A). Therefore, we conclude that the elevation of the C106  $pK_a$  by 1.0 unit in E18L DJ-1 is due principally to the loss of a single hydrogen bond donated by E18.

To improve our understanding of the influence of E18 on the  $pK_a$  of C106, we altered the strength of the hydrogen bond donated by residue 18 by generating E18Q, E18D, and E18N substitutions in DJ-1. Each of these engineered mutations at residue 18 elevates the  $pK_a$  of C106 (Table 2). The most structurally conservative substitution, E18Q, increases the  $pK_a$  by 0.3 unit and lengthens the hydrogen bond to  $S\gamma$  by 0.2 Å (Figure 3B). The electron density for C106 in E18Q DJ-1 is elongated at the  $S\gamma$  atom, consistent either with conformational disorder or minor oxidation of the cysteine residue. In contrast to the well-ordered Q18 side chain in E18Q DJ-1, the shorter side chains of the E18D and E18N substitutions result in discrete conformational disorder and increase the length of the hydrogen bond to C106 by approximately 0.4 Å compared to that of wtDJ-1 (Figure 3C,D). In addition, C106 is invariably oxidized to C106-sulfinic acid in crystals of E18N DJ-1 (Figure 3D), further lengthening the  $N\delta 2$ – $S\gamma$  distance. We note that, in contrast to the oxidized crystalline protein, the E18N DJ-1 samples used for solution  $pK_a$  determination are approximately 85% reduced as determined by electrospray mass spectrometry (University of Nebraska Mass Spectrometry Core). Oxidation of C106 to C106-sulfinic acid is well-established in DJ-1 and can occur during crystallization of an initially reduced protein (5). This combined structural and C106  $pK_a$  comparison of the E18Q, E18N, and E18D substitutions shows that even minor lengthening of the hydrogen bond donated by residue 18 significantly destabilizes the C106 thiolate, emphasizing the critical role of the E18–C106 interaction in DJ-1.

**Interactions between Acidic Residues and Cysteine in the PDB.** The set of 11765 nonredundant X-ray crystal structures determined at 3.0 Å resolution or better was extracted from the April 15, 2008, release of the PDB and searched for every instance in which one of the carboxylate oxygen atoms of aspartic or glutamic acid was within 4.0 Å of the cysteine  $S\gamma$  atom (see Experimental Procedures). This is a permissive screen for candidate D/E–C interactions and is likely to result in many false positives. We identified 6477 instances of proximal D/E–C pairs in 2349 unique structures, or roughly 20% of the models in the search set. With this set, we performed a more stringent search for all examples in which the hydrogen bond angle defined by the  $C(\delta/\gamma)$ ,  $O(\delta/\epsilon;1/2)$ , and  $S\gamma$  atoms was within  $109.5 \pm 15^\circ$  and the  $O(\delta/\epsilon;1/2)$ – $C(\delta/\gamma)$ – $O(\delta/\epsilon;1/2)$ – $S\gamma$  dihedral angle was either  $0 \pm 20^\circ$  or  $180 \pm 20^\circ$  (Figure 4A,B). These geometric criteria have been previously established for hydrogen bonds involving protonated aspartic or glutamic acids, with the  $0^\circ$  and  $180^\circ$  targets for the dihedral angle corresponding to the favored cis and trans orientations (49). For members of the DJ-1 superfamily, this dihedral is near the cis orientation. This results in 235 D/E–C pairs occurring in 149 unique structures (Supporting Information Table S2), or approximately 1.3% of the original search set. All DJ-1 superfamily members of known structure that contain a putative E–C



**FIGURE 3:** Donation of a hydrogen bond by residue 18 contributes to the depression of the  $pK_a$  of C106. In all panels,  $2mF_o - DF_c$  electron density is contoured at  $1\sigma$  (blue) and the pH dependence of the  $\Delta\epsilon_{240}$  value is shown in the bottom portion of each panel. The  $\Delta\epsilon_{240}$  value is calculated by subtraction of the  $\epsilon_{240}$  measured at the lowest pH from all other measured  $\epsilon_{240}$  values. For the  $\Delta\epsilon_{240}$  vs pH plots, measured data are shown with the best-fit curve from eq 2. Panel A shows that eliminating the hydrogen bond between residue 18 and C106 with an E18L substitution elevates the C106  $pK_a$  value by 1.0 unit to 6.4. Panel B shows that a structurally conservative E18Q substitution lengthens the hydrogen bond to S $\gamma$  of C106 by 0.2 Å and increases the C106  $pK_a$  value by 0.3 unit. Panel C shows that the E18D substitution results in spatially correlated discrete disorder at D18 and C106, lengthens the hydrogen bond by approximately 0.5 Å compared to that of wtDJ-1, and increases the C106  $pK_a$  value by 0.7 unit. Panel D shows that the E18N substitution leads to discrete disorder at N18, enhances oxidation of C106, and increases the C106  $pK_a$  value by 0.7 unit. All of the substitutions at residue 18 illustrate the requirement for a protonated glutamic acid at position 18 for maximal depression of the  $pK_a$  of C106. This figure was made with POVscript+ (61).

Table 2: C106 pK<sub>a</sub> Values and Thiolate Stabilization Free Energies in DJ-1

protein	C106 pK <sub>a</sub>	$\Delta\Delta G_{\text{ion}}^0$ (kcal/mol) <sup>a</sup>
wt	5.4 ± 0.1	-4.0
E18L	6.4 ± 0.1	-2.6
E18Q	5.7 ± 0.1	-3.5
E18D	6.1 ± 0.1	-3.0
E18N	6.1 ± 0.2	-3.0
R28Q	5.4 ± 0.1	-4.0
R48Q	5.2 ± 0.1	-4.2
R28Q/R48Q	5.0 ± 0.1	-4.5

<sup>a</sup>  $\Delta\Delta G_{\text{ion}}^0$  is the difference Gibbs free energy for ionization of the cysteine in the protein environment and free cysteine in solution and is calculated using eq 3.

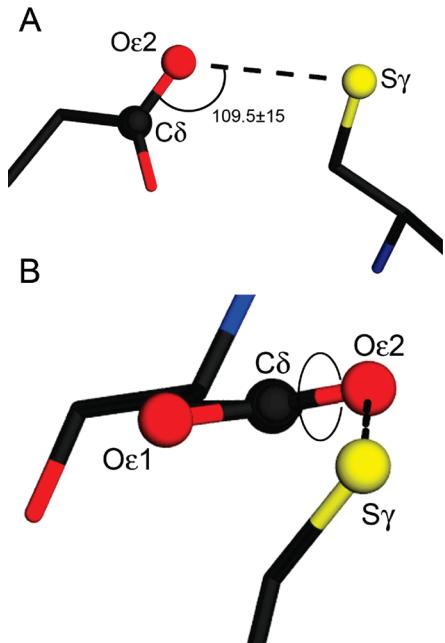


FIGURE 4: Geometric parameters for carboxylic acid–cysteine hydrogen bonding. The E18–C106 interaction is shown in both panels, with the hydrogen bond represented by the dotted line. Atoms involved in the definition of the hydrogen bond angle and dihedral angle are shown as balls and labeled. Panel A shows the definition of the hydrogen bond angle, which is set to a target value of  $109.5 \pm 15^\circ$  in the PDB search. Panel B shows the convention for defining the hydrogen bond dihedral angle, with the rotatable bond indicated by the circle. This dihedral angle has a value of  $15^\circ$  in human DJ-1, and target values of  $0 \pm 20^\circ$  (cis) and  $180 \pm 20^\circ$  (trans) were used in the PDB search. This figure was made with POVscript+ (61).

dyad are represented in this list except *E. coli* Hsp31 [PDB entry 1N57 (50)], whose  $25.7^\circ$  H-bond dihedral angle only slightly exceeds our accepted deviation tolerance of  $20^\circ$ . Moreover, the E18–C106 interaction is correctly identified for human DJ-1 [PDB entry 2RK3 (51)]. In total, the relatively small fraction of structures that returned a positive hit in our search suggests that hydrogen bonded D/E–C interactions are uncommon among all structures in the PDB, but common within the DJ-1 superfamily.

**R28 and R48 Bind Anions That Elevate the pK<sub>a</sub> of C106.** DJ-1 contains two highly conserved arginine residues (R28' and R48) that participate in an unusual guanidinium stacking interaction that spans the dimer interface and is located near C106 (Figure 5). The prime notation (e.g., R28') indicates a residue that is donated by the other monomer in the DJ-1 dimer. The charge state of these two residues is ambiguous.

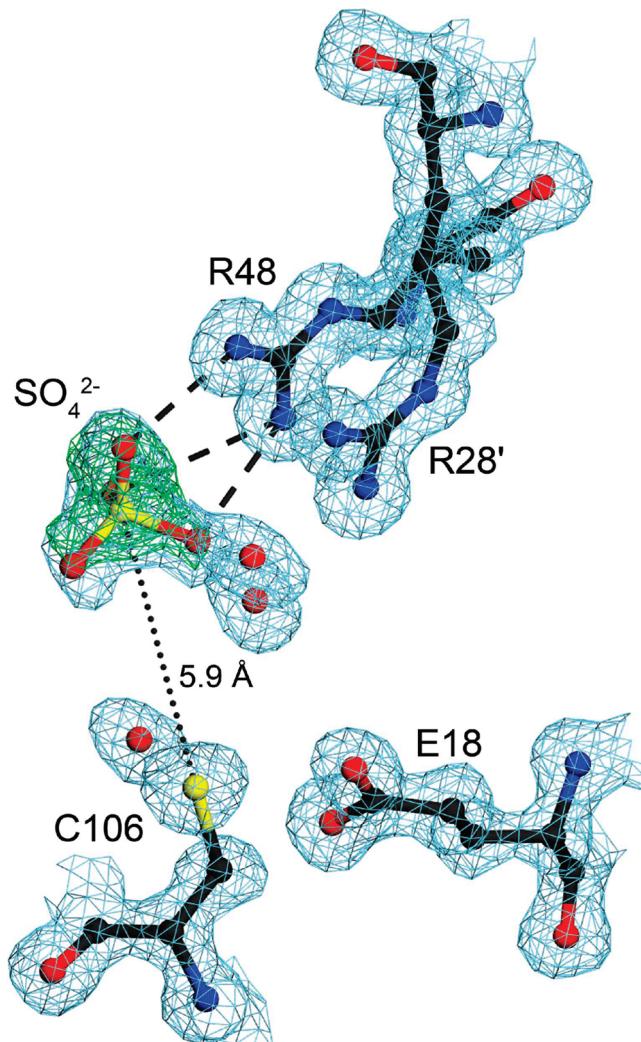


FIGURE 5: Two proximal arginine residues bind an anion near C106. The environment of the R28'/R48 dyad is shown with  $2mF_o - DF_c$  electron density contoured at  $1\sigma$  (blue) and  $mF_o - DF_c$  electron density contoured at  $4\sigma$  (green). R28 and R48 participate in an unusual guanidinium stacking interaction that spans the DJ-1 dimer interface, with R28' contributed by the other monomer in this view (indicated with a prime). A bound sulfate ion interacts with R48 through three direct hydrogen bonds to the guanidinium side chain, illustrated as dashed lines. Positive difference electron density contoured at  $4\sigma$  (green) was calculated prior to the inclusion of the sulfate ion in the model. The R28'/R48 anion binding site places a negative charge within  $5.9 \text{ \AA}$  of C106. The ordered water bound near C106 in the G75/A107 amide pocket likely initiates oxidation of this residue. This figure was made with POVscript+ (61).

The high solution pK<sub>a</sub> of the arginine side chain (12) would suggest that they are positively charged; however, the atypical  $3.6 \text{ \AA}$  stacking of the guanidinium groups of R28' and R48 is electrostatically unfavorable for two positively charged groups (Figure 5). Despite the uncertain charge state of these residues, R28' and R48 compose an anion binding site that is occupied by sulfate in both molecules in the ASU of the orthorhombic crystal structure of DJ-1 (Figure 5). The sulfur atom of the sulfate anion is  $5.9 \text{ \AA}$  from S<sub>γ</sub> of C106, and this proximal negatively charged group is expected to suppress ionization of the C106 thiol.

The bound sulfate anion makes direct hydrogen bonds only to R48, and mutation of R28' to glutamine does not affect the pK<sub>a</sub> of C106 under these experimental conditions (Figure 6A). The R48Q substitution results in a 0.2 unit decrease in

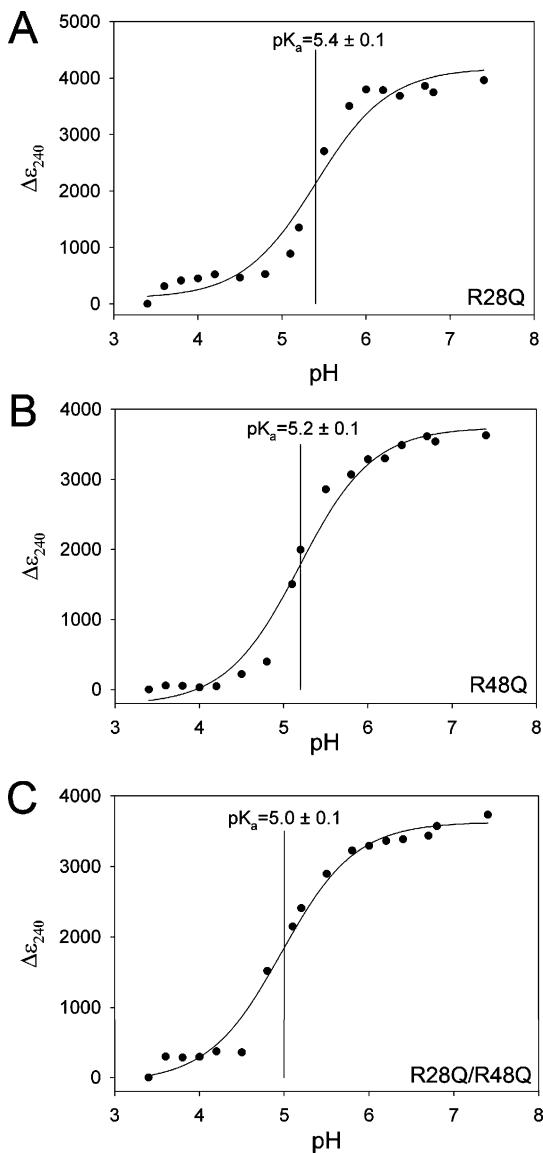


FIGURE 6: R28/R48 anion binding site elevates the  $pK_a$  of C106. Substitution of glutamine for R28 (A) results in no change in the  $pK_a$  of C106; a R48Q substitution reduces the C106  $pK_a$  by 0.2 unit, and the R28Q/R48Q double substitution decreases the C106  $pK_a$  by 0.4 unit. Measured data are shown with the best-fit curve from eq 2. Both the R48Q and double R28Q/R48Q substitutions decrease the C106  $pK_a$  value, indicating that the bound anion observed in the crystal structure suppresses the ionization of C106 in wtDJ-1.

C106  $pK_a$ , while the double R28Q/R48Q substitution results in a larger 0.4 unit C106  $pK_a$  decrease. The depression of the C106  $pK_a$  value upon removal of two nearby basic residues is surprising and strongly suggests that the R48Q and R28Q/R48Q engineered substitutions disrupt the anion binding site near C106. The loss of this bound anion alleviates electrostatic suppression of thiolate formation, thereby decreasing the  $pK_a$  of C106 (Figure 6 and Table 2). The sulfate observed in the crystal structure is not present in the double buffer used in the C106  $pK_a$  titration; thus, the R28/R48 anion binding site must be occupied by other anions in the buffer, most likely phosphate. Intriguingly, the pH-dependent  $\epsilon_{240}$  transition for both R28Q and R48Q is more cooperative than expected for a single ionization obeying the Henderson–Hasselbalch equation (Figure 6A,B), suggesting that the C106  $pK_a$  may be influenced by the ionization

of other nearby groups. A plausible candidate is the E15/E16/D24' cluster of acidic residues, which makes multiple hydrogen bonds to the R48/R28' pair and may be perturbed by the single R to Q mutations being used in this study.

## DISCUSSION

The cytoprotective activity of DJ-1 requires C106 (5, 6), and our results establish that this residue has a low  $pK_a$  of  $5.4 \pm 0.1$ . Site-directed mutation of residue 18 demonstrates that approximately one-third of the total stabilization energy for the ionization of C106 results from the  $S\gamma$  thiolate of C106 accepting a hydrogen bond from the constitutively protonated O $\epsilon 2$  atom of E18. It is noteworthy that the disordered E18D and E18N substitutions increase the C106  $pK_a$  value by comparable amounts (Table 2), despite the different electrostatic and hydrogen bonding character of these two substitutions (Table 2). Furthermore, the E18D and E18N substitutions have  $pK_a$  values near that of the E18L mutant, which eliminates this hydrogen bond entirely. We conclude that the C106–E18 interaction in DJ-1 represents an unusual type of dyad that uses a protonated carboxylic acid moiety to depress the  $pK_a$  of a reactive cysteine residue. Residues that are structurally equivalent to C106 and E18 in human DJ-1 are very highly conserved in other members of the DJ-1 superfamily, indicating that the glutamic acid–cysteine dyad is widely distributed in the superfamily (19, 20). We note that our current data do not preclude the possibility that the E18–C106 dyad represents a general acid–general base pair, although this is highly speculative in the absence of a well-established catalytic activity for DJ-1.

The influence of E18, R28, and R48 on the  $pK_a$  of C106 contradicts expectations based on the typical ionization behavior of these residues in other proteins. In this regard, DJ-1 serves as a cautionary example that illustrates the potential limitations of computational approaches to determining  $pK_a$  values, as a previous report suggested a  $pK_a$  value of 11.38 for C106 based on computational prediction (52). Furthermore, our results show that when the hydrogen bond between E18 and C106 is eliminated by mutagenesis, the  $pK_a$  of C106 is still 1.9 pH units below that of free cysteine. This suggests that additional factors contribute to the depressed  $pK_a$  of C106 in DJ-1, possibly including the strained backbone conformation around this residue or the  $\alpha$ -helix macropole of helix E (3). Future experimental studies of the structural determinants of depressed cysteine  $pK_a$  values in DJ-1 and other systems are expected to aid in the development of improved algorithms for the computational prediction of cysteine  $pK_a$  values.

Cysteine residues with depressed  $pK_a$  values often serve catalytic roles (53–55). The biochemical function of DJ-1 is uncertain, and attempts to detect various enzymatic activities have been inconclusive (15, 25, 29, 52, 56). Most published reports agree, however, that C106 is exquisitely sensitive to oxidation and readily forms the C106-sulfenic acid ( $C106-SO_2^-$ ) under physiological conditions (5, 29, 47, 52, 57). The 1.2 Å resolution crystal structure of oxidized DJ-1 containing C106- $SO_2^-$  shows that one of the two sulfenic acid oxygen atoms of C106- $SO_2^-$  makes an unusually short 2.47 Å hydrogen bond to the O $\epsilon 2$  atom of E18 (5). Therefore, the protonated carboxylic acid of E18 contributes both to rendering C106 reactive by depressing its  $pK_a$  and to the

stabilization of a specific oxidized form of this cysteine residue. The C106-SO<sub>2</sub><sup>-</sup> form of DJ-1 has been implicated in the neuroprotective function of the protein by several distinct mechanisms, including enhancement of mitochondrial DJ-1 localization (5), regulating a redox-regulated chaperone activity of the protein (15, 29) and directly scavenging reactive oxygen species (17, 58). Although further research will be required to determine which, if any, of these activities is critical for DJ-1's cytoprotective function, it is clear that the E18–C106 dyad plays a major role in directing the chemistry that can occur at C106 in DJ-1 and its many homologues.

A survey of structures in the PDB reveals that approximately 20% of the structures in our search list have at least one carboxylic oxygen atom of either glutamic and aspartic acid within 4 Å of the sulfur atom of cysteine. Of this set of possible D/E–C dyads, only 149 structures (1.3% of the original 11765 structures) contain a candidate D/E–C interaction that satisfies more stringent angular and dihedral criteria for hydrogen bonding (see Experimental Procedures). Therefore, candidate D/E–C dyads that are structurally similar to the hydrogen bonded E18–C106 interaction observed in DJ-1 are rare. We note that we have used moderately stringent criteria for identifying candidate D/E–C interactions, and protonated acidic residues can influence cysteine reactivity in diverse structural environments that do not satisfy these criteria. In the well-studied example of *E. coli* thioredoxin, a buried aspartic acid residue (D26) with an elevated pK<sub>a</sub> acts as a general acid/base that facilitates thiol–disulfide exchange involving C32 in the enzyme active site (59, 60). While the E18–C106 interaction in DJ-1 does not appear to be involved in thiol–disulfide exchange, our results suggest that functionally important interactions between carboxylates and thiols are present in several types of proteins. It is difficult to speculate about the frequency with which a protonated acidic residue depresses the pK<sub>a</sub> of cysteine because it is impractical to determine the cysteine pK<sub>a</sub> values and carboxylic acid protonation states of all of the candidate D/E–C interactions listed in Table S2 of the Supporting Information. We propose, however, that this type of interaction is common for residues that are structurally homologous to E18 and C106 in the large DJ-1 superfamily.

## ACKNOWLEDGMENT

We thank the staff of BioCARS 14 BM-C at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) for beamline support, Dr. Hideaki Moriyama for support of the rotating anode X-ray source at the University of Nebraska, and Dr. Todd Holyoak (University of Kansas Medical Center, Kansas City, KS) for helpful discussions.

## SUPPORTING INFORMATION AVAILABLE

Figure S1 shows a titration of wtDJ-1 in a double buffer supplemented with 100 mM NaCl, demonstrating that increased ionic strength has no effect on the pK<sub>a</sub> of C106 under these conditions. Table S2 lists every example of a candidate D/E–C hydrogen bonded dyad in the April 15, 2008, release of the PDB. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Kortemme, T., Darby, N. J., and Creighton, T. E. (1996) Electrostatic interactions in the active site of the N-terminal thioredoxin-like domain of protein disulfide isomerase. *Biochemistry* 35, 14503–14511.
- Jao, S. C., English Ospina, S. M., Berdis, A. J., Starke, D. W., Post, C. B., and Mieyal, J. J. (2006) Computational and mutational analysis of human glutaredoxin (thioltransferase): Probing the molecular basis of the low pKa of cysteine 22 and its role in catalysis. *Biochemistry* 45, 4785–4796.
- Kortemme, T., and Creighton, T. E. (1995) Ionisation of cysteine residues at the termini of model  $\alpha$ -helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J. Mol. Biol.* 253, 799–812.
- Naor, M. M., and Jensen, J. H. (2004) Determinants of cysteine pKa values in creatine kinase and  $\alpha$ 1-antitrypsin. *Proteins* 57, 799–803.
- Canet-Aviles, R. M., Wilson, M. A., Miller, D. W., Ahmad, R., McLendon, C., Bandyopadhyay, S., Baptista, M. J., Ringe, D., Petsko, G. A., and Cookson, M. R. (2004) The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfenic acid-driven mitochondrial localization. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9103–9108.
- Meulener, M. C., Xu, K., Thomson, L., Ischiropoulos, H., and Bonini, N. M. (2006) Mutational analysis of DJ-1 in *Drosophila* implicates functional inactivation by oxidative damage and aging. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12517–12522.
- Bonifati, V., Rizzu, P., van Baren, M. J., Schaap, O., Breedveld, G. J., Krieger, E., Dekker, M. C., Squitieri, F., Ibanez, P., Joosse, M., van Dongen, J. W., Vanacore, N., van Swieten, J. C., Brice, A., Meco, G., van Duijn, C. M., Oostra, B. A., and Heutink, P. (2003) Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science* 299, 256–259.
- Zhang, L., Shimoji, M., Thomas, B., Moore, D. J., Yu, S. W., Marupudi, N. I., Torp, R., Torgner, I. A., Ottersen, O. P., Dawson, T. M., and Dawson, V. L. (2005) Mitochondrial localization of the Parkinson's disease related protein DJ-1: Implications for pathogenesis. *Hum. Mol. Genet.* 14, 2063–2073.
- Choi, J., Sullards, M. C., Olzmann, J. A., Rees, H. D., Weintraub, S. T., Bostwick, D. E., Gearing, M., Levey, A. I., Chin, L. S., and Li, L. (2006) Oxidative damage of DJ-1 is linked to sporadic Parkinson and Alzheimer diseases. *J. Biol. Chem.* 281, 10816–10824.
- Li, H. M., Niki, T., Taira, T., Iguchi-Ariga, S. M., and Ariga, H. (2005) Association of DJ-1 with chaperones and enhanced association and colocalization with mitochondrial Hsp70 by oxidative stress. *Free Radical Res.* 39, 1091–1099.
- Meulener, M., Whitworth, A. J., Armstrong-Gold, C. E., Rizzu, P., Heutink, P., Wes, P. D., Pallanck, L. J., and Bonini, N. M. (2005) *Drosophila* DJ-1 mutants are selectively sensitive to environmental toxins associated with Parkinson's disease. *Curr. Biol.* 15, 1572–1577.
- Mitsumoto, A., and Nakagawa, Y. (2001) DJ-1 is an indicator for endogenous reactive oxygen species elicited by endotoxin. *Free Radical Res.* 35, 885–893.
- Park, J., Kim, S. Y., Cha, G. H., Lee, S. B., Kim, S., and Chung, J. (2005) *Drosophila* DJ-1 mutants show oxidative stress-sensitive locomotive dysfunction. *Gene* 361, 133–139.
- Sekito, A., Koide-Yoshida, S., Niki, T., Taira, T., Iguchi-Ariga, S. M., and Ariga, H. (2006) DJ-1 interacts with HIPK1 and affects H<sub>2</sub>O<sub>2</sub>-induced cell death. *Free Radical Res.* 40, 155–165.
- Shendelman, S., Jonason, A., Martinat, C., Leete, T., and Abeliovich, A. (2004) DJ-1 is a redox-dependent molecular chaperone that inhibits  $\alpha$ -synuclein aggregate formation. *PLoS Biol.* 2, 1764–1773.
- Taira, T., Saito, Y., Niki, T., Iguchi-Ariga, S. M., Takahashi, K., and Ariga, H. (2004) DJ-1 has a role in antioxidative stress to prevent cell death. *EMBO Rep.* 5, 213–218.
- Takahashi-Niki, K., Niki, T., Taira, T., Iguchi-Ariga, S. M., and Ariga, H. (2004) Reduced anti-oxidative stress activities of DJ-1 mutants found in Parkinson's disease patients. *Biochem. Biophys. Res. Commun.* 320, 389–397.
- Menzies, F. M., Yenissetti, S. C., and Min, K. T. (2005) Roles of *Drosophila* DJ-1 in survival of dopaminergic neurons and oxidative stress. *Curr. Biol.* 15, 1578–1582.
- Bandyopadhyay, S., and Cookson, M. R. (2004) Evolutionary and functional relationships within the DJ1 superfamily. *BMC Evol. Biol.* 4.

20. Lucas, J. I., and Marin, I. (2006) A new evolutionary paradigm for the Parkinson disease gene DJ-1. *Mol. Biol. Evol.* **24**, 551–561.
21. Honbou, K., Suzuki, N. N., Horiuchi, M., Niki, T., Taira, T., Ariga, H., and Inagaki, F. (2003) The crystal structure of DJ-1, a protein related to male fertility and Parkinson's disease. *J. Biol. Chem.* **278**, 31380–31384.
22. Huai, Q., Sun, Y. J., Wang, H. C., Chin, L. S., Li, L., Robinson, H., and Ke, H. M. (2003) Crystal structure of DJ-1/RS and implication on familial Parkinson's disease. *FEBS Lett.* **549**, 171–175.
23. Lee, S. J., Kim, S. J., Kim, I. K., Ko, J., Jeong, C. S., Kim, G. H., Park, C., Kang, S. O., Suh, P. G., Lee, H. S., and Cha, S. S. (2003) Crystal structures of human DJ-1 and *Escherichia coli* Hsp31, which share an evolutionarily conserved domain. *J. Biol. Chem.* **278**, 44552–44559.
24. Tao, X., and Tong, L. (2003) Crystal structure of human DJ-1, a protein associated with early onset Parkinson's disease. *J. Biol. Chem.* **278**, 31372–31379.
25. Wilson, M. A., Collins, J. L., Hod, Y., Ringe, D., and Petsko, G. A. (2003) The 1.1-angstrom resolution crystal structure of DJ-1, the protein mutated in autosomal recessive early onset Parkinson's disease. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9256–9261.
26. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolov, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., and Goldman, A. (1992) The  $\alpha/\beta$ -hydrolase fold. *Protein Eng.* **5**, 197–211.
27. Wilson, M. A., St Amour, C. V., Collins, J. L., Ringe, D., and Petsko, G. A. (2004) The 1.8-angstrom resolution crystal structure of YDR533Cp from *Saccharomyces cerevisiae*: A member of the DJ-1/ThiJ/Pfpl superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1531–1536.
28. Wei, Y., Ringe, D., Wilson, M. A., and Ondrechen, M. J. (2007) Identification of functional subclasses in the DJ-1 superfamily proteins. *PLoS Comput. Biol.* **3**, e10.
29. Zhou, W., Zhu, M., Wilson, M. A., Petsko, G. A., and Fink, A. L. (2006) The oxidation state of DJ-1 regulates its chaperone activity toward  $\alpha$ -synuclein. *J. Mol. Biol.* **356**, 1036–1048.
30. Holyoak, T., Fenn, T. D., Wilson, M. A., Moulin, A. G., Ringe, D., and Petsko, G. A. (2003) Malonate: A versatile cryoprotectant and stabilizing solution for salt-grown macromolecular crystals. *Acta Crystallogr. D59*, 2356–2358.
31. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, 307–326.
32. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. D61*, 458–464.
33. Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760–763.
34. Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr. A64*, 112–122.
35. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240–255.
36. Brunger, A. T. (1992) Free R-Value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.
37. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
38. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
39. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., III, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
40. Wilson, M. A., Ringe, D., and Petsko, G. A. (2005) The atomic resolution crystal structure of the YajL (ThiJ) protein from *Escherichia coli*: A close prokaryotic homologue of the Parkinsonism-associated protein DJ-1. *J. Mol. Biol.* **353**, 678–691.
41. Merritt, E. A. (1999) Expanding the model: Anisotropic displacement parameters in protein structure refinement. *Acta Crystallogr. D55*, 1109–1117.
42. Coates, L., Erskine, P. T., Crump, M. P., Wood, S. P., and Cooper, J. B. (2002) Five atomic resolution structures of endothiapepsin inhibitor complexes: Implications for the aspartic proteinase mechanism. *J. Mol. Biol.* **318**, 1405–1415.
43. Erskine, P. T., Coates, L., Mall, S., Gill, R. S., Wood, S. P., Myles, D. A. A., and Cooper, J. B. (2003) Atomic resolution analysis of the catalytic site of an aspartic proteinase and an unexpected mode of binding by short peptides. *Protein Sci.* **12**, 1741–1749.
44. Noda, L. H., Kuby, S. A., and Lardy, H. A. (1953) Properties of thioesters: Kinetics of hydrolysis in dilute aqueous media. *J. Am. Chem. Soc.* **75**, 913–917.
45. Polgar, L. (1974) Spectrophotometric determination of mercaptide ion, an activated form of SH-group in thiol enzymes. *FEBS Lett.* **38**, 187–190.
46. Deshpande, N., Addess, K. J., Bluhm, W. F., Merino-Ott, J. C., Townsend-Merino, W., Zhang, Q., Knezevich, C., Xie, L., Chen, L., Feng, Z., Green, R. K., Flippen-Anderson, J. L., Westbrook, J., Berman, H. M., and Bourne, P. E. (2005) The RCSB Protein Data Bank: A redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Res.* **33**, D233–D237.
47. Kinumi, T., Kimata, J., Taira, T., Ariga, H., and Niki, E. (2004) Cysteine-106 of DJ-1 is the most sensitive cysteine residue to hydrogen peroxide-mediated oxidation in vivo in human umbilical vein endothelial cells. *Biochem. Biophys. Res. Commun.* **317**, 722–728.
48. Engh, R. A., and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A47*, 392–400.
49. Nie, B., Stutzman, J., and Xie, A. (2005) A vibrational spectral maker for probing the hydrogen-bonding status of protonated Asp and Glu residues. *Biophys. J.* **88**, 2833–2847.
50. Quigley, P. M., Korotkov, K., Baneyx, F., and Hol, W. G. J. (2003) The 1.6-angstrom crystal structure of the class of chaperones represented by *Escherichia coli* Hsp31 reveals a putative catalytic triad. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3137–3142.
51. Lakshminarasimhan, M., Maldonado, M. T., Zhou, W., Fink, A. L., and Wilson, M. A. (2008) Structural impact of three Parkinsonism-associated missense mutations on human DJ-1. *Biochemistry* **47**, 1381–1392.
52. Andres-Mateos, E., Perier, C., Zhang, L., Blanchard-Fillion, B., Greco, T. M., Thomas, B., Ko, H. S., Sasaki, M., Ischiropoulos, H., Przedborski, S., Dawson, T. M., and Dawson, V. L. (2007) DJ-1 gene deletion reveals that DJ-1 is an atypical peroxiredoxin-like peroxidase. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14807–14812.
53. Shipton, M., Kierstan, M. P., Malthouse, J. P., Stuchbury, T., and Brocklehurst, K. (1975) The case for assigning a value of approximately 4 to pKa-i of the essential histidine-cysteine interactive systems of papain, bromelain and ficin. *FEBS Lett.* **50**, 365–368.
54. 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.
55. Darby, N. J., and Creighton, T. E. (1995) Characterization of the active site cysteine residues of the thioredoxin-like domains of protein disulfide isomerase. *Biochemistry* **34**, 16770–16780.
56. Olzmann, J. A., Brown, K., Wilkinson, K. D., Rees, H. D., Huai, Q., Ke, H., Levey, A. I., Li, L., and Chin, L. S. (2004) Familial Parkinson's disease-associated L166P mutation disrupts DJ-1 protein folding and function. *J. Biol. Chem.* **279**, 8506–8515.
57. Mitsumoto, A., Nakagawa, Y., Takeuchi, A., Okawa, K., Iwamatsu, A., and Takanezawa, Y. (2001) Oxidized forms of peroxiredoxins and DJ-1 on two-dimensional gels increased in response to sublethal levels of paraquat. *Free Radical Res.* **35**, 301–310.
58. Aleyasin, H., Rousseaux, M. W., Phillips, M., Kim, R. H., Bland, R. J., Callaghan, S., Slack, R. S., During, M. J., Mak, T. W., and Park, D. S. (2007) The Parkinson's disease gene DJ-1 is also a key regulator of stroke-induced damage. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18748–18753.
59. Chivers, P. T., Prehoda, K. E., Volkman, B. F., Kim, B. M., Markley, J. L., and Raines, R. T. (1997) Microscopic pKa values of *Escherichia coli* thioredoxin. *Biochemistry* **36**, 14985–14991.
60. Chivers, P. T., and Raines, R. T. (1997) General acid/base catalysis in the active site of *Escherichia coli* thioredoxin. *Biochemistry* **36**, 15810–15816.
61. Fenn, T. D., Ringe, D., and Petsko, G. A. (2003) POVScript+: A program for model and data visualization using persistence of vision ray-tracing. *J. Appl. Crystallogr.* **36**, 944–947.

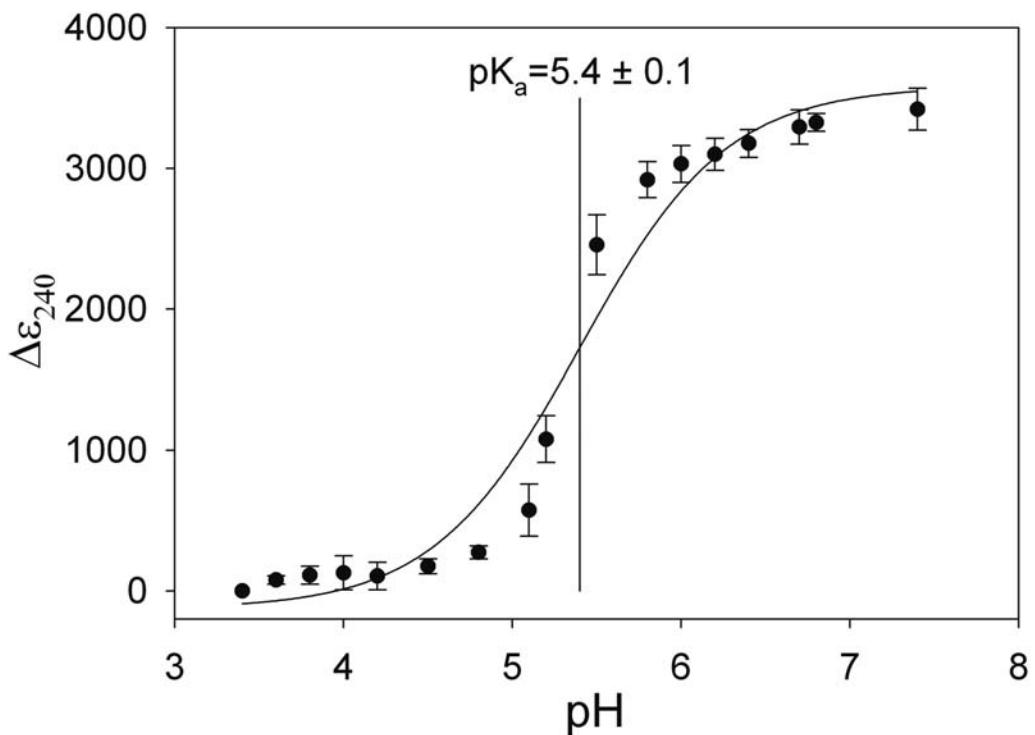
## Cysteine pK<sub>a</sub> Depression by a Protonated Glutamic Acid in Human DJ-1

Anna C. Witt<sup>†</sup>, Mahadevan Lakshminarasimhan<sup>†</sup>, Benjamin C. Remington<sup>†</sup>, Sahar Hasim<sup>†</sup>, Edwin Pozharski<sup>§</sup>, and Mark A. Wilson<sup>†\*</sup>

<sup>†</sup>Department of Biochemistry and the Redox Biology Center, The University of Nebraska-Lincoln, Lincoln, NE 68588 and <sup>§</sup>the Department of Pharmaceutical Sciences, The University of Maryland School of Pharmacy, Baltimore, MD, 21201

\*To whom correspondence should be addressed

### SUPPORTING INFORMATION



**Figure S1:** C106 pK<sub>a</sub> is unchanged by elevation of ionic strength in the double buffer. The pK<sub>a</sub> of C106 in human DJ-1 was determined by measuring the change in the extinction coefficient at 240 nm ( $\Delta\epsilon_{240}$ ) as a function of pH in 10 mM sodium citrate-sodium phosphate buffer supplemented by 100 mM NaCl. The  $\Delta\epsilon_{240}$  value is calculated by subtraction of the  $\epsilon_{240}$  measured at the lowest pH from all other measured  $\epsilon_{240}$  values. C106 has a  $pK_a=5.4\pm 0.1$  in this buffer, which is identical to the pK<sub>a</sub> measured in the double buffer alone (Figure 2A). Data were measured in triplicate and the standard deviation about the mean values is indicated with error bars.

PDB code	atom 1	res. 1 name	chain	res. 1 no.	atom 2	res. 2 name	chain	res. 2 no.	atom1-atom2 dist. (Å)	H bond angle	dihedral angle
1A4M	SG	CYS	A	153	OD2	ASP	A	181	3.52	117.40	-161.30
1A4M	SG	CYS	B	653	OD2	ASP	B	681	3.45	120.30	-164.70
1A4M	SG	CYS	C	1153	OD2	ASP	C	1181	3.30	123.30	-171.20
1ABR	SG	CYS	A	247	OE2	GLU	A	207	3.85	103.40	-172.50
1AHO	SG	BCYS	A	12	OD1	ASP	A	8	3.64	105.30	17.60
1AJ8	SG	CYS	A	21	OE2	GLU	A	187	3.71	110.30	-160.80
1AJ8	SG	CYS	B	21	OE2	GLU	B	187	3.82	108.90	-168.80
1B5L	SG	CYS	A	29	OE1	GLU	A	142	3.95	109.20	167.90
1B7B	SG	CYS	B	235	OE2	GLU	B	243	3.54	95.10	170.40
1BX7	SG	CYS	A	44	OD1	ASP	A	40	3.97	105.60	20.00
1C6V	SG	CYS	A	65	OE1	GLU	A	92	3.30	106.10	-163.70
1D2E	SG	CYS	A	222	OD2	ASP	A	228	3.41	117.10	-170.90
1D2E	SG	CYS	C	222	OD2	ASP	C	228	3.49	118.30	172.00
1DYO	SG	CYS	A	89	OE2	GLU	A	138	3.96	124.10	176.50
1DYO	SG	CYS	B	89	OE2	GLU	B	138	3.99	123.30	179.90
1E5Q	SG	CYS	A	155	OD2	ASP	A	126	3.95	104.70	163.30
1E5Q	SG	CYS	C	155	OD2	ASP	C	126	3.98	105.00	170.00
1E5Q	SG	CYS	D	155	OD2	ASP	D	126	3.96	104.10	163.20
1E5Q	SG	CYS	F	155	OD2	ASP	F	126	3.91	106.50	167.90
1E5Q	SG	CYS	G	155	OD2	ASP	G	126	3.93	105.30	165.00
1E89	SG	CYS	A	162	OD2	ASP	A	235	3.27	122.80	5.00
1EDZ	SG	CYS	A	115	OE1	GLU	A	117	3.44	108.40	-168.40
1FE0	SG	CYS	A	41	OD1	ASP	A	32	3.87	104.70	-17.60
1FX4	SG	CYS	A	959	OE1	GLU	A	955	3.85	101.60	-161.00
1FXZ	SG	CYS	A	12	OE1	GLU	A	11	3.98	117.40	-179.20
1G2I	SG	CYS	B	300	OE1	GLU	B	215	3.82	104.80	6.70
1G2I	SG	CYS	C	500	OE1	GLU	C	415	3.90	101.90	7.60
1G2I	SG	CYS	A	100	OE2	GLU	A	15	3.97	102.90	-8.80
1G2I	SG	CYS	B	300	OE2	GLU	B	215	3.93	99.10	-6.30
1G2I	SG	CYS	C	500	OE2	GLU	C	415	3.89	102.00	-7.60
1G8M	SG	CYS	A	288	OE2	GLU	A	284	3.77	118.00	169.00
1G8M	SG	CYS	B	288	OE2	GLU	B	284	3.85	116.90	169.80
1GG4	SG	CYS	B	871	OE2	GLU	B	867	3.42	94.80	-178.90
1GXC	SG	CYS	G	124	OD2	ASP	G	126	2.92	119.90	-163.60
1GYC	SG	CYS	A	117	OD1	ASP	A	456	3.71	110.20	163.20
1H2B	SG	CYS	A	112	OD1	ASP	A	109	3.85	120.00	7.60
1H2B	SG	CYS	B	112	OD1	ASP	B	109	3.97	122.40	12.70
1HCN	SG	CYS	B	26	OE2	GLU	B	19	3.42	100.30	-169.80
1HZ4	SG	CYS	A	174	OE2	GLU	A	133	3.41	115.00	169.00
1IZN	SG	CYS	C	157	OE2	GLU	C	176	3.96	108.10	178.60
1JVB	SG	CYS	A	112	OE2	GLU	A	98	3.49	99.00	174.60
1JVN	SG	CYS	B	470	OD2	ASP	B	481	3.63	108.80	179.60
1KHU	SG	CYS	C	310	OD1	ASP	C	297	3.14	122.60	18.40
1KHU	SG	CYS	D	310	OD1	ASP	D	297	3.15	120.20	15.80
1KNW	SG	CYS	A	26	OD2	ASP	A	299	3.50	109.90	0.20
1KYA	SG	CYS	A	117	OD1	ASP	A	456	3.66	105.80	163.40
1KYA	SG	CYS	B	117	OD1	ASP	B	456	3.66	105.50	162.70
1KYA	SG	CYS	C	117	OD1	ASP	C	456	3.67	106.00	164.00
1KYA	SG	CYS	D	117	OD1	ASP	D	456	3.57	107.90	162.90
1L6J	SG	CYS	A	99	OE1	GLU	A	402	2.95	121.10	17.40

1LP3	SG	CYS	A	224	OD2	ASP	A	146	2.80	104.10	17.90
1MKM	SG	CYS	B	196	OD2	ASP	B	186	3.73	116.00	-164.90
1N82	SG	CYS	A	182	OE2	GLU	A	181	3.27	120.60	165.90
1N82	SG	BCYS	B	1182	OE2	GLU	B	1181	2.97	124.40	167.30
1NEK	SG	CYS	B	75	OD2	ASP	B	63	3.17	122.70	-14.80
1NFG	SG	CYS	A	240	OE2	GLU	A	242	3.91	97.50	-165.10
1NFG	SG	CYS	B	240	OE2	GLU	B	242	3.91	97.50	-165.10
1NFG	SG	CYS	C	240	OE2	GLU	C	242	3.91	97.50	-165.10
1NFG	SG	CYS	D	240	OE2	GLU	D	242	3.91	97.50	-165.00
1NOZ	SG	CYS	B	169	OE2	GLU	B	170	3.18	97.40	179.20
1NSC	SG	CYS	A	86	OD2	ASP	A	416	3.96	101.30	175.00
1NSC	SG	CYS	B	86	OD2	ASP	B	416	3.96	102.30	175.90
1O7F	SG	CYS	A	18	OD1	ASP	A	30	3.46	107.90	161.50
1OHF	SG	CYS	B	552	OD2	ASP	C	548	3.46	113.00	11.20
1OI4	SG	CYS	B	325	OE2	GLU	B	238	3.49	112.50	11.40
1OR8	SG	CYS	A	85	OD1	ASP	A	76	3.47	105.80	-170.40
1OU0	SG	CYS	A	93	OD1	ASP	A	106	3.43	105.10	-169.00
1P35	SG	CYS	B	266	OE2	GLU	B	79	3.89	96.50	170.00
1P35	SG	CYS	C	266	OE2	GLU	C	79	3.96	104.60	164.00
1PDU	SG	CYS	B	339	OD1	ASP	B	336	3.85	96.30	-19.80
1PGS	SG	CYS	A	231	OD1	ASP	A	247	3.66	123.90	-174.10
1QDM	SG	CYS	A	37S	OE1	GLU	A	73S	3.53	124.00	160.00
1QH4	SG	CYS	A	141	OE2	GLU	A	150	3.41	106.20	-6.00
1QH4	SG	CYS	B	141	OE2	GLU	B	150	3.47	102.10	-5.20
1QH4	SG	CYS	C	141	OE2	GLU	C	150	3.42	107.20	-4.30
1QH4	SG	CYS	D	141	OE2	GLU	D	150	3.40	105.90	-4.00
1QJ4	SG	CYS	A	161	OD2	ASP	A	234	3.36	116.10	8.30
1QUB	SG	CYS	A	306	OE2	GLU	A	292	3.41	103.60	160.90
1R9D	SG	CYS	A	26	OE2	GLU	A	28	3.85	108.00	-165.30
1R9D	SG	CYS	B	26	OE2	GLU	B	28	3.93	105.10	-166.90
1RQB	SG	CYS	A	216	OD2	ASP	A	185	3.77	113.80	179.50
1RW7	SG	CYS	A	138	OE2	GLU	A	30	3.23	116.50	17.70
1RZU	SG	CYS	A	57	OE1	GLU	A	72	3.40	116.50	-167.90
1S3G	SG	CYS	A	130	OD1	ASP	A	153	3.19	116.70	-6.40
1S4E	SG	CYS	A	282	OE2	GLU	A	284	3.60	112.40	-163.50
1S4E	SG	CYS	B	282	OE2	GLU	B	284	3.49	109.20	-170.80
1S4E	SG	CYS	E	282	OE2	GLU	E	284	3.38	113.50	-168.10
1S4E	SG	CYS	F	282	OE2	GLU	F	284	3.33	117.80	-166.50
1S4E	SG	CYS	H	282	OE2	GLU	H	284	3.54	111.80	-167.50
1S6Y	SG	CYS	A	172	OE2	GLU	A	358	3.85	119.80	2.50
1SI5	SG	CYS	H	561	OD1	ASP	H	558	3.98	113.90	173.40
1SVV	SG	CYS	B	51	OD1	ASP	B	48	3.33	120.10	-19.10
1SVV	SG	CYS	A	51	OD2	ASP	A	48	3.26	113.10	-16.50
1TJO	SG	CYS	C	30	OD1	ASP	C	145	3.64	116.90	-162.10
1TJO	SG	CYS	D	30	OD1	ASP	D	145	3.64	120.80	-160.90
1U1J	SG	CYS	A	636	OD2	ASP	A	409	3.79	116.70	-170.80
1U6K	SG	CYS	A	48	OD1	ASP	A	45	3.23	109.10	-17.20
1ULZ	SG	CYS	A	228	OE1	GLU	A	295	3.49	122.00	-167.80
1V10	SG	CYS	A	117	OD1	ASP	A	455	3.17	111.40	172.30
1VPR	SG	CYS	A	972	OE1	GLU	A	1141	3.82	106.40	-175.10
1VQ2	SG	CYS	A	19	OE1	GLU	A	102	3.73	97.70	15.10
1VQ2	SG	CYS	A	19	OE2	GLU	A	102	3.65	102.20	-15.70
1W07	SG	ACYS	B	376	OE2	GLU	B	372	3.56	123.10	172.60

1WN0	SG	CYS	A	105	OD1	ASP	A	76	3.73	112.70	-171.20
1XFK	SG	CYS	A	52	OE2	GLU	A	128	3.58	118.70	166.70
1XG2	SG	CYS	B	114	OD2	ASP	B	132	3.41	110.90	-15.80
1XKL	SG	CYS	A	164	OD2	ASP	A	237	3.13	112.20	5.90
1XKL	SG	CYS	B	164	OD2	ASP	B	237	3.18	116.20	-2.00
1XKL	SG	CYS	C	164	OD2	ASP	C	237	3.18	113.80	2.20
1XKL	SG	CYS	D	164	OD2	ASP	D	237	3.16	111.10	0.10
1Y1N	SG	CYS	A	104	OD1	ASP	A	78	3.49	110.60	170.30
1YM3	SG	CYS	A	107	OD1	ASP	A	53	3.87	108.00	11.50
1YO8	SG	CYS	A	799	OE2	GLU	A	795	3.96	104.80	166.70
1YOX	SG	CYS	A	179	OD2	ASP	A	176	3.44	118.30	-176.30
1YOX	SG	CYS	E	179	OD2	ASP	E	176	3.96	103.80	174.50
1YRL	SG	CYS	C	156	OE2	GLU	C	160	3.37	112.60	-178.60
1Z08	SG	CYS	A	29	OD1	ASP	A	100	3.52	110.70	-161.70
1Z08	SG	CYS	B	29	OD1	ASP	B	100	3.59	109.00	-162.10
1Z08	SG	CYS	D	29	OD1	ASP	D	100	3.60	112.30	-162.70
2A1F	SG	CYS	D	170	OD2	ASP	D	169	3.97	109.00	-179.60
2AB0	SG	BCYS	A	106	OE2	GLU	A	17	3.02	116.70	12.90
2AB0	SG	ACYS	B	106	OE2	GLU	B	17	3.01	116.10	13.50
2B34	SG	CYS	C	114	OE2	GLU	C	110	3.17	109.70	164.30
2B3H	SG	CYS	A	202	OD1	ASP	A	176	3.56	115.70	166.50
2C1X	SG	CYS	A	368	OE2	GLU	A	383	3.90	114.20	-163.50
2CMZ	SG	CYS	A	224	OE1	GLU	A	138	3.80	119.60	-160.00
2D0V	SG	CYS	A	103	OE2	GLU	A	177	3.77	123.50	-171.00
2D2M	SG	CYS	D	2	OE2	GLU	D	1	3.61	111.90	-174.60
2DJF	SG	CYS	C	424	OE2	GLU	C	423	3.65	122.00	-161.30
2DSC	SG	CYS	A	131	OD1	ASP	B	194	3.90	111.40	160.70
2E11	SG	CYS	A	130	OE2	GLU	A	159	3.23	123.00	-6.60
2E11	SG	CYS	B	130	OE2	GLU	B	159	3.20	124.10	-7.60
2E11	SG	CYS	D	130	OE2	GLU	D	159	3.44	118.30	-6.50
2EA9	SG	CYS	A	82	OE1	GLU	A	93	3.98	117.30	-167.70
2EER	SG	CYS	A	112	OE2	GLU	A	98	3.58	99.60	-166.00
2EER	SG	CYS	B	112	OE2	GLU	B	98	3.77	98.30	-171.60
2ETL	SG	CYS	B	152	OE2	GLU	B	7	3.08	107.40	-165.30
2EZ9	SG	CYS	A	475	OD1	ASP	A	499	3.90	103.80	174.50
2EZ9	SG	CYS	B	475	OD1	ASP	B	499	3.95	104.00	174.90
2FEX	SG	CYS	A	101	OE2	GLU	A	16	3.17	121.50	13.50
2FEX	SG	CYS	C	101	OE2	GLU	C	16	3.22	121.70	17.90
2FUV	SG	ACYS	B	519	OE2	GLU	B	532	3.85	95.30	170.20
2G50	SG	CYS	A	473	OD2	ASP	A	475	3.27	118.00	-161.80
2G50	SG	CYS	G	473	OD2	ASP	G	475	3.40	117.90	-169.70
2G50	SG	CYS	H	473	OD2	ASP	H	475	3.42	116.00	-165.00
2GE3	SG	CYS	C	86	OE2	GLU	C	123	3.96	116.10	19.10
2GEL	SG	CYS	G	13	OE1	GLU	G	28	3.73	106.30	174.30
2GJ2	SG	CYS	A	46	OE1	GLU	D	31	3.20	103.90	3.80
2GJ2	SG	CYS	B	46	OE1	GLU	C	31	3.55	114.00	4.10
2GJ2	SG	CYS	C	46	OE1	GLU	B	31	3.13	123.50	2.90
2GJ2	SG	CYS	D	46	OE1	GLU	A	31	2.90	104.40	-8.30
2GL6	SG	CYS	A	238	OD1	ASP	A	96	3.69	101.50	-169.60
2GL6	SG	CYS	B	238	OD1	ASP	B	96	3.46	104.80	-170.70
2GL6	SG	CYS	C	238	OD1	ASP	C	96	3.59	108.10	-166.90
2GL6	SG	CYS	D	238	OD1	ASP	D	96	3.41	112.60	-173.50
2GL6	SG	CYS	F	238	OD1	ASP	F	96	3.56	102.00	-163.10

2GL6	SG	CYS	G	238	OD1	ASP	G	96	3.57	106.40	-172.70
2GL6	SG	CYS	H	238	OD1	ASP	H	96	3.99	95.70	-163.50
2GYQ	SG	CYS	A	79	OE1	GLU	B	88	3.41	123.30	14.10
2GYQ	SG	CYS	B	79	OE1	AGLU	A	88	3.21	106.00	7.60
2GYQ	SG	CYS	B	79	OE2	AGLU	A	88	3.38	95.00	-7.00
2H0Q	SG	CYS	A	374	OE1	GLU	A	603	3.81	117.10	165.90
2H4C	SG	CYS	C	29	OD1	ASP	C	42	3.69	120.50	170.90
2H4C	SG	CYS	H	29	OD1	ASP	H	42	3.92	115.20	-174.70
2H5E	SG	CYS	B	175	OD1	ASP	B	147	3.99	118.90	167.90
2H5U	SG	CYS	A	117	OD1	ASP	A	456	3.48	107.20	172.60
2H63	SG	CYS	C	74	OE2	GLU	C	97	3.32	122.40	-163.60
2H63	SG	CYS	D	74	OE2	GLU	D	97	3.67	119.20	164.30
2HDY	SG	CYS	A	359	OE1	GLU	A	331	3.12	113.60	179.90
2HDY	SG	CYS	B	359	OE1	GLU	B	331	3.28	106.30	-172.00
2HFK	SG	CYS	B	48	OE1	GLU	B	50	3.98	106.00	-166.50
2HL0	SG	CYS	A	127	OE2	GLU	A	134	3.69	104.90	-162.20
2HR7	SG	CYS	B	8	OE2	GLU	B	6	3.44	111.90	-178.90
2HRC	SG	CYS	A	360	OE2	GLU	A	324	3.07	103.90	-176.10
2HRG	SG	BCYS	A	117	OD2	ASP	A	453	3.66	111.00	170.70
2HZH	SG	CYS	A	117	OD2	ASP	A	456	3.95	104.80	179.00
2IV2	SG	ACYS	X	588	OD2	ASP	X	404	3.49	120.60	169.00
2J1G	SG	CYS	A	232	OD1	ASP	A	226	3.36	108.30	178.90
2JDQ	SG	CYS	A	203	OD2	ASP	A	199	3.60	113.60	-167.40
2JFK	SG	CYS	C	593	OD2	ASP	C	591	3.32	122.60	171.90
2NPI	SG	CYS	A	87	OD1	ASP	A	89	3.83	104.00	-168.50
2NTE	SG	CYS	B	743	OE1	GLU	B	740	3.90	114.30	176.30
2NYT	SG	CYS	A	131	OE1	GLU	A	60	3.65	123.20	-11.60
2NYT	SG	CYS	C	131	OE1	GLU	C	60	3.68	121.60	-10.40
2O1S	SG	CYS	B	330	OE1	GLU	B	354	3.76	118.00	-163.60
2O2P	SG	ACYS	A	707	OE1	GLU	A	673	3.85	96.60	-0.60
2O2P	SG	ACYS	B	707	OE1	GLU	B	673	3.79	97.90	0.80
2O2P	SG	ACYS	C	707	OE1	GLU	C	673	3.93	96.50	2.80
2O2P	SG	ACYS	D	707	OE1	GLU	D	673	3.78	99.80	1.30
2O2P	SG	ACYS	A	707	OE2	GLU	A	673	3.64	106.80	0.70
2O2P	SG	ACYS	B	707	OE2	GLU	B	673	3.64	105.60	-0.90
2O2P	SG	ACYS	C	707	OE2	GLU	C	673	3.71	107.50	-3.10
2O2P	SG	ACYS	D	707	OE2	GLU	D	673	3.70	104.10	-1.40
2O55	SG	CYS	A	49	OE2	GLU	A	53	3.31	112.00	-164.90
2OME	SG	CYS	A	124	OE2	GLU	A	340	3.49	114.90	-166.30
2OME	SG	CYS	H	124	OE2	GLU	H	340	3.48	112.50	-174.70
2004	SG	CYS	A	1496	OD2	ASP	A	1492	3.93	118.90	163.00
2004	SG	CYS	A	1534	OD2	ASP	A	1530	3.83	118.20	169.70
2004	SG	CYS	B	1496	OE2	GLU	B	1482	3.35	114.80	160.80
2007	SG	CYS	A	130	OD1	ASP	A	153	3.28	116.20	-10.40
2007	SG	CYS	B	130	OD1	ASP	B	153	3.30	115.20	-15.00
2OY0	SG	CYS	B	88	OD1	ASP	B	79	3.67	110.00	-172.40
2OY9	SG	CYS	B	73	OE2	GLU	B	66	3.62	123.30	-168.80
2P0I	SG	CYS	A	190	OD2	ASP	A	192	3.61	107.40	-167.50
2P26	SG	CYS	A	398	OE2	GLU	A	16	3.54	106.80	167.20
2P2W	SG	CYS	A	17	OE2	GLU	A	175	3.75	112.10	-169.10
2PLW	SG	CYS	A	37	OD2	ASP	A	28	3.29	124.40	172.60
2POZ	SG	CYS	C	383	OE2	GLU	C	57	3.66	104.30	174.40
2POZ	SG	CYS	G	383	OE2	GLU	G	57	3.93	102.30	161.40

2QBX	SG	CYS	A	105	OE2	GLU	A	117	3.88	102.60	168.50
2QBX	SG	CYS	B	105	OE2	GLU	B	117	3.46	101.60	163.40
2QH9	SG	CYS	B	18	OD2	ASP	B	11	3.94	110.10	162.40
2QT3	SG	CYS	A	300	OE2	GLU	A	324	3.58	116.80	177.80
2QT3	SG	CYS	B	300	OE2	GLU	B	324	3.61	118.40	177.20
2QZU	SG	CYS	A	158	OD2	ASP	A	136	3.40	119.50	-18.60
2RDL	SG	CYS	A	22	OE2	GLU	A	157	3.24	103.20	165.00
2RK3	SG	CYS	A	106	OE2	GLU	A	18	3.28	110.30	15.00
2TOD	SG	CYS	A	135	OD2	ASP	A	137	3.41	118.90	-169.20
2TOD	SG	CYS	B	135	OD2	ASP	B	137	3.37	116.20	177.50
2TOD	SG	CYS	C	135	OD2	ASP	C	137	3.31	116.10	178.10
2TOD	SG	CYS	D	135	OD2	ASP	D	137	3.28	123.30	-176.20
2UX0	SG	CYS	D	518	OE2	GLU	D	500	3.52	123.00	-167.50
2UX0	SG	CYS	F	518	OE2	GLU	F	500	3.67	113.80	-169.40
2V3V	SG	CYS	A	350	OE1	GLU	A	156	3.54	106.30	11.40
2V3V	SG	CYS	A	350	OE2	GLU	A	156	3.75	95.90	-10.30
2V90	SG	CYS	B	323	OE2	GLU	B	253	3.79	103.60	-165.20
2ZBA	SG	CYS	A	298	OE1	GLU	A	296	3.75	109.60	-166.10
3B89	SG	CYS	A	20	OD2	ASP	A	22	3.88	103.20	174.40
3B98	SG	CYS	B	348	OE1	GLU	B	354	3.96	111.70	-164.90
3BGE	SG	BCYS	A	319	OE2	GLU	A	327	2.91	96.70	-178.10
3BRU	SG	CYS	A	125	OE2	GLU	A	187	3.12	113.90	-9.70
3BRU	SG	CYS	B	125	OE2	GLU	B	187	3.06	118.50	-2.10
3BV6	SG	CYS	D	45	OD2	ASP	D	121	3.54	119.10	171.80
3BV6	SG	CYS	F	45	OD2	ASP	F	121	3.61	116.60	163.40
3C8X	SG	CYS	A	105	OE2	GLU	A	117	3.86	97.40	-170.20

**Table S2:** Candidate D/E-C dyads in the 4/15/08 release of the PDB. A search for all potential D/E-C candidate dyads was performed on the 11,765 non-redundant crystal structures determined at 3.0 Å or better in the 4/15/08 release of the PDB. This search required that at least one carboxylic acid oxygen from aspartic or glutamic acid be within 4.0 Å of the sulfur atom of cysteine, that the C( $\delta/\gamma$ )-O( $\delta/\epsilon;1/2$ )-S $\gamma$  hydrogen bond angle be  $109.5\pm15^\circ$ , and that the O( $\delta/\epsilon;1/2$ )-C( $\delta/\gamma$ )-O( $\delta/\epsilon;1/2$ )-S $\gamma$  dihedral angle be either  $0\pm20^\circ$  or  $180\pm20^\circ$ . Each identified interaction is tabulated above, including multiple instances of such interactions in the same set of coordinates.