# Molecular Dynamics Simulations and Oxidation Rates of Methionine Residues of Granulocyte Colony-Stimulating Factor at Different pH Values<sup>†</sup>

Jhih-Wei Chu, Jin Yin, Daniel I. C. Wang, and Bernhardt L. Trout\*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 66-458, Cambridge, Massachusetts 02139-4301

Received September 5, 2003; Revised Manuscript Received October 26, 2003

ABSTRACT: To understand the connection between the conformation of a protein molecule and the oxidation of its methionine residues, we measured the rates of oxidation of methionine residues by  $H_2O_2$  in granulocyte colony-stimulating factor (G-CSF) as a function of pH and also studied the structural properties of this protein as a function of pH via molecular dynamics simulations. We found that each of the four methionine groups in G-CSF have significant and different rates of oxidation as a function of pH. Moreover, Met<sup>1</sup>, in the unstructured N-terminal region, has a rate of oxidation as low as half that of free methionine. The structural properties of G-CSF as a function of pH are evaluated in terms of properties such as hydrogen bonding, deviations from X-ray structure, helical/helical packing, and the atomic covariance fluctuation matrix of  $\alpha$ -carbons. We found that dynamics (structural fluctuations) are essential in explaining oxidation and that a static picture, such as that resulting from X-ray data, fails in this regard. Moreover, the simulation results also indicate that the solvent-accessible area, traditionally used to measure solvent accessibility of a protein site, of the sulfur atom of methionine residues does not correlate well with the rate of oxidation. Instead, we identified a structural property, average two-shell water coordination number, that correlates well with measured oxidation rates.

The oxidation of methionine groups is one of the major degradation pathways of proteins. This occurs during the processes of drug development and production and during storage (1-7). The covalent addition of an oxygen atom to the sulfur atom of methionine changes the chemical properties of proteins and generally leads to the loss of biological function. Examples of the proteins that are destabilized by oxidation include human  $\alpha$ -1 proteinase inhibitor (8), calmodulin (9, 10), human parathoid hormone (11), antithrombin (12), glutamine synthetase (13), subtilisin (14),  $\alpha$ 1antitrypsin (15, 16), and granulocyte colony-stimulating factor (G-CSF)<sup>1</sup> (17). Stabilization against oxidation must be controlled in the product formulation of therapeutic proteins according to FDA guidelines (3). Methionine oxidation is also associated with aging, inflammatory diseases, neurologic disorders, and cataractogenesis (18-20).

Early studies suggested that the major reactive oxygen species causing the oxidation of methionine in proteins in aqueous solutions are hydroxyl radicals and hydrogen peroxide (4). Oxidation by hydroxyl radicals is catalyzed by metal ions that bind to specific sites of a protein. This type of mechanism is referred to as "site-specific" oxidation (4, 18, 21). Oxidation by hydrogen peroxide, on the other hand, is generally accepted to be an  $S_{\rm N}2$  reaction. It is referred to

as "non-site-specific" oxidation (4). The charge separation developed during the breaking of the O-O bond in H<sub>2</sub>O<sub>2</sub> contributes to most of the energetic penalty of the oxidation reaction (22-25). This observation led to the proposal in a 1994 theoretical study that a strong proton donor, such as a hydronium ion, would be necessary to affect the oxidation (26). After the 1994 study, it was observed experimentally that over a large range of pH ( $\sim$ 2-10) there is no pH dependence on the rate of oxidation (27). This observation called into question the validity of the 1994 mechanism over an intermediate range of pH. We thus performed a new study, the results of which led us to propose a new mechanism of charge stabilization resulting from specific interactions, such as hydrogen bonding, with surrounding water molecules (25). We found that these interactions stabilize the charge separation and result in the observed activation energy of the reaction (10-15 kcal/mol) (25). Our mechanism is consistent with all available experimental information about the oxidation mechanism, such as activation energies (23) and the pH dependence of oxidation rates (27). It also leads to the conclusion that exposure to solvent molecules plays an essential role in oxidation reactions. In several studies, it was shown that amino acids that are located in buried positions (determined qualitatively from static X-ray data) are less easily oxidized by exogenously added peroxide oxidants (16, 28, 29) than those at the surface. This implies that the access of solvent to methionine sites affects oxidation. While this observation is consistent with our proposed mechanism, it is qualitative at best, because it is based on static data. Clearly, proteins have essential dynamic behavior that is not captured by static X-ray structure.

<sup>&</sup>lt;sup>†</sup> This work was supported in part by Amgen, Inc., by the National University of Singapore, and by the Singapore–MIT Alliance.

<sup>\*</sup> To whom correspondence should be addressed. E-mail: trout@mit.edu.

<sup>&</sup>lt;sup>1</sup> Abbreviations: G-CSF, granulocyte colony-stimulating factor; MD, molecular dynamics; SAA, solvent-accessible area; WCN, water coordination number; RMSD, root of mean squared deviation; RMSF, root of mean squared fluctuation.

This study addresses the effects of protein dynamics and thermal fluctuations on the rates of oxidation of methionine sites in therapeutic proteins. We focus on G-CSF, a protein pharmaceutical useful in treating patients suffering from neutropenia and in mobilizing peripheral blood progenitor cells for harvesting and transplantation after myeloablative chemoradiotherapy (17). G-CSF has four methionine groups, Met<sup>1</sup>, Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>. Each of the methionine groups has been shown to be oxidized by hydrogen peroxide with distinct oxidation rates in the order Met<sup>1</sup> > Met<sup>138</sup> > Met<sup>127</sup> > Met<sup>122</sup>. Furthermore, the bioactivity measured by a cell mitogenesis bioassay was shown to decrease to 3% after all methionine groups are oxidized (17).

A measure typically used to quantify the degree to which the solvent can interact with different parts of the protein is solvent-accessible area (SAA) (30). However, using the X-ray structure by Aritoni et al. (31, 32), PDB entry 1CD9, and a probe radius of 1.6 Å, we compute that the SAA's of Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup> are zero. Since all three of these Met groups have appreciable rates of oxidation, it is unlikely that this structural quantity either provides a good measure of the degree to which solvent can interact with sites in the protein or relates to oxidation. It seems to be important at least to incorporate protein and solvent dynamics in any model that relates structural properties to rates of oxidation.

In this study, molecular dynamics simulations of G-CSF together with experiments were performed. The simulations incorporated an explicit treatment of the solvent molecules and realistic dynamical motions of each atom in the protein. Our goal was to better understand the oxidation of methionine by relating structural properties to rates of oxidation. It was hoped that we could find a structural parameter that correlates well to rates of oxidation. In addition, we measured the effects of pH on the rates of oxidation and analyzed how structural properties change as a function of pH.

# MATERIALS AND METHODS

Experimental Procedures for Measuring Oxidation Rates of Methionine Groups in G-CSF. (A) Materials. G-CSF (4 mg/mL in water, pH 4.0 via addition of HCl) was obtained from Amgen, Inc. (Thousand Oaks, CA). H<sub>2</sub>O<sub>2</sub> was purchased from Sigma (H 1009). All other reagents and chemicals used were of analytical grade, and deionized water was used through the study.

(B) Oxidation of G-CSF by Hydrogen Peroxide. G-CSF was mixed with 25 mM acetate (pH 4.5, ionic strength 0.1 M) and 30%  $\rm H_2O_2$  (final  $\rm H_2O_2$  concentration = 30, 90, or 180 mM, and G-CSF concentration = 0.1 mg/mL) incubated at 25 °C. At various time intervals, reactions were terminated by loading samples onto a disposable PD-10 desalting column (Amersham Pharmacia Biotech 17-0851-01) for removal of excess  $\rm H_2O_2$ . For each reaction, a 2.5 mL G-CSF fraction was collected for analysis.

(C) Separation of G-CSF Oxidized Forms. G-CSF oxidized forms were separated by reverse-phase HPLC (the Integral HPLC system, PerSeptive Biosystems, serial no. 176, part no. 5-1818-00). A 1.0 mL  $\rm H_2O_2$ -oxidized sample was injected into a  $\rm C_8$  column (Vydac 208TP54; 300 Å, 5  $\mu$ m, 4.6 mm i.d.  $\times$  25 cm). The mobile phases used were solvent A [0.3% (w/v) trifluoroacetic acid (TFA)] and solvent B [0.255% (w/v) TFA in 90% acetonitrile]. The column was

initially equilibrated with 56% solvent B at a flow rate of 0.8 mL/min. After sample injection, the column was washed with 56% solvent B for 10 min. The separation was then performed by a linear gradient of 56% solvent B to 66% solvent B for 100 min and 66% solvent B to 100% solvent B for 10 min and then by isocratic elution at 100% solvent B for 10 min with a UV detector set at 215 nm.

(D) Trypsin Peptide Map Analysis for the Assay of Met<sup>1</sup>. G-CSF in 25 mM acetate was buffer-exchanged into 10 mM Tris-HCl, 100 mM NaOAc, and 1 mM CaCl<sub>2</sub>, pH 8.3, by using a disposable PD-10 desalting column. A 2.5 mL fraction was collected and digested with trypsin (Sigma T 8658), using an enzyme-to-substrate ratio of 1:50 for 18 h at 30 °C for a subsequent reverse-phase HPLC peptide map analysis of Met<sup>1</sup>. A 1.0 mL trypsin-digested G-CSF mixture was injected into a C<sub>8</sub> column (Vydac 208TP54; 300 Å,  $5\,\mu\text{m},~4.6$  mm i.d.  $\times~25$  cm). The mobile phases used included solvent A [0.1% (w/v) TFA] and solvent B [0.1% (w/v) TFA in 90% acetonitrile]. The column was initially equilibrated with 25% solvent B at a flow rate of 0.8 mL/min. After sample injection, the column was washed with 25% solvent B for 10 min. The separation was then performed by a linear gradient of 25% solvent B to 40% solvent B for 30 min and 40% solvent B to 100% solvent B for 10 min and then by isocratic elution at 100% solvent B for 10 min with a UV detector set at 215 nm.

(E) Glu-C Peptide Map Analysis for the Assay of Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>. G-CSF in 25 mM acetate was bufferexchanged into 2.5 mL of 6 M guanidine hydrochloride (GdnHCl) and 0.3 M Tris-HCl, pH 8.5, by the use of a PD-10 desalting column. Modification of the disulfide bonds of G-CSF was performed by the addition of 35 mM dithiothreitol (DTT) (Sigma D 5545). After incubation at room temperature for 45 min, G-CSF was buffer-exchanged into 2.5 mL of 6 M GdnHCl and 0.3 M Tris-HCl, pH 8.5, containing 4 mM 7-chloro-4-nitrobenz-2-oxa-1,3-diazole (NbdCl) (Sigma C 5261) by the use of a PD-10 desalting column. After incubation at room temperature for 60 min, G-CSF was buffer-exchanged into 2 M urea and 0.1 M Tris-HCl, pH 7.8, by the use of a PD-10 desalting column and digested with endoproteinase Glu-C (Sigma P 6181), using an enzyme-to-substrate ratio of 1:25 for 26 h at 25 °C, for subsequent reverse-phase HPLC peptide map analysis of Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>. A 1.0 mL Glu-C-digested G-CSF mixture was injected into a C<sub>4</sub> column (Vydac 214TP54; 300 Å, 5  $\mu$ m, 4.6 mm i.d.  $\times$  25 cm). The mobile phases used were solvent A [0.1% (w/v) TFA] and solvent B [0.1% (w/v) TFA in 90% acetonitrile]. The column was initially equilibrated with 3% solvent B at a flow rate of 0.8 mL/min. After sample injection, the column was washed with 3% solvent B for 10 min first. The separation was performed by a linear gradient of 3% solvent B to 30% solvent B for 40 min, 30% solvent B to 55% solvent B for 60 min, and 55% solvent B to 100% solvent B for 10 min and by isocratic elution at 100% solvent B for 10 min with the UV detector set at 215 nm.

(F) Peptide Map Analyses of Methionine Oxidation in G-CSF. The extent of oxidation of individual methionine residues in G-CSF was quantified by trypsin (for Met<sup>1</sup>) and endoproteinase Glu-C (for Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>) peptide map analyses. Electrospray mass spectrometry was performed at the MIT Biopolymers Laboratory using a Perkin-Elmer

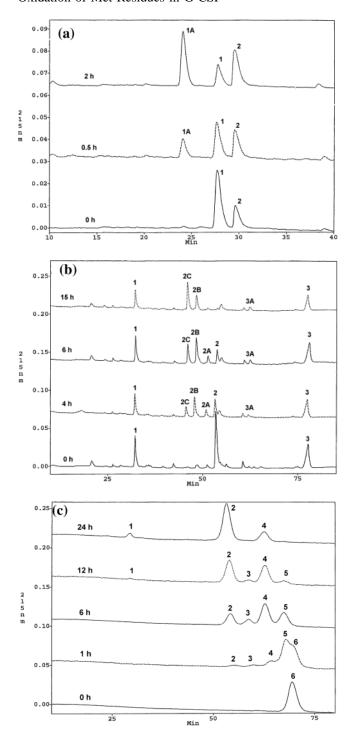


FIGURE 1: Reverse-phase HPLC of G-CSF samples. Oxidation reactions were carried out with 30 mM H<sub>2</sub>O<sub>2</sub> and 0.1 mg/mL G-CSF in 25 mM acetate at a pH of 4.5 at 25 °C. (a) Trypsin peptide map of G-CSF samples taken at 0, 0.5, and 2 h during H<sub>2</sub>O<sub>2</sub> oxidation. (b) Endoproteinase Glu-C peptide map of G-CSF samples taken at 0, 4, 6, and 15 h during H<sub>2</sub>O<sub>2</sub> oxidation. (c) Oxidized forms of G-CSF. Peak 6 represents the intact G-CSF.

Sciex Model 365 triple stage mass spectrometer (Perkin-Elmer Analytical Instruments, Norwalk, CT). Peptides collected from reverse-phase HPLC separations were dried, reconstituted in 50% acetonitrile with 0.2% acetic acid, and injected into the electrospray interface. Mass spectra were acquired over an m/z range of 400–2300. Figure 1 compares the trypsin peptide map of intact G-CSF (0 h) with those of the samples taken at 0.5 and 2 h during H<sub>2</sub>O<sub>2</sub> (30 mM) oxidation. In Figure 1a, peptide 1 is a 17-mer peptide containing Met<sup>1</sup> and peptide 2 is a 19-mer peptide (see Table 1). Peptide 1A is the oxidation product of peptide 1. An increase of 16 mass units found for peptide 1A represents the oxidation of Met<sup>1</sup> in peptide 1.

The endoproteinase Glu-C peptide map of intact G-CSF (0 h) with those of the samples taken at 0, 4, 6, and 15 h during H<sub>2</sub>O<sub>2</sub> oxidation is presented in Figure 1b. Peptide 1 is a 12-mer, peptide 2 is a 39-mer containing Met<sup>127</sup> and Met<sup>138</sup>, and peptide 3 is a 25-mer containing Met<sup>122</sup>, as summarized in Table 2. Peptide 3A is the oxidation product of peptide 3. An increase of 16 mass units found for peptide 3A represents the oxidation of Met<sup>122</sup> in peptide 3. Likewise, peptides 2A, 2B, and 2C are the oxidation products of peptide 2. An increase of 32 mass units found for peptide 2C represents the oxidation of both Met<sup>127</sup> and Met<sup>138</sup> in peptide 2 while an increase of 16 mass units found for peptides 2A and 2B represents the oxidation of either Met<sup>127</sup> or Met<sup>138</sup>, as shown in Table 2. We used cyanogen bromide (33) to cleave peptides 2A and 2B and confirmed that 2A represents the oxidation of Met<sup>127</sup> and 2B represents the oxidation of Met<sup>138</sup>. Peptide 2A was not identified in an earlier study (17).

(G) Oxidized Forms of G-CSF. As a validation, reversephase HPLC was used to separate different G-CSF forms from G-CSF samples that had been oxidized by H<sub>2</sub>O<sub>2</sub> at different times. As shown in Figure 1c, six distinct peaks, labeled 1-6, were obtained. Each peak was collected and subjected to electrospray mass spectrometry in order to measure the molecular weight, and the results are presented in Table 3. In addition, each peak was also subjected to Glu-C and trypsin peptide map analyses. Results indicate that peak 1 contains G-CSF with all four methionines oxidized, peak 2 contains oxidized Met<sup>1</sup>, Met<sup>127</sup>, and Met<sup>138</sup>, peak 3 contains oxidized Met<sup>1</sup> and Met<sup>127</sup>, peak 4 contains oxidized Met<sup>1</sup> and Met<sup>138</sup>, peak 5 contains only oxidized Met<sup>1</sup>, and peak 6 contains intact G-CSF. Peaks 1, 2, 4, 5, and 6 were also identified by Lu et al. (17). Peak 3, with G-CSF oxidized at Met1 and Met127, was identified for the first time in this work. On the basis of the above results, we conclude that the order of oxidation rates of methionine groups at pH 4.5 is  $Met^{1} > Met^{138} > Met^{127} > Met^{122}$ .

Molecular Dynamics Simulations. Molecular dynamics simulations of G-CSF were performed using the CHARMM software (34) and the CHARMM 22 all-atom potential (35). The initial structure of G-CSF was obtained from the X-ray structure of Aritomi et al. (31, 32), determined at pH 7.5, PDB entry 1CD9. The coordinates of residues missing in the PDB data (residues 1-6) were generated using the internal coordinate table; the coordinates of hydrogen atoms were then generated using the HBUILD (34) function of CHARMM. One hundred steepest descent steps followed by 500 conjugate gradient steps were applied to the initial geometry with heavy atoms fixed at their PDB coordinates. The resulting geometry was then further minimized with 300 conjugate gradient steps with heavy atoms restrained by their PDB coordinates. The restraint force constant was 30 kcal  $\text{mol}^{-1} \text{ Å}^{-1}$  and was decreased by 10 kcal  $\text{mol}^{-1} \text{ Å}^{-1}$  at every 100 minimization steps. The resulting geometry had a 0.12 Å all-atom RMSD (root of mean squared deviation) from the original X-ray coordinates and was used as the reference structure for analysis. G-CSF has five histidine residues, and the protonation states were determined by examining the

Table 1: Structural Characterization of the Trypsin-Digested Peptides of G-CSF in Figure 1a

peptides	sequences	MW detected <sup>a</sup>	mass difference	oxidized methionines
1 (17-mer)	MTPLGPASSLPQSFLLK	$1786.82 \pm 0.02$		
1A		$1802.82 \pm 0.27$	16.00	$Met^1$
2 (19-mer)	AGGVLVASHLQSFLEVSYR	$2032.84 \pm 0.04$		

<sup>&</sup>lt;sup>a</sup> Obtained from electrospray mass spectrometry.

Table 2: Structural Characterization of G-CSF Endoproteinase Glu-C Digested Peptides in Figure 1b

peptides	sequences	MW detected <sup>a</sup>	mass difference	oxidized methionines
1 (12-mer)	VSYRVLRHLAQP	$1438.36 \pm 0.40$		
2 (39-mer)	LGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLE	$4027.44 \pm 1.15$		
2A		$4044.00 \pm 1.34$	16.56	Met <sup>127</sup>
2B		$4044.59 \pm 0.27$	17.15	Met <sup>138</sup>
2C		$4059.29 \pm 0.83$	31.85	Met <sup>127</sup> , Met <sup>138</sup>
3 (25-mer)	LGPTLDTLQLDVADFATTIWQQMEE	$2836.13 \pm 0.24$		
3Å		$2852.27 \pm 0.70$	16.14	Met <sup>122</sup>
- 61 1 1 6				

<sup>&</sup>lt;sup>a</sup> Obtained from electrospray mass spectrometry.

Table 3: Characterization of Intact G-CSF and Its Derivatives Shown in Figure 1c

G-CSF	peaks	$\mathbf{M}\mathbf{W}^{a}$	mass difference	oxidized methionines
native	6	$18798.73 \pm 0.85$		
	5	$18816.53 \pm 2.32$	17.80	Met <sup>1</sup>
	4	$18833.60 \pm 2.88$	34.87	Met <sup>1</sup> , Met <sup>138</sup>
oxidized	3	$18832.39 \pm 2.18$	33.66	Met <sup>1</sup> , Met <sup>127</sup>
	2	$18849.29 \pm 1.23$	50.56	Met <sup>1</sup> , Met <sup>127</sup> , Met <sup>138</sup>
	1	$18868.40 \pm 3.26$	69.67	Met <sup>1</sup> , Met <sup>122</sup> , Met <sup>127</sup> ,
				Met <sup>138</sup>

<sup>&</sup>lt;sup>a</sup> Obtained from electrospray mass spectrometry.

relative distances to surrounding hydrogen bond donors and acceptors.

The reference structure was then solvated in a preequilibrated (295 K, 1 atm) truncated octahedral water cell with a lattice length of 68 Å containing 8280 water molecules. Counterions were placed using the SOLVATE program (36). Water molecules with oxygen atoms within 2.4 Å of the heavy atoms of G-CSF and counterions were removed. Periodic boundary conditions were applied, and the longrange Coulombic interactions were calculated using the particle-mesh Ewald summation method (37, 38). The cutoff radius for electrostatic interactions was 15 Å. Covalent bonds were fixed during the simulations using the SHAKE (39) algorithm, thus allowing a 0.002 ps integration time step. Fifty steepest descent steps followed by 200 conjugate gradient minimization steps were performed first with protein atoms fixed. Next, the system was heated to 295 K by gradually scaling the velocities (3 K/0.1 ps) for 10 ps. During the heating period, the backbone atoms of G-CSF were restrained to the reference structure by a force constant of 10 kcal  $\text{mol}^{-1}$  Å<sup>-1</sup>. After the heating period, a 30 ps equilibration simulation at 295 K was performed using the Nosé-Hoover thermostat (40). During the equilibration period, the restraint force constant for the protein backbone atoms was gradually decreased by 5 kcal mol<sup>-1</sup> Å<sup>-1</sup> every 10 ps. For the last 10 ps of equilibration, no restraint force was applied. Then the molecular dynamics simulation was performed in the canonical ensemble at 295 K for 2.96 ns, during which averaged properties were computed. Trajectories were saved every 0.1 ps for analysis, and the statistical uncertainties of averaged properties were estimated on the basis of standard procedures (39).

Simulations were performed separately for three regimes of pH. The first is called "high-pH", in which only one of the nitrogen atoms of histidine residues was protonated. The carboxyl oxygen of aspartic acid and glutamic acid residues was not protonated. The second is called "medium-pH", in which both nitrogen atoms of the histidine residues were protonated but not the carboxyl oxygen of aspartic acid and glutamic acid residues. The last is called "low-pH", in which all acidic residues were protonated.

To compare the results of molecular dynamics simulations with the measured rates of oxidation of methionine residues, a mapping between protonation states and pH values is needed. G-CSF has five histidine groups. From the X-ray structure, the protonation states of His<sup>80</sup> and His<sup>157</sup> may affect the fluorescence intensity of the nearby Trp<sup>59</sup>(His<sup>157</sup>) and Trp<sup>119</sup>(His<sup>80</sup>) residues (41, 42). The pH dependence of the fluorescence intensities of G-CSF suggests that the  $pK_a$ 's of His<sup>80</sup> and His<sup>157</sup> are around 5.8 (43). The rest of histidine residues (His<sup>44</sup>, His<sup>53</sup>, His<sup>171</sup>) are more exposed to the solvent, and we assume that the  $pK_a$ 's are close to the value of the free residue, 6.1. G-CSF also contains four Asp and nine Glu residues over the protein surface, and we also assume that their  $pK_a$ 's do not deviate significantly from the values of the free residues. On the basis of the above information, we assume that the high-pH protonation configurations correspond roughly to pH 8.5, the medium-pH configurations correspond roughly to pH 5.5, and the low-pH configurations correspond roughly to pH 2.

To evaluate the effects of different protonated configurations of His<sup>80</sup> and His<sup>157</sup>, three different protonation configurations of His<sup>80</sup> and His<sup>157</sup> are examined. They are all considered to be in the high-pH regime. The first configuration is called "high-pH-1". In this configuration, the  $\epsilon$ -nitrogen of His<sup>80</sup> and the  $\delta$ -nitrogen of His<sup>157</sup> are protonated. The second configuration, in which the  $\delta$ -nitrogen of His<sup>80</sup> and the  $\delta$ -nitrogen of His<sup>157</sup> are protonated, is called "high-pH-2". The last configuration, in which the  $\epsilon$ -nitrogen of His<sup>80</sup> and the  $\epsilon$ -nitrogen of His<sup>157</sup> are protonated, is called "high-pH-3".

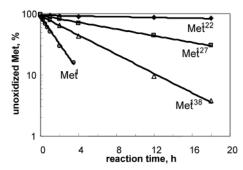


FIGURE 2: Time course of the percentage of unoxidized methionine residues on a semilogarithmic scale. Oxidation reactions were carried out with 0.1 mg/mL G-CSF at a pH of 4.5 and 25 °C. H<sub>2</sub>O<sub>2</sub> concentrations were 30 mM.

Table 4: Pseudo-First-Order Oxidation Rate Constants of Methionine Groups in G-CSF at Different pH Values<sup>a</sup>

pН	Met <sup>1</sup>	Met <sup>122</sup>	Met <sup>127</sup>	Met <sup>138</sup>
2	0.566	0.024	0.086	0.253
3	0.509	0.021	0.083	0.221
4	0.495	0.019	0.082	0.194
4.5	0.470	0.020	0.087	0.190
$4.5^{b}$	0.486	0.002	0.180	0.263
5.5	0.483	0.025	0.093	0.196
6.5	0.575	0.037	0.079	0.194
7.5	0.792	0.036	0.083	0.194
8.5	0.834	0.044	0.101	0.215

<sup>&</sup>lt;sup>a</sup> Oxidation reactions were carried out with 30 mM H<sub>2</sub>O<sub>2</sub> and 0.1 mg/mL G-CSF at 25 °C. The data recorded are rate constant ( $h^{-1}$ ) × 100. b Results from Lu et al. (17), converted to 30 mM H<sub>2</sub>O<sub>2</sub>.

#### RESULTS AND DISCUSSION

Oxidation Rates of Methionine Residues of G-CSF at Different pH Values. In this section, the experimental results of the oxidation rates of methionine groups in G-CSF are presented. Since the concentration of H<sub>2</sub>O<sub>2</sub> (30 mM and greater) is much higher than that of the protein (0.1 mg/mL = 5.3  $\mu$ M), its concentration is assumed to be constant in the kinetic analysis. In Figure 2, the concentration of unoxidized methionine residues in G-CSF at pH = 4.5 is normalized to their initial concentration and plotted versus time on a semilogarithmic scale. The very high linearity  $(R^2 > 0.98)$  of Figure 2 indicates that the reaction is pseudo first order for each methionine residue. The first-order nature of the reaction also indicates that the oxidation of one methionine residue does not affect the rates of oxidation of other methionine residues in the experimental time scale; i.e., a significant change of the protein conformation did not occur. For Met<sup>122</sup>, on the other hand, there indeed exists a different oxidation rate constant after the rest of methionine residues are oxidized, as reported earlier (17) and as has also been confirmed by our experiments (results not shown). In this study, our focus is on the stability of native G-CSF, i.e., initial rates of oxidation, and, thus, change in the conformation of G-CSF as a result of oxidation is not a concern. The reaction order of H<sub>2</sub>O<sub>2</sub> was also examined by performing the oxidation experiments at different H<sub>2</sub>O<sub>2</sub> concentrations, and the results indicate that the reaction is also of first order with respect to H<sub>2</sub>O<sub>2</sub>. The pseudo-first-order rate constants of each methionine residue as well as the results in Lu et al. (17) are shown in Table 4. For Met<sup>1</sup>, our measured rate constant is very close to Lu's result. However, for Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>, our rate constants are significantly

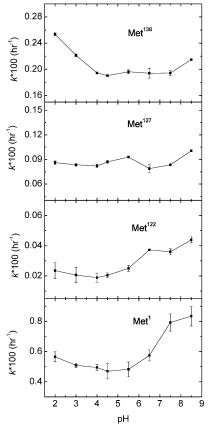


FIGURE 3: Rates of oxidation in G-CSF at different pH values. The pseudo-first-order rate constants in units of h<sup>-1</sup> are plotted as a function of pH for each methionine residue. Oxidation reactions were carried out with 0.1 mg/mL G-CSF at 25 °C. H<sub>2</sub>O<sub>2</sub> concentrations were 30 mM.

different. There are several possible reasons for these differences. First, the difference in the forced oxidation conditions may be important. The H<sub>2</sub>O<sub>2</sub> concentration is 30 mM in this study and 300 mM in the study of Lu et al. (17). Second, our data may be more precise; this is demonstrated by the fact that we were able to resolve an additional peak (Met1/Met127 oxidized form of G-CSF) in the mass spectral experiments.

The oxidation rate of each methionine residue was then measured at an H<sub>2</sub>O<sub>2</sub> concentration of 30 mM at different pH values. The buffers used were phosphate at pH 2.0, citrate at pH 3.0, 5.5, and 6.5, acetate at pH 4.0, and Tris-HCl at pH 8.5. The pH dependence of oxidation rates is shown in Figure 3. The oxidation rates of Met<sup>1</sup> show a minimum at pH 4.5 and are 20% higher at pH 2 and 77% higher at pH 8.5. Met<sup>138</sup> is oxidized the second fastest and shows a moderate pH dependence of oxidation rates with a minimum at pH = 4.5. The oxidation rate of  $Met^{138}$  is 33% higher at pH 2 and 13% higher at pH = 8.5 than at pH 4.5. For Met<sup>127</sup>, the pH dependence of oxidation rates is small (the difference between the maximum and minimum rate constant is <15%). Met<sup>122</sup> oxidizes at the lowest rate among the methionine residues of G-CSF and shows a weak pH dependence of oxidation rates between pH 2 and pH 5.5 but increases by 120% from pH 4.5 to pH 8.5.

The experimental results indicate that each methionine residue in G-CSF has a different rate of oxidation and that the rates of oxidation of each methionine residue are different functions of pH. Note that we measured the rates of oxidation

Table 5: Distribution of Charged Residues and Net Charges of Subdomains of G-CSF

	domain													
	$L_0^a$ (1-11)	$A^a$ (12-40		AE <sup>a</sup> -44)	E <sup>a</sup> (45-55)	L <sub>EB</sub> <sup>a</sup> (56-72)	B <sup>a</sup> (73-93)	L <sub>BC</sub> <sup>a</sup> (94-99)	$C^a$ (100–12	25)	L <sub>CD</sub> <sup>a</sup> (126-143)	D <sup>a</sup> (144-1	75)	total <sup>b</sup>
Arg Lys		23 17, 24, 3	5 4	41								147, 148, 1	67, 170	7 4
His Asp		28		44	53		80		105, 110,	113		157, 171		5 4
Glu		20, 34			46, 47			94, 99	123, 124			163		9
							1	net charge						
		$L_0^a$	$A^a$	$L_{AE}$	a E	a L	EB <sup>a</sup>	$\mathbf{B}^{a}$	$L_{BC}^{a}$	$\mathbb{C}^a$	$L_{CI}$	o <sup>a</sup> ]	$D^a$	

	$L_0^a$	$A^a$	$L_{AE}{}^{a}$	$\mathbf{E}^{a}$	$L_{\mathrm{EB}}{}^{a}$	$\mathbf{B}^{a}$	$L_{\mathrm{BC}}^{a}$	$\mathbb{C}^a$	$L_{CD}^{a}$	$\mathbf{D}^{a}$	
$pH^c$	(1-11)	(12-40)	(41-44)	(45-55)	(56-72)	(73 - 93)	(94-99)	(100-125)	(126-143)	(144-175)	$total^b$
low	0	4	2	1	0	1	0	0	0	6	14
medium	0	1	2	-1	0	1	-2	-5	0	5	1
high	0	1	1	-2	0	0	-2	-5	0	3	-4

<sup>&</sup>lt;sup>a</sup> The numbers shown in the column are the residue numbers of charged residues. <sup>b</sup> The total number of charged residues in G-CSF. <sup>c</sup> The protonated state of each pH category is described in the text.

of free methionine at different pH values and found no pH dependence between pH 2 and pH 8.5 (rate  $\times$  100 = 0.98 h<sup>-1</sup> at a H<sub>2</sub>O<sub>2</sub> concentration of 30 mM). Also note that a previous study on the oxidation of dimethyl sulfide (DMS), which represents part of the methionine side chain, reports no pH dependence between pH 2 and pH 10 (27). Furthermore, our proposed mechanism, recently reported, does not lead to a pH dependence of the oxidation of organic sulfides by hydrogen peroxide over intermediate ranges of pH (25). All of these results suggest that the pH dependence of the rates of oxidation of a given methionine is due to the changes in the local environment in response to solution pH. This hypothesis is examined in detail below using the results of molecular dynamics simulations.

Identification of the Governing Factors of Oxidation of Methionine via Molecular Dynamics Simulations. The structures and conformation of G-CSF at different pH values were investigated via molecular dynamics simulations of G-CSF with different protonation configurations in the unoxidized form. Since all methionine groups show first-order kinetics in the early phase of oxidation, using the unoxidized form of G-CSF to relate G-CSF conformation to its initial rates of oxidation is valid. In this section, we first describe the structural properties of G-CSF at different pH values and then study the correlation between the structural properties of G-CSF and the initial rates of oxidation of its methionine residues.

(A) Labeling of Regions in G-CSF. We first label regions of the protein and next describe their conformational properties. G-CSF belongs to the four-helix-bundle structural superfamily of growth factors (44). The ribbon representation of the X-ray structure of G-CSF in Figure 4 shows that G-CSF has four main helices, A (residues 12-40), B (residues 73–93), C (residues 100–125), and D (residues 143-173), and a short helical region, E (residues 45-55). The four main helices are arranged in an up-up-downdown fashion by the long loops  $L_{EB}$  (residues 56–72), connecting helices E and B, and L<sub>CD</sub> (residues 126-172), connecting helices C and D. Residues 1-11 of G-CSF, denoted by L<sub>0</sub>, do not form a well-defined secondary structure. The short segment between helices A and E (residues 41-44) is denoted by LAE, and the segment between helices B and C (residues 94-99) is denoted as

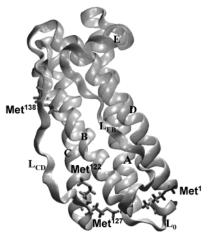


FIGURE 4: Ribbon representation of the X-ray structure of G-CSF (PDB code 1CD9). The figure was prepared using the VMD package (56). The secondary structure was identified using DSSP (57). The main bundle helices A (residues 12–40), B (residues 73–93), C (residues 100–125), and D (residues 144–173) are labeled near their N termini. The short  $3_{10}$  (residues 45–48) and  $\alpha$  (residues 51–55) helices are also indicated.

 $L_{BC}$ . The net charges of G-CSF and its subdomains at various pH ranges are summarized in Table 5. Basic residues are mostly located in helices A and D, and acidic residues are mostly located in helix C. Note that helices A and C also contain the receptor binding domains of G-CSF (31, 45). Loops  $L_0$ ,  $L_{BE}$ , and  $L_{CD}$  do not have charged residues.

(B) Overall Structure and Fluctuations. The time course of the root mean square deviations (RMSD) of the backbone heavy atoms of G-CSF from the X-ray reference structure is shown in Figure 5. Backbone RMSDs for the helical domains (A-D), for L<sub>0</sub>-excluded residues (denoted as L<sub>0</sub>-ex in Figure 5), and for all residues are presented. The helical regions have relatively small RMSDs from the X-ray structure for all protonation configurations throughout the simulations. The largest average helical region RMSD (last 1.5 ns) occurs in the low-pH simulation (1.2 Å). When we include residues other than those in the L<sub>0</sub> region, the RMSDs from the X-ray structure become larger but stay constant after 1.5 ns in the simulation for all protonation configurations. Including the L<sub>0</sub> region further increases the RMSD values, indicating that the loop regions of G-CSF have deviated more from the reference structure than the helical domains have.

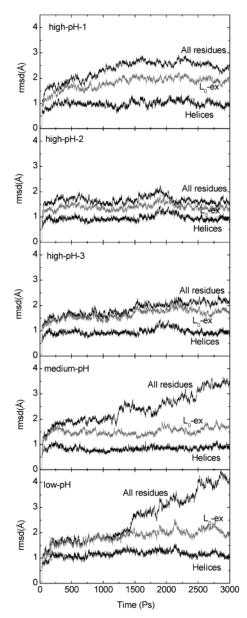


FIGURE 5: Time course of the root-mean-square deviations (RMSD) of backbone atoms from the X-ray structure (PDB code 1CD9) of G-CSF. RMSDs for the helical domains (A-D), for all residues except the  $L_0$  region ( $L_0$ -ex), and for all residues are plotted. The data points are shown for every 1 ps.

The small backbone RMSDs for helical domains in the simulation of all protonation configurations suggest that the secondary as well as the up-up-down-down arrangement of the helical domains of G-CSF is stable for a wide pH range, even at low pH values where many proteins may exhibit the pH-induced partially unfolded "A" state (46, 47). The conformational stability of G-CSF at low pH values has also been reported using fluorescence and circular dichroism spectra (48).

In the high-pH regime, the high-pH-2 simulation has the smallest average backbone all-residue (1.71 Å) and L<sub>0</sub>-ex (1.44 Å) RMSDs among all three cases (high-pH-1, 1.96 Å L<sub>0</sub>-ex RMSD, 2.58 Å all-residue RMSD; high-pH-3, 2.0 Å L<sub>0</sub>-ex RMSD, 1.8 Å all-residue RMSD). These differences highlight the fact that changes in the protonation of the nitrogen atoms in histidine residues may result in distinct protein conformational changes. The medium-pH simulation,

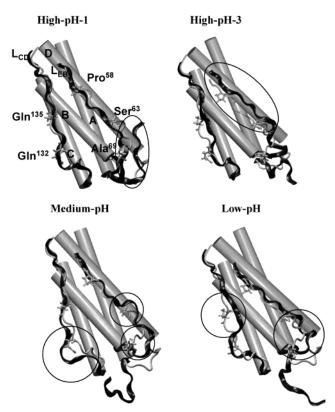


FIGURE 6: Representations of characteristic loop conformations of G-CSF at different protonation states. The X-ray structure (in gray) is also plotted for comparison.

on the other hand, has a quite small  $L_0$ -ex RMSD (1.56 Å), but the all-residue RMSD is significantly larger (2.86 Å). This indicates that the position of the L<sub>0</sub> region has deviated significantly from its initial structure, but the rest of the protein units maintain positions similar to those in the X-ray structure. Finally, the low-pH simulation has the largest average L<sub>0</sub>-ex RMSD (2.03 Å) and all-residue RMSD (3.4 Å).

Comparison of the MD simulations with the *B* factor from X-ray experiments (49) and the order parameter from NMR experiments (50, 51) indicates that the simulations provide an accurate representation of G-CSF structure and dynamics. Details of the comparisons can be found in the Supporting Information.

(C) Conformational Properties at Different pH Values. Titratable and charged residues of G-CSF are mainly contained in helical domains; the two long loops, L<sub>EB</sub> and L<sub>CD</sub>, however, do not contain these groups (Table 5). The change of charge distribution upon changing the protonation configuration affects the interactions among different regions in G-CSF and thus its structure. The change of conformation of G-CSF in the loop regions is shown in Figure 6. Structural changes can be also captured by properties such as hydrogenbonding patterns and helical-helical distances and angles (52).

Hydrogen bonds (HB) not involved in forming the helical structures (see Figure 7) are analyzed, and the results are shown in Table 6. Analyses show that the protonation/ deprotonation of acidic groups significantly affects the helical/helical packing between helices A and C and between helices B and C since acidic residues are more populated in these regions of G-CSF. The change of helical packing also

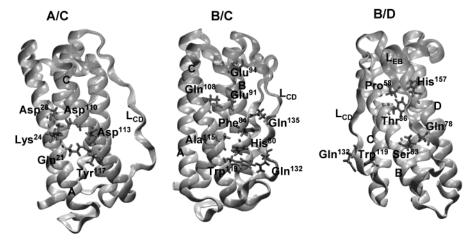


FIGURE 7: Representations of the persistent nonhelical hydrogen bonds of G-CSF.

Table 6: Strengths of Several Persistent (See Text) Nonhelical Hydrogen Bonds of G-CSF $^a$ 

			streng	th	
hydrogen-bonded residues	high- pH-1	high- pH-2	high- pH-3	medium- pH	low- pH
	A/C				
$Gln^{21} (O^{SC}) - Thr^{117} (H^{SC})$	56	0	73	19	56
$Gln^{21}$ (HSC) $-Asp^{113}$ (OSC)	11	41	48	24	0
Lys <sup>24</sup> (H <sup>SC</sup> )-Asp <sup>110</sup> (O <sup>SC</sup> )	94	96	89	96	0
$Asp^{28} (O^{SC}) - Asp^{110} (H^{SC})$	0	0	0	0	21
	B/C				
Glu <sup>94</sup> (OSC)-Gln <sup>108</sup> (HSC)	0	0	0	0	45
Glu <sup>91</sup> (H,O <sup>SC</sup> )-Gln <sup>108</sup> (H,O <sup>SC</sup> )	34	10	19	12	0
His <sup>80</sup> (H <sup>SC</sup> )—Ala <sup>115</sup> (O <sup>BB</sup> )	0	11	0	56	0
	$B/L_{\rm CD}$				
His <sup>80</sup> (H <sup>SC</sup> )-Gln <sup>132</sup> (O <sup>SC</sup> )	15	0	0	0	0
$Phe^{84} (O^{BB}) - Gln^{135} (H^{SC})$	68	77	59	69	39
	C/L <sub>CD</sub>				
$Trp^{119} (H^{SC}) - Gln^{132} (O^{SC})$	20	13	1.5	8	17
	$L_{EB}/D$				
Pro <sup>58</sup> (O <sup>BB</sup> )-His <sup>157</sup> (H <sup>SC</sup> )	4.5	54	0	91	87
	L <sub>EB</sub> /B				
Pro <sup>58</sup> (O <sup>BB</sup> )-Thr <sup>86</sup> (H <sup>SC</sup> )	0	9	20	0	0
$Ser^{63}$ (H,O <sup>BB</sup> )- $Gln^{78}$ (H,O <sup>SC</sup> )	76	70	0	0	5

<sup>&</sup>lt;sup>a</sup> The strength is the percentage of the dynamic trajectory satisfying the 2.4 Å donor/acceptor distance and 150° D−H···A angle criteria (58). The strength of hydrogen bonds is estimated by computing the percentage satisfying the above criteria in the last 1.5 ns of the simulations. Superscript SC denotes side-chain atoms, and superscript BB denotes backbone atoms.

leads to the change of loop conformation such as shown in Figure 6. The protonation/deprotonation of histidine groups, on the other hand, mainly affects the conformation in the loop regions, especially His<sup>80</sup> and His<sup>157</sup>, which form hydrogen bonds with other residues. G-CSF conformation is sensitive to the protonation states of His<sup>80</sup> and His<sup>157</sup>; the results can be seen by the difference in the loop conformation and hydrogen bond strengths as shown in Figure 6 and Table 6 for high-pH-1, high-pH-2, and high-pH-3 configurations in which the net charge of G-CSF remains the same but the protonation states of His<sup>80</sup> and His<sup>157</sup> are different. These analyses show that although the overall tertiary structure of G-CSF is stable across different pH values, there exists a difference in conformation, even for the exposed  $L_0$  region, at different pH values. Calculation of the covariance

fluctuation matrix between  $\alpha$ -carbons, which measures the correlation of motion between each pair of atoms, also shows that there exist different types of long-range correlation in different protonation states. These various types of correlations can be related to the conformational properties of each protonation configuration. A more detailed version of conformational analysis is available in the Supporting Information.

In summary, the molecular dynamics simulations show that altering the protonation states of the titratable residues of G-CSF, located mainly in the helical regions, can lead to distinct conformations for each protonation state. Despite the fact that the up-up-down-down arrangement of the helical domains of G-CSF remains stable, different conformational features, such as hydrogen-bonding patterns, helix-helix packing, and pairwise correlations, can be identified in each protonation configuration even in the most exposed  $L_0$  region. Since our proposed oxidation mechanism suggests that interaction of methionine sites with solvent molecules is governing, the oxidation rate must be a function of the local conformation of solvent and protein surrounding each methionine residue. The detailed nature of these effects is currently being studied in our group. Here we propose that the pH dependence of the conformation of G-CSF is most likely responsible for the observed pH dependence of the rates of oxidation (Figure 3). In the next section, we focus on evaluating why different methionine residues exhibit different rates of oxidation.

(D) Characterization of Solvent Configuration and Correlation with Rates of Oxidation. Differences in the oxidation rates of different methionine residues have been proposed to be due to the different structural environments of individual methionine residues (7, 16, 17). This difference could be due to differences in interactions among residues in a protein and differences in the interaction between the protein and the solvent. A better understanding these interactions would lead to a better understanding of oxidation and how to control it. Moreover, having a useful tool to estimate the relative oxidation rates would be helpful. This property could be taken from the X-ray structure or from molecular dynamics simulations, but as noted above, properly treating thermodynamic fluctuations is important, and it is unlikely that static X-ray data will be sufficient. We address the correlation of structural properties to the oxidation rates of methionine residues in this section.

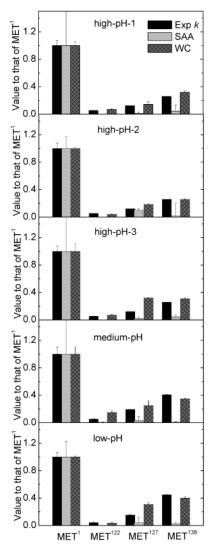


FIGURE 8: Comparison of structural properties to the rates of oxidation of methionine residues in G-CSF. Solvent-accessible area (SAA) of methionine sulfur atoms (1.6 Å probe radius), water coordination number (WCN) of methionine sulfur atoms with a cutoff radius of 5.5 Å, and measured rates of oxidation of methionine residues (Exp) are normalized to the value of Met1 at each pH value. The average values over the last 1.5 ns are shown.

Solvent-accessible area (30) (SAA) is traditionally used to characterize the degree of interaction between solvent molecules and sites in a protein. Since the access of the sulfur atom to oxidants in the solution is required for the oxidation reaction, assuming that the oxidation rate is proportional to SAA is the simplest model. The SAA (probe radius 1.6 Å) of the sulfur atom of methionine determined using the X-ray structure is only nonzero for Met<sup>1</sup> and is zero for Met<sup>122</sup>, Met<sup>127</sup>, and Met<sup>138</sup>. Clearly, SAA's computed from X-ray structure are not good structural properties for describing oxidation.

Comparison of average SAA's of methionine residues computed from molecular dynamics simulation results (last 1.5 ns) to the rates of oxidation at each protonation state is shown in Figure 8. All methionine residues have nonzero average SAA's, suggesting that each methionine residue is liable to oxidation. This result is qualitatively consistent with the experimental results shown earlier. This result also further emphasizes that using the SAA's from the X-ray structure of methionine residues is not sufficient, even qualitatively,

for making reasonable estimates of relative rates of oxidation of G-CSF and that considering the dynamic nature of protein motion is essential. Nevertheless, the SAA's captured from simulation do not show good quantitative agreement with the oxidation rates of methionine residues. Therefore, we conclude that SAA is not a good measure of the susceptibility of a methionine site to oxidation.

Another measure of the degree to which the solvent interacts with sites on the protein is the water coordination number of the sulfur atom (WCN), i.e., the number of water molecules within a sphere of a predefined cutoff radius from the atom. Ab initio calculations indicate that water molecules near the reaction site (the sulfur atom) stabilize the charge separation occurring during the breaking of the O-O bond of hydrogen peroxide in the oxidation process (25). The results discussed below use a cutoff radius of 5.5 Å to compute the WCN. This cutoff radius corresponds to two solvation shells around the sulfur atom (from the pair correlation function of the sulfur atom to the oxygen atoms of water). The choice of including two solvation shells was made on the basis of our proposed oxidation mechanism (25). Results are shown in Figure 8. Adjusting the cutoff radius does not lead to a better correlation between the rate constants and WCN's, also leading to the conclusion that two solvation shells participate in the oxidation processes. Here, the WCN's are normalized to the WCN's of Met<sup>1</sup> at each pH. In the high-pH-1 simulation, the correlation between the WCN's and relative oxidation rates of each methionine residues is very good, providing strong evidence that the rates of oxidation can be characterized by the local conformational and structural environment of each methionine residue. In the high-pH-2 simulation, the agreement is also good, in general, but the WCN is a bit larger than the oxidation rate of Met<sup>127</sup>. In the high-pH-3 simulation, however, the WCN's of Met<sup>127</sup> and Met<sup>138</sup> are in reverse order to the rate constants. We interpret these results to mean that the protonation configuration adapted in the case of high-pH-1 is more representative of the experimental system than those of highpH-2 and high-pH-3 at pH 8.5. In the medium-pH simulation, relative oxidation rates and WCN also correlate well at pH 5.5, but the normalized WCN of Met<sup>122</sup> is larger than the normalized rate of oxidation. In the low-pH simulation, the WCN's of Met<sup>122</sup> and Met<sup>138</sup> correlate well to the experimental data, while the normalized WCN of Met<sup>127</sup> is larger than the normalized rate of oxidation. At all pH values (taking the high-pH-1 simulation as representative of the high-pH case), the order of the normalized rates of oxidation found via experiment, Met<sup>1</sup> > Met<sup>138</sup> > Met<sup>127</sup> > Met<sup>122</sup>, is consistent with the order of the normalized WCN's for the methionine residues of G-CSF. The success in the correlation of the WCN's of sulfur atoms to the relative rates of oxidation and the failure of using SAA's also indicates that taking into account solvent effects in an explicit manner is essential.

Some additional understanding of the relative rates of oxidation can be gained from viewing the results in Figure 8 in light of the protein structure in Figure 4. Met<sup>1</sup> is the most exposed methionine residue and, therefore, has the highest methionine oxidation rate. Met<sup>138</sup> and Met<sup>127</sup> are located in the  $L_{CD}$  region (Figure 4). Met<sup>138</sup> is close to helix B and near the strong  $Trp^{119}-Gln^{132}$  HB region of  $L_{CD}$ (Figures 7 and 6). Met 127 is close to the C terminus of helix

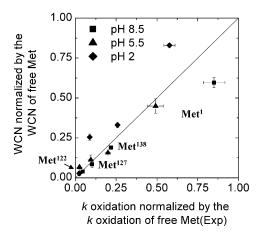


FIGURE 9: Correlation between the water coordination numbers (WCN) of the methionine sulfur atoms (cutoff radius 5.5 Å) and the rates of oxidation of methionine residues of G-CSF at different pH values. The rates of oxidation of methionine residues and methionine sulfur WCN's are normalized to the values of the free methionine amino acid.

C (residues 100–125) and has a slower oxidation rate than Met<sup>138</sup> at all pH values (2–8.5) and also smaller WCN's. Finally, Met<sup>122</sup> is buried in helix C and has the lowest oxidation rate.

A more general comparison between the WCN and the rates of oxidation of methionine residues can be made by normalizing the WCN'S and oxidation rate against the WCN's and oxidation rates of the free methionine residue (free Met). This is done in Figure 9. Note that the WCN's of the free methionine are determined in separate simulations, and the rate of oxidation of the free methionine acid does not show a pH dependence. As shown in Figure 9, the oxidation rate of Met<sup>1</sup> at pH 5.5 is 49% of that of free Met, and similarly, the WCN of Met<sup>1</sup> at that pH is 45% of that versus that of the free Met. This indicates that the sulfur atom of Met<sup>1</sup> is not as exposed as that of free Met. Both the oxidation rates and WCN's of Met1 at different pH values are 49-85% of the corresponding values of free Met. The differences, particularly the minimum at pH 5.5, indicate that the local conformation around Met<sup>1</sup> is also pH dependent. Similarly, both the oxidation rates and WCN's of Met<sup>138</sup> depend on pH (minimum at pH 5.5) and agree well with each other quantitatively. For Met<sup>127</sup>, the pH dependence of oxidation rates is not significant, and the WCN's correlate well with oxidation rates at pH 5.5 and 8.5. The WCN is, however, larger than the normalized rate of oxidation at pH 2. For Met<sup>122</sup>, the pH dependence of oxidation rates is also not significant, and the WCN's correlate well with the oxidation rates.

The above results demonstrate that the rates of oxidation of different Met groups correlate well to the WCN's of different Met groups. However, differences in the rates of oxidation of a particular Met residue as a function of pH can be explained only in part by differences in WCN, mainly because the effects of pH on conformation are subtle, as analyzed briefly above and in more detail in the Supporting Information.

Therefore, WCN can be used as a quantitative measure of the relative susceptibility to oxidation of methionine residues by  $H_2O_2$  in a protein molecule. In this model, the effects of surrounding side-chain residues on the oxidation

reaction are not considered. The fundamental mechanism of oxidation described earlier (25) does not rule out the possibility that surrounding residues of a methionine side chain with hydrogen bond donors and acceptors can play a similar role as solvent molecules, stabilize the charge separation when the O-O bond of H<sub>2</sub>O<sub>2</sub> breaks. Moreover, it has been shown that carboxylic acids can form a reactive intermediate with H<sub>2</sub>O<sub>2</sub> and catalyze the oxidation of various compounds (53, 54). Along the same line, Glu and Asp residues may also have catalytic effects on the oxidation of surrounding methionine residues. However, for methionine residues in G-CSF, the closest Asp or Glu residues are 10 Å away from the Met residues during the MD simulations. Therefore, the Asp and Glu residues in G-CSF likely do not affect the oxidation reaction; as a result, WCN's provide an accurate description on the relative rates of oxidation. The effects of residues with hydrogen bond donors and acceptors, if present, on the rates of oxidation of methionine residues are currently under investigation in our laboratory via ab initio calculations.

Moreover, correlating a property such as WCN to rates of oxidation also assumes that each water molecule surrounding the sulfur atom, on average, stabilizes the oxidation reaction equally. The water molecules within the protein are sometimes bound (55) and therefore may not be as available as the nonbound water molecules to stabilize the oxidation reaction. Such bound water molecules around methionine side chains have not been observed in the MD simulations as estimated by monitoring the duration of each water molecule surrounding the sulfur atoms. This is probably because G-CSF is a protein with four compactly packed helices and does not have water binding sites near methionine residues. Moreover, we did not find a significant difference in the lifetime of a water molecule within 4 Å of the sulfur atom for different methionine groups. A more detailed analysis on the effects of protein environment of the oxidation reaction is ongoing by QM/MM simulations.

## **CONCLUSIONS**

In this work, we measured the rates of oxidation by  $H_2O_2$  of the Met groups as a function of pH. We also obtained detailed information on the local environment of the methionine residues of G-CSF via molecular dynamics simulations. We found that the water coordination number (WCN) of each sulfur atom correlates well with the rate of oxidation. Thus, WCN can be used as a semiquantitative measure of oxidation. Note that solvent-accessible area (SAA), typically used in the analysis of proteins, does not correlate well with the rates of oxidation. Furthermore, static structural data from X-ray measurements cannot provide much insight into the oxidation process; dynamic information, such as that resulting from molecular dynamics (MD) simulations, is needed.

The MD simulations show that the conformation of the protein is significantly different at different pHs. These differences can be characterized via properties such as hydrogen bonding, helix packing, and the covariance fluctuation matrix. Experiments show differences in the rates of oxidation of a Met site with pH. These results, plus the fact that there is no pH dependence of the oxidation rate of free methionine amino acids, suggest that the observed pH dependence of methionine residues in a protein molecule is

very likely to be due to the change of protein conformation as a function of pH. Determining exactly how these changes in conformation affect the rates of oxidation of Met groups is currently under investigation.

#### ACKNOWLEDGMENT

We thank David Brems and Margaret Speed Ricci of Amgen, Inc., for very helpful input into this project.

### SUPPORTING INFORMATION AVAILABLE

A comparison of the MD results with X-ray *B* factors and NMR parameters and analyses of the conformation change of G-CSF as a function of pH. This material is available free of charge via the Internet at http://pubs.acs.org.

#### REFERENCES

- Cleland, J. L., Powell, M. F., and Shire, J. (1993) Crit. Rev. Ther. Drug Carrier Syst. 10, 307–377.
- 2. Wang, W. (1999) Int. J. Pharm. 185, 129-188.
- 3. Cleland, J. L., and Langer, R. (1994) ACS Symp. Ser., 1-19.
- 4. Li, S., Schöneich, C., and Borchardt, R. T. (1995) *Biotechnol. Bioeng.* 48, 490–500.
- 5. Nguyan, T. H. (1994) ACS Symp. Ser., 59-71.
- 6. Shahrokh, Z. (1997) ACS Symp. Ser. 675, 1.
- 7. Meyer, D. J., Ho, B., and Manning, C. M. (2002) in *Rational Design of Stable Protein Formulations* (Carpenter, F. J., and Manning, C. M., Eds.) pp 85–107, Kluwer Academic/Plenum Publishers, New York.
- 8. Matheson, N. R., Wang, P. S., and Travis, J. (1979) *Biochem. Biophys. Res. Commun.* 88, 402–409.
- Yao, Y., Yin, D., Jas, G. S., Kuczera, K., Williams, T. D., Schöneich, C., and Squier, T. C. (1996) *Biochemistry* 35, 2767– 2787.
- Gao, J., Yin, D. H., Yao, Y., Sun, H., Qin, Z., Schöneich, C., Williams, T. D., and Squier, T. C. (1998) *Biophys. J.* 74, 1115– 1134.
- Nabichi, Y., Fujiwara, E., Kuboniwa, H., Asoh, Y., and Ushio, H. (1998) *Anal. Chim. Acta 365*, 301–307.
- Patten, S. M. V., Hanson, E., Bernasconi, R., Zhang, K., Manavalan, P., Cole, E. S., McPherson, J. M., and Edmunds, T. (1999) *J. Biol. Chem.* 274, 10268–10276.
- 13. Levine, R. L. (1983) J. Biol. Chem. 258, 11828-11833.
- 14. Stauffer, P. K., and Etson, D. (1969) J. Biol. Chem. 244, 5333-5338
- Griffiths, S. W., and Cooney, C. L. (2002) J. Chromatogr. A 942, 133–143.
- Griffiths, S. W., and Cooney, C. L. (2002) Biochemistry 41, 6245

  6252.
- Lu, H. S., Fausset, P. R., Narho, L. O., Horan, T., Shinagawa, K., Shimamoto, G., and Boone, T. C. (1999) Arch. Biochem. Biophys. 362, 1–11.
- Stadtman, E. R., and Oliver, C. N. (1991) J. Biol. Chem. 266, 2005–2008.
- 19. Stadtman, E. R. (1992) Science 257, 1220-1224.
- Merker, K., Sitte, N., and Grune, T. (2000) Arch. Biochem. Biophys. 375, 50-54.
- 21. Stadtman, E. R. (1993) Annu. Rev. Biochem. 62, 797-821.
- John, O. E. (1960) in *Peroxide Reaction Mechanisms* (John, O. E., Ed.) pp 67–106, Interscience, New York.
- Dankleff, M. A. P., Ruggero, C., Edwards, J. O., and Pyun, H. Y. (1968) J. Am. Chem. Soc. 90, 3209.

- Bach, R. D., Su, M. D., and Schlegel, H. B. (1994) J. Am. Chem. Soc. 116, 5379-5391.
- 25. Chu, J. W., and Trout, B. L. (2003) J. Am. Chem. Soc. (in press).
- Bach, R. D., Su, M. D., and Schlegel, H. B. (1994) J. Am. Chem. Soc. 116, 5379-5391.
- 27. Amels, R., Elias, H., and Wannowius, K.-J. (1997) *J. Chem. Soc., Faraday Trans.* 93, 2537–2544.
- 28. Pearlman, R., and Bewley, T. A. (1993) Stability and Characterization of Human Growth Hormone, Plenum Press, New York.
- 29. Schöneich, C., Zhao, F., Wilson, G. S., and Borchardt, R. T. (1993) *Biochim. Biophys. Acta 1158*, 307–322.
- 30. Lee, B., and Richards, F. M. (1971) J. Mol. Biol. 55, 379-400.
- 31. Aritomi, M., Kunishima, N., Okamoto, T., Kuroki, R., Ota, Y., and Morikawa, K. (1999) *Nature 401*, 713.
- 32. Aritomi, M., Kunishima, N., Okitsu, N., Shimizu, M., Ota, Y., and Morikawa, K. (2000) *Acta Crystallogr. D56*, 751–753.
- 33. Kaiser, R., and Metzka, L. (1999) Anal. Biochem. 266, 1-8.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* 4, 187–217.
- MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) J. Phys. Chem. B 102, 3586-3616.
- Grubmuller, H. (1996) Theoretical Biophysics Group, Institut fur Medizinische Optik, Ludwig-Maximilians-Universitat Munchen, Munchen, Germany.
- Darden, T., York, D., and Pedersen, L. (1993) J. Chem. Phys. 93, 10089–10092.
- 38. Essmann, U., Perers, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) *J. Chem. Phys.* 103, 8577–8593.
- Allen, M. P., and Tildesley, D. J. (1987) Computer Simulation of Liquids, Oxford, New York.
- 40. Nosé, S. (1991) Prog. Theor. Phys., Suppl. 103, 1-46.
- Chen, Y., and Barkley, M. D. (1998) Biochemistry 37, 9976– 9982.
- 42. Ricci, M. S., Sarkar, C. A., Fallon, E. M., Lauffenburger, D. A., and Brems, D. N. (2003) *Protein Sci.* 12, 1030–1038.
- Narhi, L. O., Kenney, W. C., and Arakawa, T. (1991) J. Protein Chem. 10, 359.
- 44. Bazan, J. F. (1990) Immunol. Today 11, 350-354.
- 45. Reidhaar-Olson, J. F., De Souza-Hart, J. A., and Selick, H. E. (1996) *Biochemistry* 35, 9304–9341.
- Fink, A. L., Calciano, L. J., Goto, Y., Kurotsu, T., and Palleros, D. R. (1994) *Biochemistry 33*, 12504–12511.
- Fink, A. L., Calciano, L. J., Goto, Y., Nishimura, M., and Swedberg, S. A. (1993) *Protein Sci.* 2, 1155–1160.
- Kolvenbach, C. G., Narhi, L. O., Philo, J. S., Li, T., Zhang, M., and Arakawa, T. (1997) J. Pept. Res. 50, 310–318.
- McCammon, J. A., and Harvey, S. C. (1987) Dynamics of Proteins and Nucleic Acids, Cambridge University Press, New York.
- 50. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546–
- Zink, T., Ross, A., Lueers, K., Cieslar, C., Rudolph, R., and Holak, T. A. (1994) *Biochemistry* 33, 8453

  –8463.
- Loof, H. D., Nilsson, L., and Rigler, R. (1992) J. Am. Chem. Soc. 114, 4028–4035.
- Chothia, C., Levitt, M., and Richardson, D. (1981) Mol. Biol. 145, 215–250.
- 54. Case, D. A. (1994) Curr. Opin. Struct. Biol. 4, 285-190.
- Humphrey, W., Dalke, A., and Scholten, K. (1996) J. Mol. Graphics 14, 33–38.
- Kabsch, W., and Sander, C. (1983) *Biopolymers 12*, 2577–2637.
   BI0356000