Mutation Choice to Eliminate Buried Free Cysteines in Protein Therapeutics

XUE XIA, LIAM M. LONGO, MICHAEL BLABER

Department of Biomedical Sciences, Florida State University, Tallahassee, Florida 32306-4300

Received 4 August 2014; revised 3 September 2014; accepted 9 September 2014

Published online 13 October 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.24188

ABSTRACT: Buried free-cysteine (Cys) residues can contribute to an irreversible unfolding pathway that promotes protein aggregation, increases immunogenic potential, and significantly reduces protein functional half-life. Consequently, mutation of buried free-Cys residues can result in significant improvement in the storage, reconstitution, and pharmacokinetic properties of protein-based therapeutics. Mutational design to eliminate buried free-Cys residues typically follows one of two common heuristics: either substitution by Ser (polar and isosteric), or substitution by Ala or Val (hydrophobic); however, a detailed structural and thermodynamic understanding of Cys mutations is lacking. We report a comprehensive structure and stability study of Ala, Ser, Thr, and Val mutations at each of the three buried free-Cys positions (Cys16, Cys83, and Cys117) in fibroblast growth factor-1. Mutation was almost universally destabilizing, indicating a general optimization for the wild-type Cys, including van der Waals and H-bond interactions. Structural response to Cys mutation characteristically involved changes to maintain, or effectively substitute, local H-bond interactions—by either structural collapse to accommodate the smaller oxygen radius of Ser/Thr, or conversely, expansion to enable inclusion of novel H-bonding solvent. Despite the diverse structural effects, the least destabilizing average substitution at each position was Ala, and not isosteric Ser. © 2014 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci 104:566–576, 2015

Keywords: protein aggregation; protein structure; stability; crystal structure; thermodynamics; protein engineering; protein half-life; hydrogen bond; fibroblast growth factor-1

INTRODUCTION

Among the least abundant amino acids found in proteins,[1] cysteine (Cys) satisfies a number of unique and critically important structural and functional roles—including disulfide bonds, structurally important H-bonds, metal coordination, activity regulation via posttranslational modification, and redox-active as well as non-redox catalysis.[1,2] A statistical analysis of Cys residues from a set of 378 non-redundant protein X-ray structures reported that Cys occupies 0.73% of all positions.[3] Free-Cys residues (i.e., those not involved in disulfide bonds) are approximately evenly distributed between solvent-exposed and buried positions within proteins.[1] Free-Cys residues have a reactive sulfhydryl group, which in buried positions is structurally protected from redox reactions that can result in thiol adducts or other oxidized derivatives. However, because of the dynamic nature of protein structures (e.g., transient global unfolding or localized structural dynamics), buried free-Cys residues can become exposed and undergo redox chemistry. Oxidized derivatives of free-Cys residues, including formation of thiol adducts or cysteic acid, can effectively block refolding—resulting in an irreversible pathway that, by Le Chatelier's principle, will continuously drive protein unfolding. Consequently, buried free-Cys residues can play a key role in regulating protein functional half-life.[4–8]

The presence of buried free-Cys residues in protein biopharmaceuticals, notably monoclonal antibodies, is problematic because they can effectively limit shelf-life as well as induce immunogenic aggregates.[9–12] A recent study of IgG1 and IgG2 antibodies showed that roughly one third of recombinant antibody molecules from a standardized industrial preparation contain a single open disulfide bond (that is, two free-Cys residues that should form a disulfide bond in the native state structure).[9,12] In addition, all free-Cys residues within monoclonal IgGs are buried in the hydrophobic core and mediate aggregation in response to agitation stress (as can occur during shipment).[12] Cys→Ser mutation is a common approach to substitute Cys because Ser is isosteric (substituting the Ser Oγ for the Cys Sγ). Several engineered protein therapeutics have successfully introduced Cys→Ser mutations to improve shelf life, including interleukin-2 (Proleukin®), interferon β1b (Betaseron®), and granulocyte colony-stimulating factor (Neulasta™).[13] Studies have shown that while Cys→Ser mutations in fibroblast growth factor-1 (FGF-1) can significantly reduce protein stability, the mutant proteins nonetheless exhibit increased functional half-life.[4,5,8] However, Cys residues have a comparatively high hydrophobicity index, similar to Met, Phe, and Tyr, which suggests chemical compatibility if buried within the hydrophobic core environment.[3] Thus, it is not clear whether Cys→Ser mutation is the best *de facto* substitution to effectively eliminate buried free-Cys residues, or whether small hydrophobic amino acids such as Ala or Val (isovolumetric to Cys) might be more appropriate substitutions. Indeed, an analysis of statistically preferred mutations of Cys among related proteins (e.g., BLOSUM45, NCBI) identifies Ala, Ser, Thr, and Val as the most common naturally substituted amino acids. An

Abbreviations used: PDB, protein data bank; NTA, nitrilotriacetic acid; GuHCl, guanidine hydrochloride; RMSD, root-mean-square deviation; WT, wild-type.

Correspondence to: Michael Blaber (Telephone: +850-644-3361; Fax: +850-644-5781; E-mail: michael.blaber@med.fsu.edu)

This article contains supplementary material available from the authors upon request or via the Internet at http://onlinelibrary.wiley.com/.

The three letter amino acid code is utilized throughout; while the single letter amino acid code is utilized in figures.

Journal of Pharmaceutical Sciences, Vol. 104, 566–576 (2015)

important unresolved question in engineering protein therapeutics is whether there is a universally preferred substitution for buried free-Cys residues, or whether each situation must be considered as uniquely different.

We address this question by performing detailed structure and thermodynamic analyses of Ala, Ser, Thr, and Val mutations at three buried free-Cys positions (Cys16, Cys83, Cys117; numbering scheme of the 140 amino acid form) in FGF-1. FGF-1 has therapeutic potential to accelerate wound healing but has low thermostability and a proclivity to unfold, form mixed thiols, and aggregate.[14–18] The side chains of Cys16 and Cys83 are fully buried in the protein interior, whereas Cys117 exhibits two conformers in which one is buried and the other is partially exposed to solvent (~12 Å^2 accessible surface area).[19,20] In a previous report, Cys83→Ala, Ser, Thr, and Val substitutions were constructed in FGF-1 and subjected to both X-ray structure and thermodynamic analyses.[5] In the present report, Ala, Ser, Thr, and Val substitutions are characterized at positions Cys16 and Cys117. Combined with the prior Cys83 data, the present results provide insight into the diverse structural and stability properties of buried free-Cys residues. The results show that there is a general structural optimization to accommodate Cys in the wild-type (WT) protein. Such optimization is described by a set of local H-bond interactions involving the buried Cys, combined with a general structural rigidity, such that all of the (longer) H-bond interactions involving the WT Cys Sγ cannot be effectively maintained by the Oγ in Ser or Thr mutations. Consequently, the protein response to Cys mutation follows one of two general solutions in an attempt to satisfy local H-bond interactions: the first is a limited structural collapse able to maintain a subset of the local H-bond distances, and the other is expansion to permit introduction of a buried solvent to serve as a novel H-bond partner. In this regard, expansion appears to be more energetically costly as regards mutant protein stability. Although the most stable mutation was different for each of the three buried Cys positions in FGF-1, the best average mutant as regards thermostability was Ala.

MATERIALS AND METHODS

Materials

All chemicals used in this work were of reagent grade or higher quality. Mutations were performed using the QuikChange™ site-directed mutagenesis protocol (Agilent Technologies, Santa Clara, California). Ni-NTA affinity resin was purchased from Qiagen (Valencia, California) and heparin Sepharose resin was purchased from GE Life Sciences (Pittsburgh, Pennsylvania). Flourescence spectroscopy data were collected on a Varian Eclipse spectrophotometer (Varian Medical Technologies, Palo Alto, California). Crystallization reagents and cryoloops were purchased from Hampton Research (Aliso Viejo, California).

Methods

Protein Mutagenesis and Expression

A codon-optimized synthetic gene encoding the 140 amino acid form of human FGF-1 with an N-terminal six His tag was cloned into the pET21a(+) expression vector (the His tag has shown no influence upon stability and mitogenic activity).[21] Residue Cys16 and Cys117 were mutated to Ala, Ser, Thr, or Val using the QuikChange™ site-directed mutagenesis protocol (Agilent Technologies), and each mutation was confirmed by DNA sequencing (Biomolecular Analysis Synthesis and Sequencing Laboratory, Florida State University). Recombinant WT FGF-1 and mutant proteins were expressed from pET21a(+)/BL21(DE3) *Escherichia coli*. After induction with 1 mM isopropyl-β-D-thiogalactoside, the incubation temperature was decreased from 37°C to 16°C for overnight expression. The expressed protein was purified utilizing sequential column chromatography on Ni-NTA affinity resin (Qiagen) and heparin Sepharose resin (GE Life Sciences). Purified protein was buffer exchanged into "crystallization buffer" (50 mM sodium phosphate, 100 mM sodium chloride, 10 mM ammonium sulfate, and 2 mM dithiothreitol pH 7.5) using an 8 kDa molecular weight cutoff membrane tubing (Spectrum Industries Inc., Chippewa Falls, Wisconsin). The resulting proteins were judged to be >98% pure as determined by gel densitometry of Coomassie blue stained SDS-PAGE.

Isothermal Equilibrium Denaturation

Fibroblast growth factor-1 contains a single buried tryptophan residue (Trp107) whose fluorescence is internally quenched in the native state, and quenching is subsequently released upon denaturation, thereby providing a spectroscopic probe of unfolding. Isothermal equilibrium denaturation by guanidine hydrochloride (GuHCl) was quantified by probing fluorescence intensity. WT FGF-1 and mutant protein samples were equilibrated overnight in crystallization buffer at 298 K in 0.1 M increments of GuHCl with final protein concentration of 5 μM. Triplicate scans were collected and averaged; buffer background was collected and subtracted from the protein scans. Fluorescence scans were integrated to quantify the total fluorescence as a function of denaturant concentration. Data were analyzed using a six parameter two-state model[22]:

$$F = \frac{F_{0\mathrm{N}} + S_{\mathrm{N}}[D] + (F_{0\mathrm{D}} + (S_{\mathrm{D}}[D]))\, e^{-(\Delta G_0 + m[D])/RT}}{1 + e^{-(\Delta G_0 + m[D])/RT}}$$

where $[D]$ is the denaturant concentration, $F_{0\mathrm{N}}$ and $F_{0\mathrm{D}}$ are the 0 M denaturant intercepts for the native and denatured state baselines, respectively, and S_{N} and S_{D} are the slopes of the native and denatured state baselines, respectively. ΔG_0 and m describe the linear function of the unfolding free energy versus denaturant concentration. The effect of a given mutation upon the stability of the protein ($\Delta\Delta G$) was calculated by taking the difference between the C_{m} values for WT and mutant proteins and multiplying by the average of the m values, as described by Pace and Scholtz[23]:

$$\Delta\Delta G = (C_{\mathrm{m\,WT}} - C_{\mathrm{m\,mutant}})(m_{\mathrm{WT}} + m_{\mathrm{mutant}})/2$$

where C_{m} is the denaturant concentration at the midpoint of denaturation, and m is the slope of $\Delta G_{\mathrm{unfolding}}$ as a function of denaturant. A negative value of $\Delta\Delta G$ indicates the mutation is stabilizing in comparison to the reference (i.e., WT) protein.

Crystallization, X-ray Data Collection and Structural Refinement of FGF-1 Mutant Proteins

Fibroblast growth factor-1 mutant protein in crystallization buffer was concentrated to 6–10 mg/mL for crystallization trials. Crystals were grown in a 25°C incubator using the hanging drop vapor diffusion method (combining 5.5 μL of