# Site-Specific Protein PEGylation

## Application to Cysteine Analogs of Recombinant Human Granulocyte Colony-Stimulating Factor

**Mary S. Rosendahl, Daniel H. Doherty, Darin J. Smith, Alison M. Bendele, and George N. Cox**

Granulocyte colony-stimulating factor (G-CSF) is a 19-kDa glycoprotein that stimulates proliferation, maturation, and functional activation of cells in the granulocyte lineage. Recombinant G-CSF is widely used to ameliorate neutropenia resulting from myelosuppressive chemotherapy and bone marrow transplantation as well as to mobilize peripheral blood progenitor cells for transplantation and blood banking (**1**, **2**). G-CSF has a short circulating half-life, which necessitates that the protein be administered to patients through daily subcutaneous injection. Development of modified forms that last longer in vivo — and thus can be administered less frequently — is of significant interest to patients and healthcare providers.

**PRODUCT FOCUS:** PROTEIN DRUGS WITH SHORT IN-VIVO HALFLIVES

**PROCESS FOCUS:** PRODUCTION, PURIFICATION, AND TESTING

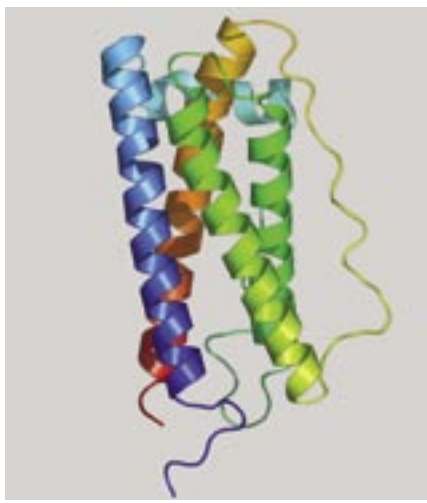
**WHO SHOULD READ:** PRODUCT DEVELOPMENT SCIENTISTS, ANALYTICAL LABORATORY TECHNICIANS

**KEYWORDS:** PEGYLATION, GENETIC ENGINEERING, PRECLINICAL TESTING, CONJUGATE PROTEINS

**LEVEL:** ADVANCED

Covalent modification of G-CSF with polymers such as polyethylene glycol (PEG) increases its circulating half-life in animals and humans (**3–6**). Covalent attachment of PEG to a protein increases that protein's effective size and reduces its rate of clearance from the body. Previous studies have used amine-reactive PEGs to modify G-CSF at exposed amine groups on lysine residues and the N-terminal amino acids. Such an approach is not optimal for G-CSF, however, because the protein contains four lysine residues in addition to its N-terminal amino acid, and those lysine residues are located in regions implicated in receptor binding (**7–10**).

Not surprisingly, modification of G-CSF with amine-reactive PEG reagents reduces in vitro biological activity of the protein by three- to 50-fold, depending on the number and sizes of attached PEG molecules (**3**, **4**). Loss of in vitro bioactivity is greatest when G-CSF is modified with large PEGs — e.g., 20-kDa PEGs, which are most useful for extending the protein's half-life. Amine-PEGylated G-CSF also is heterogeneous, occurring as a complex mixture of at least four isoforms and multiple molecular weight species, all of which have different specific activities (**3–5**). Less heterogeneous PEG–G-CSF conjugates with less severe loss of in vitro bioactivity can be prepared by preferential attachment of PEG to the N-terminus of G-CSF (**5**, **6**, **11**).



**Model of the three dimensional structure of human G-CSF: Alpha helices are indicated by the spiral shaped ribbons.** ADAPTED FROM ARITOMI M, ET AL. ATOMIC STRUCTURE OF THE GCSF–RECEPTOR COMPLEX SHOWING A NEW CYTOKINE-RECEPTOR RECOGNITION SCHEME. *NATURE* 401, 1999: 713–717.

An alternative method for PEGylating proteins covalently attaches PEG to cysteine residues using cysteine-reactive PEGs such as maleimide PEGs. At near-neutral pH values, these PEG reagents selectively attach to the thiol groups of "free" cysteine residues (those not involved in disulfide bonds). Because most proteins contain no surface-exposed free cysteine residues, attachment of a PEG molecule can be targeted to a specific site by introducing a free cysteine residue into a protein using site-directed mutagenesis followed

## Genetic Engineering: Constructing the Cysteine Analogs

Plasmid pBBT165, which contains a cDNA-encoding human G-CSF has been described previously (**13**). For our purposes, DNA encoding the leader sequence of the *Escherichia coli* heat-stable enterotoxin (STII) gene (**14**) was fused to the coding sequence for mature G-CSF using polymerase chain reaction (PCR) methods (**15, 16**). We added a TAA stop codon following the carboxy-terminal residue P174. Codons for prolines at positions 2, 5, and 10 were all changed to CCG; the arginine codon at position 22 was changed to CGT; the leucine codon at position 15 was changed to CTG; and the leucine codon at position 18 was changed to CTC. G-CSF (C17S) and G-CSF (C17S) cysteine analogs were constructed using site-directed PCR-based mutagenesis (**15, 16**). We confirmed that the DNA sequences for G-CSF and G-CSF cysteine analogs were correct. STII-G-CSF, STII-G-CSF (C17S), and STII-G-CSF (C17S) cysteine analog genes were cloned into plasmid pCYB1 (from New England BioLabs, www.neb.com) for expression in *E. coli*.

**Periplasmic Expression of Wild-Type G-CSF and G-CSF (C17S) Proteins in *E. coli*:** Intracellular expression of human G-CSF in *E. coli* yields a protein with an N-terminal methionine residue not present in the natural human version. To express a human G-CSF protein without that nonnatural methionine, we fused the mature coding sequence of G-CSF to the *E. coli* STII signal sequence. G-CSF contains five cysteine residues that form two disulfide bonds (C36–C42 and C64–C74). The fifth, C17, is not involved in a disulfide bond (**18**). An identical construct was prepared for a G-CSF variant in which C17 is changed to serine (C17S). Both proteins were expressed under control of the IPTG-inducible *tac* promoter of plasmid pCYB1 in *E. coli* strain W3110.

Following overnight induction at 28 °C, SDS-PAGE of induced G-CSF wild type and C17S cultures showed expression of a new 19-kDa protein. Western blot analyses revealed that it reacted with an antihuman G-CSF antiserum and comigrated with a commercial *E. coli*-expressed human met–G-CSF standard (data not shown). The latter result suggested that the STII leader peptide had been removed, which is consistent with G-CSF having been secreted to the periplasm. Removal of the STII signal sequence was confirmed by N-terminal protein sequencing studies. We subjected the induced cells to osmotic shock (**17**) and separated the soluble periplasmic proteins from cytoplasmic, insoluble periplasmic, and cell-associated components by centrifugation. Most G-CSF proteins remained associated with the osmotic shock pellets, indicating that the secreted G-CSF molecules were insoluble or membrane-associated (data not shown).

by modification of the added residue with a maleimide-PEG — known as site-specific PEGylation (**12**). If the target site is nonessential, PEGylation will not significantly alter in vitro bioactivity of the protein. The resulting conjugate will have a homogeneous structure, which simplifies analytical characterization of the product. In addition, protein structure and function information can be used to aid in the rational design of PEG–protein conjugates.

### Materials and Methods

After pCYB1 plasmids encoding the various G-CSF proteins (as described in the "Genetic Engineering" box) were transformed into *E. coli* strain W3110, saturated overnight cultures were diluted to an optical density at 600 nm of ~0.05 in Luria broth containing 100 µg/mL ampicillin. When the optical densities of the cultures reached 0.5–0.7, we added isopropyl-ß-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and incubated the cultures overnight for 16 hours at 28 °C. In some experiments, we subjected the cells to osmotic shock (**17**) to release soluble periplasm contents. For most experiments, cells were pelleted by centrifugation and stored at –80 °C.

Cell pellets from 400-mL cultures were treated with 5 mL of B-PER bacterial protein extraction reagent (Pierce Chemical Company, www.piercenet.com) according to the manufacturer's protocols. We treated this mixture with lysozyme (200 µg/mL) for 10 min, then added $MgCl_2$ (10 mM final concentration) and protease-free DNAse (2 µg/mL). Insoluble material was collected by centrifugation, washed with water, and recentrifuged. The resulting pellets were dissolved in 20 mL of 8-M urea with 25-mM cysteine in 20-mM Tris base. The mixture was stirred for 30 min at room temperature, then diluted into 100 mL of 40-mM sodium phosphate with 40-µM copper sulfate and 15% glycerol at pH 8.0.

After storing it for two days at 4 °C, we adjusted the pH of the refold mixture to 4.0 using HCl and centrifuged the mixture. Supernatant was loaded onto a 5-mL S-Sepharose HiTrap column (Amersham Biosciences, now GE Healthcare, www.amershambiosciences.com) equilibrated in 40-mM sodium phosphate at pH 4.0 (Buffer A). Bound proteins were eluted with a linear salt gradient of 0–100% Buffer B (500-mM NaCl, 40-mM sodium phosphate, pH 4.0). Column fractions were analyzed by nonreducing sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing G-CSF proteins and no visible impurities were pooled. We measured in vitro bioactivities of the G-CSF proteins using the

Table 1: In vitro bioactivities of G-CSF, G-CSF (C17S) and G-CSF (C17S) cysteine analogs

| G-CSF Protein | Added Cysteine Location [a] | Yield (mg) [b] | $EC_{50}$ (pg/mL) [c] |
|---|---|---|---|
| met-G-CSF [d] | - | - | 18.6 ± 6.6 |
| G-CSF [e] | - | 1.1 | 10.2 ± 1.6 |
| G-CSF (C17S) | - | 3.3 | 7.2 ± 2.0 |
| *-1C / C17S | N-terminus | 2.5 | 7.0 ± 1.3 |
| T1C / C17S | N-terminus | 2.5 | 7.8 ± 2.9 |
| L3C / C17S | Preceding A Helix | 2.7 | 8.0 ± 2.2 |
| A6C / C17S | Preceding A Helix | 2.9 | 8.2 ± 3.3 |
| S7C / C17S | Preceding A Helix | 2.7 | 7.3 ± 3.2 |
| E93C / C17S | B–C loop | 0.2 | 7.6 ± 0.9 |
| A129C / C17S | C–D loop | 2.5 | 6.0 ± 0.0 |
| T133C / C17S | C–D loop | 1.9 | 6.6 ± 1.2 |
| A136C / C17S | C–D loop | 0.5 | 8.3 ± 1.3 |
| A139C / C17S | C–D loop | 3.1 | 5.2 ± 0.3 |
| A141C / C17S | C–D loop | 2.3 | 8.9 ± 1.1 |
| Q173C / C17S | Following D Helix | 3.5 | 6.2 ± 1.3 |
| *175C / C17S | C-terminus | 2.9 | 5.6 ± 0.4 |

[a] G-CSF structure from Ref. 19

[b] Amount of protein recovered from 400 mL *E. coli* culture after refolding and S-Sepharose column chromatography

[c] Means ± standard deviation of at least three assays for each protein

[d] met-G-CSF standard (from R&D Systems, Inc.)

[e] Wild-type G-CSF prepared by us

Table 2: In vitro bioactivities of PEGylated G-CSF cysteine analogs

| G-CSF Analog | % PEGylation Efficiency [a] | $EC_{50s}$ (pg/mL) [b] 5-kDa PEG | $EC_{50s}$ (pg/mL) [b] 20-kDa PEG |
|---|---|---|---|
| *-1C / C17S | 60 | 5.6 ± 0.3 | |
| T1C / C17S | 45 | 7.0 ± 1.0 | |
| L3C / C17S | 45 | 5.5 ± 0.6 | 8.8 ± 1.0 |
| A6C / C17S | 26 | 6.9 ± 1.0 | |
| A129C / C17S | 23 | 7.1 ± 3.4 | |
| T133C / C17S | 51 | 7.4 ± 3.1 | 9.0 ± 2.9 |
| A136C / C17S | 34 | 6.9 ± 1.1 | |
| A139C / C17S | 34 | 6.8 ± 2.8 | |
| A141C / C17S | 38 | 7.1 ± 0.6 | 9.3 ± 3.8 |
| Q173C / C17S | 50 | 7.0 ± 2.6 | 11.0 ± 1.8 |
| *175C / C17S | 52 | 11.0 ± 1.0 | 14.0 ± 2.3 |

[a] Percent PEGylated protein estimated by S-Sepharose column chromatography analysis of 5-kDa PEG PEGylation reactions

[b] Means ± standard deviation of at least three assays for each protein


murine NFS-60 cell line essentially as described in the literature (13). Serial threefold dilutions of the protein samples were prepared and analyzed in triplicate.

**Preparation of PEGylated G-CSF (C17S) Cysteine Analogs:** The analogs were incubated at pH 8.5 and room temperature for an hour, with a 10-fold molar excess of TCEP (Tris[2-carboxyethylphosphine]hydrochloride) from Pierce Chemical Company and a 20-fold molar excess of 5-kDa or 20-kDa maleimide-PEG from Nektar Therapeutics, Inc. (www.nektar.com). Wild-type and C17S G-CSF were exposed to the same reaction conditions. At the end of the incubation period, we diluted each PEGylation mixture 10-fold with 40-mM sodium phosphate (monobasic) and adjusted the pH to 4.0. PEGylated protein was separated from non-PEGylated protein by S-Sepharose column chromatography under the conditions described above. Purified PEGylated proteins contained less than 0.1 endotoxin units per mg, as determined using QCL-1000 *Limulus* amebocyte lysate kits from BioWhittaker (now Cambrex BioProducts, www.cambrex.com/content/bioscience/bioProducts.asp).

**Biochemical Methods:** N-terminal protein sequencing was performed by Macromolecular Resources at Colorado State University (http://mmr.bmb.colostate.edu) using an Edman degradation procedure. Protein samples were analyzed by SDS-PAGE on precast Tris–glycine gels from Invitrogen Corporation (www.invitrogen.com) and stained with Coomassie Blue. Proteins were reduced by addition of 2-mercaptoethanol to a final concentration of 1–5% (V/V) in sample buffer. We determined protein concentrations using a Bradford assay kit from Bio-Rad Laboratories (www.bio-rad.com), using bovine serum albumin as our protein standard. Western blots were performed using standard procedures.

The primary antibody we used was a polyclonal goat anti-human G-CSF antisera from R&D Systems, Inc. (www.rndsystems.com), and the secondary antibody was an alkaline phosphatase–conjugated rabbit anti-goat IgG antisera from Pierce Chemical Company. Western blots were developed using an NBT/BCIP staining kit from Promega Corporation (www.promega.com). We obtained the human met-G-CSF standard (*E. coli*-expressed) from R&D Systems.

**Animal Experiments:** Our animal experiments were performed with approval of Bolder BioPATH's institutional animal care and use committee. In groups of three, male Sprague Dawley rats weighing approximately 350 g each received a single intravenous injection (in the lateral tail vein) of recombinant wild-type G-CSF or 20-kDa PEG-L3C at a dose of 100 µg protein/kg. At selected time points (0.25, 1.5, 4, 8, 12, 16, 24, 48, 72, 96, and 120 hours postinjection), we drew blood samples (0.3–0.4 mL) from the rats into EDTA anticoagulant tubes. A portion of each sample was used for complete blood cell count analysis (performed by IDEXX Laboratories, www.idexx.com). The remainder was centrifuged so its plasma could be frozen at –80 °C.

A predose baseline sample was obtained for each rat 24 hours before injecting the test compounds. We

measured the plasma protein concentrations using human G-CSF ELISA kits from R&D Systems. Serial dilutions of plasma samples from one animal in each test group were analyzed by in vitro bioassay to identify dilutions that fell within the ELISA's linear range (39–2500 pg/mL). Duplicate samples of appropriate dilutions from all rats were then tested. We analyzed pharmacokinetic parameters using the WinNonlin software package from Pharsight Corporation (www.pharsight.com). Then we compared statistical analyses between samples (blood counts) using a student's T-test with the significance set at $p \leq 0.05$.

## Results

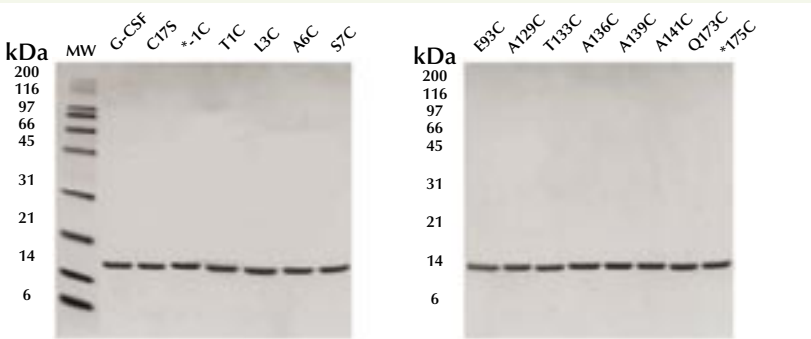
*E. coli* cells expressing the G-CSF wild type and C17S proteins were treated with a mild detergent mixture (B-PER) to release their soluble components. Particulate material was collected by centrifugation, and the insoluble proteins were denatured, reduced, and refolded as described above. We purified the refolded G-CSF proteins to greater than 90% purity by S-Sepharose column chromatography. Both wild-type and C17S G-CSF eluted as single peaks at about 300-mM NaCl. Nonreducing SDS-PAGE analysis of column fractions across the major peaks showed a single main refold species migrating at 19 kDa for both proteins. We pooled fractions containing the major refold species and no visible impurities.

Final yields of G-CSF wild type and C17S after the column step were about 1.1 mg and 3.3 mg, respectively, from 400 mL of culture (Table 1). Purified wild-type G-CSF and C17S comigrated in reducing and nonreducing SDS-PAGE (Figure 1). Reduced G-CSF migrates with a slightly higher apparent molecular mass than nonreduced (19 kDa and 17 kDa, respectively) due to reduction of the protein's two native disulfide bonds. Wild-type G-CSF and C17S had mean $EC_{50}$ potencies of 10 and 7 pg/mL, respectively, in an in vitro cell proliferation assay using mouse NFS-60 cells (Figure 2, Table 1). C17S was reproducibly 1.5- to twofold more potent than the wild type and two- to threefold more potent than a commercial met–G-CSF standard (Table 1).

### Cysteine Mutants

Five cysteine analogs were constructed in the amino-terminal region preceding Helix A: *-1C (the addition of a cysteine residue preceding the natural amino terminus), T1C, L3C, A6C, and S7C. One cysteine analog was constructed in the B–C loop (E93C). Five cysteine analogs were constructed in the C–D loop: A129C, T133C, A136, A139C, and A141C. Two cysteine analogs were constructed in the carboxy-terminal region following Helix D: Q173C and *175C (the addition of a cysteine residue following the natural carboxy-terminus). The G-CSF cysteine analogs were constructed in a C17S background to prevent any difficulties/ambiguities that could be caused by the unpaired C17 residue in wild-type G-CSF.

**Figure 1:** Nonreducing SDS-PAGE analysis of column pools for purified recombinant wild-type G-CSF, G-CSF (C17S), and G-CSF (C17S) cysteine analogs. MW indicates molecular weight markers.

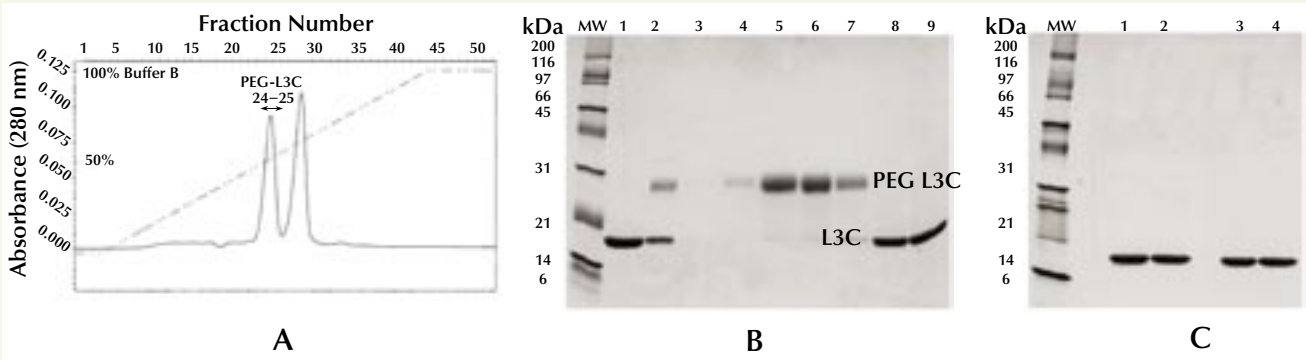



**Construction and Characterization of G-CSF (C17S) Cysteine Analogs:** G-CSF has a compact globular structure, comprising four alpha helices joined by loops (19). The four alpha helices are labeled A–D beginning from the N-terminus. The loop regions are referred to by the helices they join (e.g., the A–B loop joins helices A and B). We constructed 13 mutant G-CSF genes, each of which contains a single cysteine substitution or addition. The "Cysteine Mutants" box provides more detail. Those cysteine analogs were expressed, refolded, and purified using the protocols described for wild type G-CSF and G-CSF (C17S).

Protein recoveries for most of the C17S cysteine analogs after an S-Sepharose column step were similar to the recovery of C17S itself and greater than the recovery of wild-type G-CSF (Table 1). Figure 1 shows a nonreducing SDS-PAGE analysis of the 13 purified cysteine analogs. They were recovered predominantly as monomers that comigrated with wild-type G-CSF and C17S under reducing and nonreducing conditions. All but one eluted from the ion-exchange column at a salt concentration (300–325 mM NaCl) similar to wild-type G-CSF and C17S. The one exception, E93C, eluted at about 400-mM NaCl. In vitro biological activity ($EC_{50}$ potencies) of the 13 were indistinguishable from that of G-CSF (C17S) in the NFS-60 cell proliferation assay (Table 1). Figure 2 shows examples of dose-response curves for representative cysteine analogs.

**Preparation and In Vitro Bioactivity of PEGylated G-CSF Cysteine Analogs:** We tested aliquots of the purified G-CSF cysteine reacting with a 5-kDa maleimide-PEG under the conditions described above. We included the TCEP reducing agent in the PEGylation reaction to partially reduce the protein and thus expose

**Figure 3:** Purification of PEG-L3C by S-Sepharose column chromatography. Panel A shows the S-Sepharose column chromatogram. Panel B is nonreducing SDS-PAGE of purified L3C (lane 1), the L3C PEGylation reaction (lane 2), S-Sepharose fractions 22–26 (lanes 3–7), and fractions 28–29 (lanes 8–9). Fractions 24 and 25 (lanes 5–6) containing PEG-L3C were pooled. Fractions 28–29 (Lanes 8–9) contain unreacted L3C. Panel C shows nonreducing SDS-PAGE analysis of G-CSF and G-CSF (C17S) subjected to the identical PEGylation reaction conditions used for L3C. G-CSF (lane 1), G-CSF PEGylation reaction (lane 2), G-CSF (C17S) (lane 3), and G-CSF (C17S) PEGylation reaction (lane 4). MW indicates molecular weight markers.




the free cysteine residue, which forms a mixed disulfide with cysteine during refolding. Control experiments indicated that the cysteine analogs needed to be partially reduced to PEGylate efficiently (data not shown). PEGylation reactions were terminated by diluting the samples and adjusting the pH to 4.0. Figure 3B (lane 3) shows nonreducing SDS-PAGE analysis of one representative PEGylation reaction using the L3C protein. The only detectable PEGylated species is the mono-PEGylated-L3C protein, which migrates with an apparent molecular mass of 28 kDA. Wild-type G-CSF and C17S were not modified under identical conditions (Figure 3C).

PEGylated L3C was separated from non-PEGylated protein and excess PEG reagent by S-Sepharose column chromatography (Figure 3A). Chromatograms show two major protein peaks eluting at about 275-mM NaCl and 300–325 mM NaCl. SDS-PAGE analysis indicated that the early eluting major peak was mono-PEGylated L3C, and the later-eluting major peak was unreacted L3C protein (Figure 3B). Fractions containing predominantly mono-PEGylated protein (fractions 24 and 25, lanes 5–6 in Figure 3B) were pooled and used for bioactivity measurements of 5-kDa PEG-L3C. The other G-CSF cysteine analogs were PEGylated and purified by the same protocol. All reacted with the PEG reagent to yield mono-PEGylated proteins. Relative PEGylation efficiencies were estimated from the chromatogram peaks and varied 23–60%, with most cysteine analogs being in the 30–60% range (Table 2). Efficiency can be increased by optimizing reaction conditions (data not shown).

Five cysteine analogs (L3C, T133C, A141C, Q173C and *175C) were also modified with a 20-kDa maleimide-PEG, and they eluted when purified at about 250-mM NaCl. Figure 4 shows nonreducing SDS-PAGE analyses of the purified PEGylated cysteine analogs. We subjected the 20-kDa PEG-L3C protein to N-terminal amino acid sequencing, which yielded the sequence TPXGPAS (X indicates a blank). That matches the N-terminal sequence of mature G-CSF (**20**), indicating that the STII signal sequence was correctly processed. The presence of a blank at the third position is consistent with a cysteine and/or PEG-cysteine located at that position (the third amino acid in wild-type G-CSF is leucine).

We measured in vitro biological activities of 11 purified 5-kDa PEG-G-CSF (C17S) cysteine analogs and five purified 20-kDa PEG-G-CSF (C17S) cysteine analogs using the NFS-60 cell proliferation assay. $EC_{50}$ potencies for the PEGylated C17S cysteine analogs ranged 5–12 pg/mL and were comparable to those of both G-CSF (C17S) and wild-type G-CSF (Table 2). Figure 5 provides dose-response curves for representative PEG-G-CSF (C17S) cysteine analogs.

**Pharmacokinetic and Efficacy Experiments in Rats:** We determined the circulating halflives of both the 20-kDa PEG-L3C protein and wild-type G-CSF following intravenous administration to rats at a dose of 100 µg of protein/kg (Figure 6A). Plasma protein levels were measured by ELISA 0.25–120 hours post-injection. Wild-type G-CSF cleared rapidly, reaching undetectable levels by 24 hours. The terminal halflife was calculated to be 1.8 hours, similar to the 1.82 hours reported literature for met-G-CSF (**3**). By contrast, the 20-kDa PEG-L3C protein showed a significantly longer circulating halflife that was biphasic: From 1.5 to 48 hours, the halflife was 13.7 hours. From 48 to 96 hours, when blood neutrophil levels were elevated (see below), the 20-kDa PEG-L3C protein cleared more rapidly, with a halflife of 4.7 hours.

Both wild-type G-CSF and 20-kDa PEG-L3C stimulated time-dependent increases in peripheral white blood cell and neutrophil counts over baseline values (Figure 6 B and C). Counts for the test groups receiving wild-type G-CSF peaked 10–24 hours postinjection and returned to baseline values by 48 hours, similar to results reported for met-G-CSF (**3, 13**). By contrast, those cell counts for the rats receiving PEG-L3C did not peak until 48–72 hours postinjection and

**Figure 4:** Nonreducing SDS-PAGE analysis of column pools for purified 5-kDa and 20-kDa PEGylated G-CSF (C17S) cysteine analogs. MW indicates molecular weight markers. G-CSF (C17S) is shown for comparison.

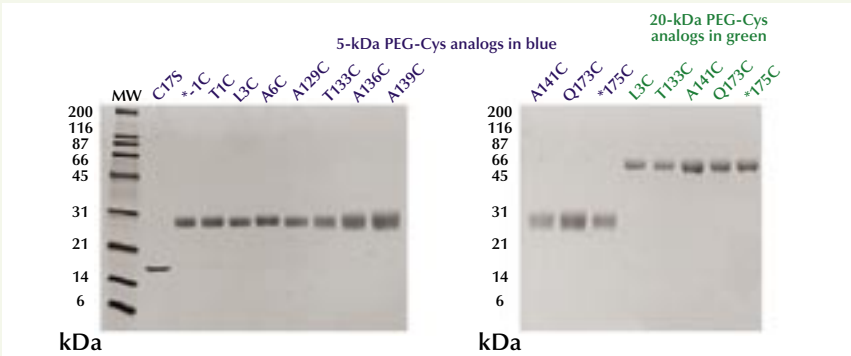
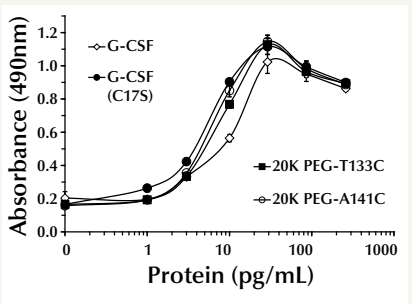



**Figure 5:** Dose-response curves for representative PEGylated G-CSF (C17S) cysteine analogs, G-CSF (C17S), and wild type G-CSF for stimulating proliferation of NFS-60 cells. Data are means ± standard deviations of triplicate wells from representative experiments.




returned to baseline values about 96–120 hours postinjection. White blood cell and neutrophil levels were significantly higher in the rats receiving 20-kDa PEG-L3C than in those receiving wild-type G-CSF at 48 and 72 hours postinjection ($p < 0.05$).

We noted no differences in red blood cell or platelet counts between the groups (data not shown). Red blood cell counts in both groups decreased about 15% during the first 12 hours postinjection, possibly due to repeated blood sampling (data not shown).

## Discussion

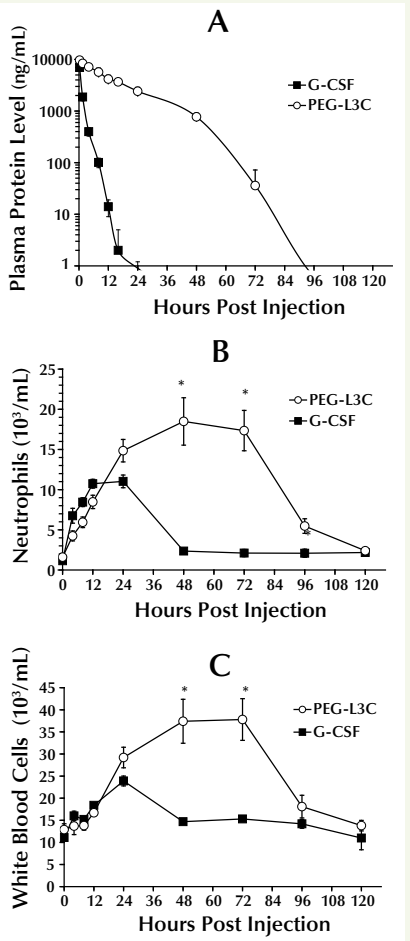
Modification of proteins with amine-reactive PEGs has proven to be a useful technology for prolonging circulating halflives and improving in vivo efficacy of protein therapeutics (**6, 21, 22**). However, amine-PEGylated proteins often occur in heterogeneous product mixtures that have significantly reduced (often by 10–100 fold) specific activities (**4, 12, 22**) due to reaction of the PEG reagent with multiple lysine residues in protein molecules. Some of those residues can lie at or near the active or receptor binding site of a given protein. Loss of bioactivity increases the amount (and cost) of drug required by patients and thus may fall short of maximizing its potential therapeutic benefit.

G-CSF is an example of a protein with in vitro bioactivity significantly reduced (up to 50-fold) by modification with conventional amine-reactive PEGs (**4**). G-CSF contains three lysines in helix A and a fourth at the N-terminal end of its A–B loop (**19, 20**). Those locations, along with helix D, have been identified as critical receptor binding sites (**7–10, 23**). So it is likely that modification of one or more lysine residues in any of the critical regions leads to poor in vitro bioactivity of amine-PEGylated G-CSF.

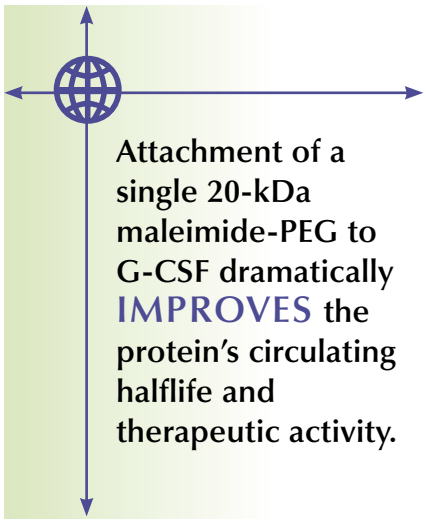
Our data indicate that it is possible to create PEGylated G-CSF (C17S) cysteine analogs with minimal, if any, loss of in vitro biological activity. This is accomplished by targeting attachment of the PEG molecule to specific predetermined sites on the molecule using site-specific PEGylation technology. We identified three regions — the N-terminal region preceding helix A, the C–D loop, and the C-terminal region following helix D — where cysteine residues can be introduced without appreciable effects on in vitro bioactivity. All our purified C17S cysteine analogs reacted readily with maleimide PEGs and yielded mono-PEGylated proteins, whereas wild-type G-CSF and C17S did not react with the PEG reagents under identical conditions. The latter result strongly suggests that the PEG molecules attach to our added cysteine residues. Failure of C17 in wild type G-CSF to react with the PEG reagent under the conditions we used is probably due to the fact that C17 is partially buried (**19, 24**). A previous report indicated that it is unreactive to alkylating agents (**24**).

All the PEGylated G-CSF (C17S)

**Figure 6:** Changes in protein plasma levels (Panel A), blood neutrophil counts (Panel B), and white blood cell counts (Panel C) following a single intravenous injection in rats of 100 µg protein/kg 20-kDa-PEG-L3C and wild type G-CSF. Data are means ± standard deviations for three rats per group. Asterisks (*) denote cell numbers significantly different ($p < 0.05$) between groups.



cysteine analogs we studied retained complete or near-complete in vitro biological activity, even when modified with large 20-kDa PEGs. The large number of fully active cysteine analogs (even PEGylated molecules) identified probably comes from our focusing mutagenesis efforts on the nonhelical regions of the protein, which lie outside its major receptor binding sites. Our finding that large polymers can be attached to many amino acids in those regions without appreciable affects on in vitro bioactivity provides further evidence that the regions examined (the one preceding helix A, the C–D loop, and the region following helix D) are largely nonessential for in vitro biological activity of G-CSF. Our data regarding the region preceding helix A, in particular, coincide with studies showing that PEG-G-CSF proteins showing only moderate loss of in vitro bioactivity (two- to threefold) can be prepared by preferential attachment of PEG to the N-terminus of G-CSF or met-G-CSF (**5, 11**).




Attachment of a single 20-kDa maleimide-PEG to G-CSF (C17S/L3C) dramatically improves the protein's circulating halflife and therapeutic efficacy over those of wild-type G-CSF in rats — similar to what has been reported for met-G-CSF modified with one or multiple amine-reactive PEGs (**3, 4, 6**). The 20-kDa PEG-L3C stimulated increases in blood neutrophils and white blood cells that were significantly greater and longer lasting than those seen with the wild type molecule.

Kidney filtration and receptor-mediated endocytosis have been implicated as dual mechanisms regulating G-CSF metabolism and neutrophil homeostasis in vivo (**25**). G-CSF receptors are abundant on neutrophils and bone marrow cells. The more rapid clearance of PEG-L3C observed beginning 48 hours postinjection correlates with peak neutrophil levels in the rats. Metabolism of PEG-L3C by bone marrow cells and/or neutrophils is probably responsible for the more rapid clearance of PEG-L3C observed at that point. In studies to be presented elsewhere (manuscript in preparation), we find that PEG-L3C is also significantly more effective than met-G-CSF at accelerating recovery from chemotherapy-induced neutropenia following single or multiple subcutaneous injections in rats.

We have shown that mono-PEGylated G-CSF (C17S) proteins with preserved in vitro biological activity and improved in vivo therapeutic efficacy can be created by targeting attachment of the PEG molecule to nonessential sites on the protein using site-specific PEGylation. Targeted attachment leads to homogeneously modified proteins with significantly greater in vitro biological activities than PEG-G-CSF proteins prepared using nonspecific amine-PEGylation technologies. G-CSF has a four-helix bundle structure shared by more than 20 different cytokines and growth factors, collectively referred to as the growth hormone supergene family (**19, 26, 27**). Information gained from our studies with G-CSF should prove useful for creating long-acting, site-specific PEGylated forms of other members of that important family.

## Acknowledgments

We are grateful to Dr. James Ihle of St. Jude Children's Research Hospital (Memphis, TN) for kindly providing the NFS-60 cell line and to Sharon Carlson of Bolder BioTechnology, Inc. for preparing the figures included herein. Our work was supported by grants 1R43-CA78094 and 2R44-CA78094 to George Cox from the National Cancer Institute. This publication's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

## References

**1** Anderlini P, et al. Biologic and Clinical Effects of Granulocte Colony-Stimulating Factor in Normal Individuals. *Blood* 88, 1996: 2819–2825.

**2** Welte K, et al. Filgrastim (r-metHuG-CSF): The First 10 Years. *Blood* 88, 1996: 1907–1929.

**3** Tanaka H, et al. Pharmacokinetics of Recombinant Human Granulocyte Colony-Stimulating Factor Conjugated to Polyethylene Glycol in Rats. *Cancer Res.* 51, 1991: 3710–3714.

**4** Bowen S, et al. Relationship Between Molecular Mass and Duration of Activity of Polyethylene Glycol Conjugated Granulocyte Colony-Stimulating Factor Mutein. *Exp. Hematol.* 27, 1999: 425–432.

**5** Kinstler OB, et al. Characterization and Stability of N-Terminally PEGylated rhG-CSF. *Pharm. Res.* 13, 1996: 996–1002.

**6** Molineux G, et al. A New Form of Filgrastim with Sustained Duration In Vivo and Enhanced Ability to Mobilize PBPC in Both Mice and Humans. *Exp. Hematol.* 27, 1999: 1724–1734.

**7** Nagata S, et al. The Chromosomal Gene Structure and Two mRNAs for Human Granulocyte Colony-Stimulating Factor. *EMBO J.* 5, 1986: 575–581.

**8** Kuga T, et al. Mutagenesis of Human Granulocyte Colony Stimulating Factor. *Bioch. Biophys. Res. Comm.* 159, 1989: 103–111.

**9** Kubota N, et al. Structural Characterization of Natural and Recombinant Human Granulocyte Colony-Stimulating Factors. *J. Biochem.* 107, 1990: 486–492.

**10** Layton JE, et al. Identification of a Functional Domain of Human Granulocyte Colony-Stimulating Factor Using Neutralizing Monoclonal Antibodies. *J. Biol. Chem.* 266, 1991: 23815–23823.

**11** Gaertner HF, Offord RE. Site-Specific Attachment of Functionalized Poly(ethylene Glycol) to the Amino Terminus of Proteins. *Bioconjug. Chem.* 7, 1996: 38–44.

**12** Goodson RJ, Katre NV. Site-Directed PEGylation of Recombinant Interleukin-2 at its Glycosylation Site. *Biotechnol.* 8, 1990: 343–346.

**13** Cox GN, et al. Enhanced Circulating Half-Life and Hematopoietic Properties of a Human Granulocyte Colony-Stimulating Factor (G-CSF)–Immunoglobulin Fusion Protein. *Exp. Hematol.* 32, 2004: 441–449.

**14** Picken RN, et al. Nucleotide Sequence of the Gene for Heat-Stable Enterotoxin II of *Escherichia coli*. *Infect. Immun.* 42, 1983: 269–275.

**15** Higuchi R. Chapter 22: Recombinant PCR. *PCR Protocols: A Guide to Methods and Applications.* Innis MA, et al., eds. Academic Press: San Diego, CA, 1990: 177–183.

**16** Horton RM. Chapter 25: In Vitro Recombination and Mutagnesis of DNA — SOIng Together Tailor-Made Genes. Methods in Molecular Biology, Vol. 15: *PCR Protocols: Current Methods and Applications.* White BA, ed. Humana Press: Totawa, NJ, 1993: 251–266.

**17** Koshland D, Botstein D. Secretion of Beta-Lactamase Requires the Carboxy End of the Protein. *Cell* 20, 1980: 749–760.

**18** Lu HS, et al. Disulfide and Secondary Structures of Recombinant Human Granulocyte Colony-Stimulating Factor. *Arch. Biochem. Biophys.* 268, 1989: 81–92.

**19** Hill CP, Osslund TD, Eisenberg D. The Structure of Granulocyte-Colony-Stimulating Factor and Its Relationship to Other Growth Factors. *Proc. Natl. Acad. Sci. USA* 90, 1993: 5167–5171.

**20** Souza LM, et al. Recombinant Human Granulocyte Colony-Stimulating Factor: Effects on Normal and Leukemic Myeloid Cells. *Science* 232, 1986: 61–65.

**21** Abuchowski A, et al. Cancer Therapy with Chemically Modified Enzymes, I: Antitumor Properties of Polyethylene Glycol–Asparaginase Conjugates. *Cancer Biochem. Biophys.* 7, 1984: 175–186.

**22** Bailon P, et al. Rational Design of a Potent, Long-Acting Form of Interferon: A 40-kDa Branched Polyethylene Glycol-Conjugated Interferon α2a for the Treatment of Hepatitis C. *Bioconjugate Chem.* 12, 2001: 195–202.

**23** Young DC, et al. Characterization of the Receptor Binding Determinants of Granulocyte Colony Stimulating Factor. *Prot. Sci.* 6, 1997: 1228–1236.

**24** Arakawa T, et al. Cysteine 17 of Recombinant Human Granulocyte Colony-Stimulating Factor Is Partially Solvent-Exposed. *J. Protein Chem.* 12, 1993: 525–531.

**25** Kuwabara T, et al. Nonlinear Pharmacokinetics of a Recombinant Human Granulocyte Colony-Stimulating Factor Derivative (Nartograstim): Species Differences Among Rats, Monkeys, and Humans. *JPET* 271, 1994: 1535–1543.

**26** Bazan F. Haemopoietic Receptors and Helical Cytokines. *Immunol. Today* 11, 1990: 350–354.

**27** Mott HR, Campbell ID. Four-Helix Bundle Growth Factors and Their Receptors: Protein–Protein Interactions. *Cur. Op. Struct. Biol.* 5, 1995: 114–121.

***Mary S. Rosendahl*** *is senior scientist,* ***Daniel H. Doherty*** *is director of research,* ***Darin J. Smith*** *is senior research associate, and corresponding author* ***George N. Cox*** *is chief scientific officer with Bolder BioTechnology, Inc., 4056 Youngfield Street, Wheat Ridge, CO 80033; 1-303-420-4420, fax 1-303-420-4426; jcox@bolderbio.com.* ***Alison M. Bendele*** *is president and chief executive officer of Bolder BioPATH, Inc., Campus Box 345, University of Colorado, Boulder, CO 80309.*