

Synthesis of [U - ^{13}C , ^{15}N]-Cysteine Hydrochloride: An Important Tool for Heteronuclear, Multi-Dimensional NMR Studies of Proteins

Michael J. Panigot^a, Stephen W. Fesik^b, and Robert W. Curley, Jr.^{a*}

^aDivision of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, The Ohio State University, Columbus, Ohio 43210, and ^bPharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064

Summary

Protein structure determination by modern NMR techniques is greatly facilitated using ^{15}N - and ^{13}C -labeled proteins. Labeling of proteins that are overexpressed in mammalian cells is a difficult task that requires a growth medium consisting of algal hydrolysates supplemented with labeled amino acids. Although most of the amino acids can be obtained to prepare an isotopically labeled growth medium for mammalian cells, [U - ^{13}C , ^{15}N]-cysteine is not available, hampering the backbone and cysteine side-chain assignments and structure determination in the vicinity of the cysteine residues. A synthesis of D,L-[U - ^{13}C , ^{15}N]-cysteine hydrochloride in good overall yield is described which makes use of readily available ^{15}N - and ^{13}C -labeled starting materials and will facilitate heteronuclear multidimensional NMR studies of proteins that are overexpressed in mammalian cells.

Key Words: Synthesis, cysteine, carbon-13, nitrogen-15, NMR, protein

Introduction

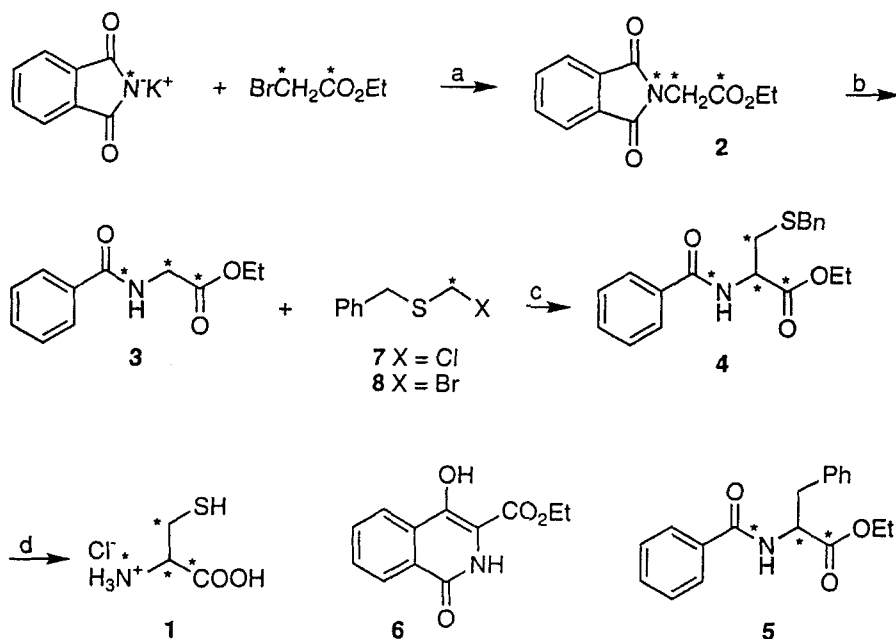
Three-dimensional structure determination of proteins by NMR is greatly aided by the use of ^{15}N -, ^{13}C -labeled proteins and recently developed heteronuclear multi-dimensional NMR techniques.¹ Although isotope labeling has been readily achieved for proteins that can be overexpressed in bacteria, labeling in mammalian cells has proved to be more difficult, requiring a growth medium supplemented with isotopically labeled amino acids.² Most of the isotopically labeled amino acids that are required for mammalian cell growth can be prepared from an acid hydrolysate of isotopically labeled bacterial or algal cell extracts. However, under the conditions for protein hydrolysis, cysteine, among other amino acids, is destroyed.³ Although ^{15}N -labeled cysteine can be commercially obtained, [U - ^{13}C , ^{15}N]cysteine, important for the production of uniformly ^{15}N -, ^{13}C -labeled proteins, is not readily available. The lack of ^{13}C -labeled cysteine makes it difficult to assign the protein backbone and cysteine sidechains and obtain an accurate structure in the vicinity of the cysteine residues. This is especially problematic for proteins obtained from mammalian cells that often contain multiple disulfide linkages.

Herein we describe the synthesis of [*U*-¹³C, ¹⁵N]-cysteine hydrochloride (**1**) which makes use of readily available labelled starting materials and can be modified to permit the selective labelling of any of the carbon atoms of the cysteine. Successful incorporation into mammalian proteins of interest will now allow the complete assignment of the NMR resonances and more accurate solution structures of proteins that can only be obtained from mammalian cells in their proper folded form.

Results and Discussion

The synthesis of isotopically labeled cysteine is shown in Scheme I. Gabriel synthesis⁴ of protected uniformly labeled glycine **2** proceeded in good yield (85%) utilizing commercially available [*1,2*-¹³C₂]-ethyl bromoacetate and [¹⁵N]-potassium phthalimide. We found, as have others,^{5,6} that base-catalyzed methylene alkylation of **2** was not an efficient means to introduce a protected thiomethyl group because of the rearrangement of **2** to the isoquinoline **6** under these conditions. Nonetheless, **2**

Scheme 1^a



^aReagents: (a) DMF, rt, 16h, 85%; (b) i, 6N HCl, reflux, 16h; ii, PhCOCl, NaOH, H₂O, 1h; iii, EtOH, cat. H₂SO₄, 16h, 52%; (c) 2 LDA, -78 °C to -42 °C, 2h, 77%; (d) i, 6N HCl, reflux, 16h; ii, 48% HBr, reflux, 96h, Dowex 50W / HCl, reversed-phase preparative HPLC, 48%

represents a desirable starting point for preparing labeled cysteine because of the good yield in which it can be prepared using the relatively inexpensive [^{15}N]-potassium phthalimide as the nitrogen source.

Acid-catalyzed hydrolysis of **2** to glycine followed by N-benzoylation and esterification of the unisolated hippuric acid produced labeled ethyl hippurate (**3**) in 52% overall yield. Earlier syntheses of cysteine have used base-catalyzed reactions between compounds like **3** and chloromethylsulfide **7** to prepare cysteine-like molecules.⁷ However, we found little reaction occurs between the dianion prepared from **3** and **7**, presumably due to the poor electrophilicity of **7**.⁸ Others have used the bromomethylsulfide **8** in situations where **7** has proven too unreactive^{9a}. Thus, we prepared [^{13}C]-**8** by treatment of benzylmercaptan with [^{13}C]-paraformaldehyde and HBr.^{9b} Careful distillation of the product mixture prepared **8** essentially free of by-product benzyl bromide. The presence of residual contaminating benzyl bromide was found to promote substantial formation of protected phenylalanine (**5**) in addition to the desired **4** upon treatment of the dianion of **3** with **8**. Alkylation of **3** with purified [^{13}C]-**8** produced labeled **4** in 77% yield free of detectable **5**. Attempts to effect a one-pot deprotection of **4** by prolonged heating in 48% HBr resulted in a complex, difficult to purify mixture which was toxic to the mammalian cells overexpressing the desired proteins. Therefore, a two-step deprotection was planned to remove the ester and amide protecting groups followed by more forcing conditions to hydrolyze the benzyl thioether. Overnight reflux of **4** in 6N HCl produced S-benzylcysteine hydrochloride which was used as obtained after isolation by $CHCl_3$ extraction of benzoic acid. Although published procedures¹⁰ are available for debenylation of S-benzylcysteine using CF_3SO_3H , in our hands these methods frequently led to cysteine degradation. Moreover, other methods for debenylation¹¹ can promote excessive dimerization of cysteine to cystine and/or proceed in poor yield. However, 96 hr reflux of uniformly labeled S-benzylcysteine hydrochloride in 48% HBr produced cysteine hydrobromide which was isolated by ion-exchange (Dowex 50W; HCl). Further purification by reversed-phase HPLC provided **1** in 48% yield. The facile oxidation of cysteine to its disulfide dimer cystine leads to the isolation of varying quantities (10-30%) of cystine using this procedure. However, if necessary, this mixture can be readily reduced solely to cysteine using a number of procedures such as Na/NH_3 or Sn/HCl .¹² Also, if necessary for a particular application, enzymatic resolution of racemic N-acetylcysteines can provide the natural L-enantiomer in good yield.¹³

In summary, we have developed a convenient synthesis of uniformly ^{15}N - and ^{13}C -labeled cysteine hydrochloride which will allow complete exploitation of the practical, cost-effective production of

uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled proteins expressed in mammalian cells by addition to labeled algal protein hydrolysates. The synthesis uses readily available labeled starting materials and at present costs ca. \$300 to produce the 80 mg of cysteine required to grow one liter of mammalian cells.³ By supplementing algal protein hydrolysates with uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled cysteine, all of the amino acid residues in protein from mammalian cells can be labeled, facilitating the backbone and cysteine side chain assignments and structure determination in the vicinity of the cysteine residues. In addition, the synthetic approach allows selectively labeled cysteines to be produced which may be useful for other applications (e.g. study of a peptide bound to its target site).

Experimental

General Procedures: All stable isotope-labeled compounds were from Cambridge Isotope Laboratories, Andover, MA. Other reagents were purchased from Aldrich Chemical Co. Dry THF was prepared by distillation from Na/benzophenone. Silica gel 60 (70-230 mesh) was from EM Reagents. Dowex 50W - X12 was obtained from BioRad. HPLC was conducted on a Beckman Instruments model 332 chromatograph with a model 164 UV detector using a 9.4 x 250 mm Zorbax ODS column (DuPont Instruments). ^1H and ^{13}C NMR spectra were collected at 250.13 and 62.9 MHz respectively on an IBM AC250 spectrometer with spectra recorded in CDCl_3 , referenced to residual CHCl_3 , and those measured in D_2O referenced to HOD and external dioxane/ C_6D_6 for ^1H and ^{13}C measurements respectively. FTIR spectra were measured using an Analect RFX40 instrument. Mass spectra were obtained in EI mode at The Ohio State University Campus Chemical Instrument Center on a Kratos MS-30 instrument.

[1,2- $^{13}\text{C}_2$, ^{15}N]-N-phthaloylglycine ethyl ester (2): To a solution of 2.0 g (11.83 mmol) of [1,2- $^{13}\text{C}_2$]-ethyl bromoacetate in 25 mL of DMF was added 2.4 g (12.9 mmol) of [^{15}N]-potassium phthalimide. The mixture was stirred at room temperature for 16 h, poured into 300 mL of ice water, and the product isolated by filtration and recrystallized from 95% EtOH to give 2.42 g (87%) of **2**: ^1H NMR (CDCl_3) δ 1.25 (t, 3H, $J = 7.15$ Hz), 4.19 (dq, 2H, $J = 7.15, 0.8$ Hz), 4.40 (dd, 2H, $J = 148, 0.8$ Hz), 7.7 - 7.9 (m, 4H); ^{13}C NMR (CDCl_3) δ 39.0 (dd, $J = 62, 12$ Hz), 167.2 (d, $J = 62$ Hz); IR (KBr) cm^{-1} , 3046, 2994, 2977, 2935, 1789, 1774, 1722, 1409, 1376, 1207, 1097, 1087, 944, 734, 710; HRMS calcd for $^{12}\text{C}_{10}$ $^{13}\text{C}_2$ $^{15}\text{NO}_4$ H_{11} 236.0725, found 236.0724.

[1,2- $^{13}\text{C}_2$, ^{15}N]-N-benzoylglycine ethyl ester (3): Compound **2** (3.57 g, 15.1 mmol) was refluxed in 100 mL of 6N HCl overnight. Phthalic acid was removed by filtration and the aqueous phase concentrated under reduced pressure. The crude glycine-HCl was dissolved in 100 mL of 2N NaOH, cooled to 0°C,

and 10 mL of benzoyl chloride was added. The mixture was stirred at 0°C for 30 min then stirred at room temperature for 1 h, poured into 100 mL of conc. HCl, the resulting solid filtered, and the filter cake washed with 200 mL of CHCl_3 . The remaining solid was dissolved in 150 mL of 95% EtOH containing 5 mL of conc. H_2SO_4 and refluxed overnight. The EtOH was evaporated and the residue partitioned between 100 mL each of ether and saturated aqueous NaHCO_3 . The organic layer was dried (MgSO_4), filtered, and the solvent removed under reduced pressure to give a product which can be crystallized from ether/petroleum ether. Yield 1.45 g (46%): ^1H NMR (CDCl_3) δ 1.29 (t, 3H, $J = 7.1$ Hz), 4.21 (dt, 2H, $J = 142, 5.6$ Hz), 4.22 (dd, 2H, $J = 7.1, 3.1$ Hz), 6.67 (brd, $J = 91$ Hz), 7.3-7.8 (m, 5H); ^{13}C NMR (CDCl_3) 41.9 (dd, $J = 61.2, 13.5$ Hz), 170.0 (d, $J = 61.2$ Hz); IR (KBr) cm^{-1} , 3330, 3068, 3056, 3031, 3002, 2975, 2964, 2931, 2883, 1714, 1641, 1602, 1579, 1517, 1488, 1463, 1448, 1398, 1369, 1348, 1313, 1303, 1297, 1255, 1176, 1168, 1153, 1054, 1016, 981, 694; HRMS calcd for $^{12}\text{C}_9$, $^{13}\text{C}_2$, H_{13} , $^{15}\text{NO}_3$ 210.0933, found 210.0931.

Synthesis of [^{13}C]-Benzyl bromomethylsulfide 8: This compound was prepared according to the method of Reich.^{10b} ^1H NMR (CDCl_3) δ 3.89 (d, 2H, $J = 6$ Hz), 4.43 (d, 2H, $J = 165$ Hz), 7.2-7.4 (m, 5H); ^{13}C NMR (CDCl_3) δ 36.54.

[1,2,3- ^{13}C , ^{15}N]-N-benzoyl-S-benzylcysteine ethyl ester (4): A solution of LDA (11.08 mmol, 2.2 eq) containing TMEDA (1.29 g, 2.2 eq) was prepared in 100 mL of dry THF and cooled to -78°C . A solution of 1.06 g (5.03 mmol) of **3** in 20 mL of THF was added over 15 min and the mixture was stirred at -78°C for 1 h then allowed to warm to -42°C . A solution of **8** (1.2 g, 5.54 mmol) in 25 mL of THF was added, the mixture stirred at -42°C for 2 h and quenched with 2 mL of saturated aqueous NH_4Cl . Solvent was evaporated and the residue dissolved in 100 mL of ether and washed with 2 x 100 mL H_2O and 2 x 100 mL 1N HCl, dried (MgSO_4), filtered, and solvent removed under reduced pressure. The residue was chromatographed on silical gel (1:2 EtOAc/hexane) to yield 1.15 g (66%) of **4**: ^1H NMR (CDCl_3) δ 1.27 (t, 3H, $J = 7.1$ Hz), 3.0 (dm, 2H, $J = 136$ Hz), 3.71 (d, 2H, $J = 4.1$ Hz), 4.20 (dq, 2H, $J = 7.1, 3.1$ Hz), 4.98 (dm, 1H, $J = 140$ Hz), 6.92 (dd, 1H, $J = 91, 7.1$ Hz), 7.2-7.8 (m, 10H); ^{13}C NMR (CDCl_3) δ 33.71 (d, 36 Hz), 52.2 (ddd, $J = 62, 36, 12.9$ Hz), 170.8 (d, $J = 62$ Hz); IR (KBr) cm^{-1} , 3340, 3068, 3027, 2983, 2933, 1708, 1687, 1635, 1600, 1579, 1509, 1484, 1305, 1292, 1261, 1224, 1187, 1031, 719, 696, 688; HRMS calcd for $^{12}\text{C}_{16}$, $^{13}\text{C}_3$, H_{21} , $^{15}\text{NO}_3\text{S}$ 347.1314, found 347.1313.

[1,2,3- ^{13}C , ^{15}N]-Cysteine hydrochloride (1): A suspension of **4** (0.772 g, 2.22 mmol) was refluxed in 50 mL of 6N HCl for 16 h, cooled, and washed with 3 x 50 mL of CHCl_3 . The aqueous layer was

concentrated under reduced pressure and the residue refluxed in 35 mL of 48% HBr for 96 h. The volume was reduced to 10 mL, diluted to 100 mL with H₂O, washed with CHCl₃ (2 x 25 mL), and concentrated to dryness. The residue was loaded onto 15 g of Dowex 50W-X12 (H⁺ form) and eluted with 1N HCl. The eluent was concentrated and further purified by reversed-phase HPLC (100% H₂O at 3 mL/min, UV detection at 215 nm). The HPLC fractions were concentrated and dissolved in 50 mL of H₂O, washed with 2 x 50 mL of CHCl₃, 1 x 50 mL of hexanes, and concentrated to give 192 mg (48%) of 1: ¹H NMR (D₂O) 2.94 (br d, 2H, J = 145 Hz), 3.73 (br dd, 1H, J = 145, 5.8 Hz); ¹³C NMR (D₂O) 26.4 (d, J = 35 Hz), 57.1 (br dd, J = 60, 35 Hz), 172.4 (d, J = 60 Hz); HRMS calcd for ¹³C₃H₇¹⁵NO₂S 125.0269, found 125.0266.

References

1. Clore G.M. and Gronenborn A.M. - *Science* **252**:1390 (1991).
2. (a) Thomas, J.N. - In *Large-scale Mammalian Cell Culture Technology*, Lubiniecki A., Ed., Marcel Dekker, New York, 1990, p.93. (b) Eagle H. - *Science* **122**:501 (1955).
3. Hansen A.P., Petros A.M., Mazar A.P., Pederson T.M., Rueter A. and Fesik S.W. - *Biochemistry* **31**:12713 (1992).
4. Gibson M.S. and Bradshaw R.W. - *Angew. Chem. Int. Ed. Engl.* **7**:919 (1968).
5. Gabriel S. and Colman J. - *Chem. Ber.* **33**:980 (1900).
6. Caswell L.R. and Campbell R.D. - *J. Org. Chem.* **26**:4175 (1961).
7. Wood J.L. and duVigneaud V. - *J. Biol. Chem.* **131**:267 (1939).
8. Krapcho A.P. and Dundulis E.A. - *Tetrahedron Lett.* 2205 (1976).
9. (a) Evans D.A., Mathre D.J. and Scott W.L. - *J. Org. Chem.* **50**:1830 (1985). (b) Reich H.J., Jasperse C.P. and Renga J.M. - *J. Org. Chem.* **51**:2981 (1986).
10. Yajima H., Fujii N., Ogawa H. and Kawatani H. - *J. Chem. Soc. Chem. Commun.* 107 (1974).
11. (a) Corrie J.E.T., Hlubucek J.R. and Lowe G. - *J. Chem. Soc. Perkin Trans. I* 1421 (1977). (b) Patterson W.I. and duVigneaud V. - *J. Biol. Chem.* **111**:393 (1935). (c) Hoffmann K., Bridgwater A. and Axelrod A.E. - *J. Am. Chem. Soc.* **71**:1253 (1949).
12. (a) Wood J.L. and du Vigneaud V. - *J. Biol. Chem.* **130**:109 (1939). (b) Clarke H.T. and Inouye J.M. - *J. Biol. Chem.* **94**:541 (1931).
13. Fischman A.J., Live D.H., Wyssbrod H.R., Agosta W.C. and Cowburn, D. - *J. Am. Chem. Soc.* **102**:2533 (1980).