Synthesis of Biarsenical Compound AsCy3 for HIV Imaging

Dr. Haitao Wu

Imaging Probe Development Center, National Heart, Lung, and Blood Institute, National Institutes of Health, 9800 Medical Center Drive, Rockville, MD 20850, ipdc@nhlbi.nih.gov

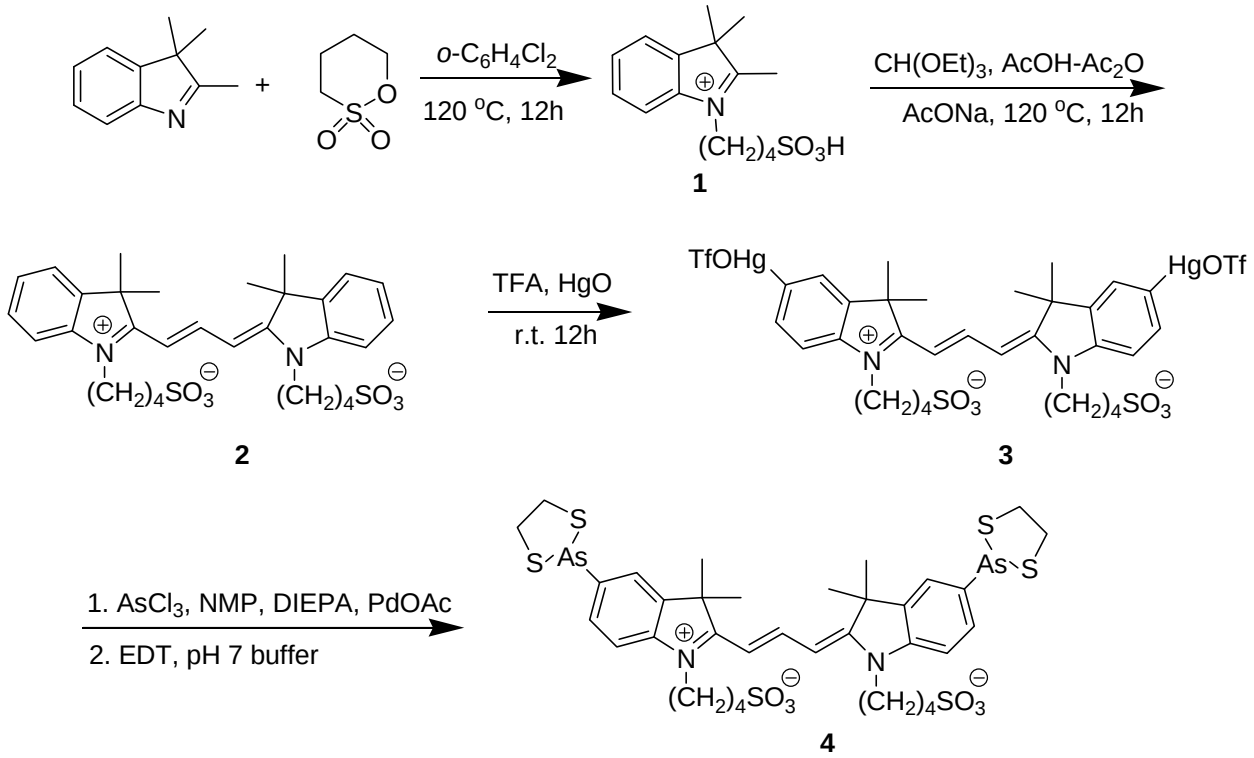
Background

HIV is a retrovirus belonging to the viral family *Retroviridae*.[1] Expression of a single protein, termed Gag, is sufficient for efficient assembly of retrovirus-like particles in mammalian cells.[2, 3] Imaging of HIV-1 Gag has been carried out mostly with fluorescent proteins (FPs).[4, 5] However, there are several drawbacks for the imaging of retrovirus assembly. The relatively large size of FPs can alter the properties of retroviral proteins and FPs take a long time to mature ($t_{1/2}$ > 1hr), a time frame which is long relative to HIV assembly and budding. Finally, it is difficult to engineer infectious particles containing FPs and to test the effects of the FP tag on the functional properties of the protein. In contrast, the tetracysteine tag/biarsenical dye labeling approach presents several advantages. Tetracysteine tags are as small as, minimally, 6 amino acids that are required to for protein insertion. They do not disrupt protein-protein interactions nor perturb the correct trafficking of tagged proteins[6, 7] and biarsenical dye binds to the tag very rapidly making this system very attractive for imaging HIV-1 proteins in live cells. This study was designed to use a biarsenical compound[8, 9] to study folding, maturation, and trafficking of the HIV-1 Gag protein by exploring the molecular mechanisms in the assembly process, and in the trafficking of Gag within virus-producing cells.

Chemistry

The target compound was obtained according to literature[8] methods and the synthesis is outlined in Scheme 1.

Scheme 1. Synthesis of AsCy3.



Experimental

General. Reagents were obtained from Sigma-Aldrich, Inc. unless otherwise noted, and used without further purification. All synthetic manipulations were performed under a dry nitrogen atmosphere using standard techniques. ^{1}H and ^{13}C NMR spectra were acquired using a Varian spectrometer operating at 400MHz and 100MHz respectively. Chemical shifts are reported in parts per million (δ) and referenced to tetramethylsilane (TMS). ESI mass spectrometry (ESI MS) was performed on LC/MSD TrapXCl Agilent Technologies instrument. A Beckman Coulter System Gold HPLC was used for analyses and product purification.

1-δ-Sulfobutyl-2,3,3-trimethyl-indolenine (1). A mixture of 2,3,3-trimethyl-indole (0.80g, 5.0mmol) and 1,4-butane sulfone (2.25g, 16mmol) in 1,2-dichlorobenzene (7mL) was refluxed at 120°C for 12 hours under a nitrogen atmosphere. After cooling, the reaction mixture was added into cold acetone (100mL). The precipitate was filtered, washed with acetone (3x10mL) and dried *in vacuo* to afford the title compound as a pink powder (1.1g, 55% yield). ^{1}H NMR (DMSO-d_6): δ 1.52 (s, 6H), 1.72 (m, 2H), 1.95 (m,2H), 2.49 (m, 2H), 2.84 (s, 3H), 4.48 (t, J =

7.8Hz, 2H), 7.61 (m, 2H), 7.82 (m, 1H), 8.03 (m, 1H). ^{13}C NMR (DMSO-d_6): δ 13.8, 22.0, 22.1, 26.0, 47.1, 50.1, 54.1, 115.6, 123.4, 128.9, 129.3, 141.2, 141.8, 196.5. ESI MS: m/z calcd. $[MH]^+$ 296.1, found 296.4.

3-(3, 3-dimethyl-1-(4-sulfobutyl)indolin-2-ylidene)prop-1-enyl)-3, 3-dimethyl-1-(4-sulfobutyl) -3*H*-indolium (Cy3) (2). Compound **1** (0.5g, 1.7mmol) and triethyl orthoformate (0.5g, 3.4mmol) were heated in a buffered acetic anhydride/ acetic acid mixture (7mL acetic acid, 5 mL acetic anhydride and 0.32g sodium acetate) at 120°C for 12 hours. After the reaction, the solvent was removed *in vacuo* and the residue was dissolved in a minimum amount of dimethylformamide and purified by HPLC monitored at 540 nm on an Agilent Zorbax C18 (21.5x250 mm, 5µm) column run at a 12 mL/min flow rate. The solvent system consisted of solvent A (0.05% trifluoroacetic acid in water) and B (0.1% trifluoroacetic acid in acetonitrile) with gradient of the solvent B from 40 to 90% over 15 minutes. The product eluted at 14.8 minutes. The product fraction was concentrated and then lyophilized to afford the title compound as a pink powder (0.11g, 22 % yield). ^{1}H NMR (DMSO-d_6): δ 1.69 (s, 12H), 1.84-1.73 (m, 8H), 2.56 (m, 4H), 4.14 (t, J = 7.0 Hz, 4H), 6.57 (d, J = 13.6 Hz, 2H), 7.29 (t, J = 7.6 Hz, 2H), 7.43 (t, J = 7.6 Hz, 2H), 7.45 (d, J = 8.0 Hz, 2H), 7.62 (d, J = 6.8 Hz, 2H), 8.35 (t, J = 13.6 Hz, 1H). ^{13}C NMR (DMSO- d_6): δ 22.3, 26.1, 27.4, 44.7, 48.8, 50.6, 50.7, 78.6, 78.9, 79.2, 102.7, 111.6, 122.3, 125.1, 128.6, 140.5, 141.8, 149.9, 173.7. ESI MS: m/z calcd. $[MH]^+$ 600.3, found 600.8.

3-(5-(1, 3, 2-dithiarsolan-2-yl)-3, 3-dimethyl-1-(4-sulfobutyl)indolin-2-ylidene)prop-1-enyl)-5-(1, 3, 2-dithiarsolan-2-yl)-3, 3-dimethyl-1-(4-sulfobutyl)-3*H*-indolium (AsCy3) (3 and 4). The cyanine dye compound (**2**) (95mg, 0.15mmol) and mercuric oxide (108mg, 0.30mmol) were stirred in trifluoroacetic acid (2.0mL) at room temperature overnight. The solvent was evaporated *in vacuo* and the residue was washed thoroughly with water to afford a pink solid, Cy3-bis-mercuric trifluoroacetate, (**3**) (180mg, 75% yield). After drying under high vacuum with phosphorous pentoxide (P_2O_5) overnight, Cy3-bis-mercuric trifluoroacetate (76mg, 0.06mmol) was suspended in anhydrous *N*-methyl pyrrolidinone (NMP, 0.6mL) containing arsenic trichloride (75µL, 1.5mmol, 26 equiv.), diisopropyl ethyl amine (DIPEA, 0.6mmol, 55µL, 10 equiv.) and a catalytic amount of palladium acetate (5mg). The reaction mixture was stirred at room temperature for 8 hours. Phosphate buffer (5mL 0.25M, pH 7) and 1, 2-ethanedithiol

(2.7mol, 226mL 44 equiv.) were then added. The solution was extracted with chloroform ($CHCl_3$) (3×10mL) and dried over sodium sulfate (Na_2SO_4). The solvent was evaporated under reduced pressure. The crude product residue was dissolved in dimethyl sulfoxide (1 mL) and purified by semi-preparative HPLC on an Agilent Eclipse XDB C18 (9.4 x 250 mm, 5 µm) column run at a 3mL/min flow rate. The solvent system consisted of solvent A (0.05% trifluoroacetic acid in water) and B (0.05% trifluoroacetic acid in methanol) with a gradient of solvent B from 70 to 95% over 12 minutes, then maintaining at 95% of B for 2 minutes and then to 70% of B over 1 minutes. The final product eluted at 13 min. The fraction containing the product was concentrated under reduced pressure at ambient temperature and then lyophilized to afford a purple solid (35 mg, 62% yield). Quality control by analytical HPLC was performed using an Agilent Eclipse XDB C18 (4.6x150 mm, 5 µm) column at a 0.5 mL/min flow rate. The solvent system consisted of solvent A (0.05% trifluoroacetic acid in water) and B (0.05% trifluoroacetic acid in acetonitrile) with a gradient of solvent B from 50 to 95% over 10 minutes. The product was eluted at 7.1 minutes. ^{1}H NMR (DMSO-d_6): δ 1.69-1.82 (m, 20H), 3.22 (m, 4H), 3.37(m, 4H), 4.14 (m, 4H), 6.60 (d, J = 13.2 Hz, 2H), 7.53 (d, J = 8.4 Hz, 2H), 7.71 (d, J = 8.4 Hz, 2H), 7.87 (s, 2H), 8.33 (t, J = 13.4 Hz,1H). ^{13}C NMR (DMSO-d_6): δ 22.5, 27.4, 28.5, 41.6, 48.8, 50.8, 103.4, 111.5, 124.4, 129.8, 131.3, 140.6, 140.7, 142.9, 173.7. ESI MS: m/z calcd. for $[MH]^+$ 933.0, found 933.2.

References

1. Varmus H. Retroviruses. Science. 1988 Jun 10;240(4858): 1427-35. PMID 3287617

2. Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO and Rein A. Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. Proc Natl Acad Sci U S A. 1998 Jun 23;95(13): 7299-304. PMID 9636143

3. Fuller SD, Wilk T, Gowen BE, Krausslich HG and Vogt VM. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr Biol. 1997 Oct 1;7(10): 729-38. PMID 9368755

4. Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998 67: 509-44. PMID 9759496

5. Chalfie M, Tu Y, Euskirchen G, Ward WW and Prasher DC. Green fluorescent protein as a marker for gene expression. Science. 1994 Feb 11;263(5148): 802-5. PMID 8303295

6. Giepmans BN, Adams SR, Ellisman MH and Tsien RY. The fluorescent toolbox for assessing protein location and function. Science. 2006 Apr 14;312(5771): 217-24. PMID 16614209

7. Andresen M, Schmitz-Salue R and Jakobs S. Short tetracysteine tags to beta-tubulin demonstrate the significance of small labels for live cell imaging. Mol Biol Cell. 2004 Dec;15(12): 5616-22. PMID 15469986

8. Cao H, Xiong Y, Wang T, Chen B, Squier TC and Mayer MU. A red cy3-based biarsenical fluorescent probe targeted to a complementary binding peptide. J Am Chem Soc. 2007 Jul 18;129(28): 8672-3. PMID 17585763

9. Cao H, Chen B, Squier TC and Mayer MU. CrAsH: a biarsenical multi-use affinity probe with low non-specific fluorescence. Chem Commun (Camb). 2006 Jun 28;(24): 2601-3. PMID 16779491