J|A|C|S
ARTICLES
Published on Web 02/04/2004

Functional Characterization and NMR Spectroscopy on Full-Length Vpu from HIV-1 Prepared by Total Chemical Synthesis

Gerd G. Kochendoerfer,*,† David H. Jones,‡ Sangwon Lee,‡ Myrta Oblatt-Montal,§ Stanley J. Opella,‡ and Mauricio Montal§

Gryphon Therapeutics, 600 Gateway Blvd., South San Francisco, California 94080, Section of Neurobiology, Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0366, and Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093

Received October 10, 2003; E-mail: gerd@gryphonRX.com

Abstract: Vpu is an 81-residue integral membrane protein encoded in the HIV-1 genome that is of considerable interest because it plays important roles in the release of virus particles from infected cells and in the degradation of the cellular receptor. We report here the total chemical synthesis of full-length Vpu(1−81) as well as a site-specifically ^{15}N-labeled analogue, Vpu(2−81), using native chemical ligation methodologies and also report a structural and functional comparison of these constructs with recombinant protein obtained via bacterial expression. The structures of the synthetic and expressed polypeptides were similar in lipid micelles using solution NMR spectroscopy. Solid-state NMR spectra of the polypeptides in aligned hydrated lipid bilayers indicated that their overall topologies were also very comparable. Further, the channel activity of the synthetic protein was found to be analogous to that previously characterized for the recombinant protein. We have thus demonstrated that using solid phase peptide synthesis and chemical ligation it is feasible to obtain large quantities of a purified and homogeneous membrane protein in a structurally and functionally relevant form for future structural and characterization studies.

Introduction

Vpu is an 81-residue membrane protein encoded in the HIV-1 genome.[1−3] It has been defined as an accessory protein because the virus can be cultured in vitro in its absence. However, there is a growing body of evidence indicating that Vpu function is vital for in vivo replication and contributes to the virulence of HIV-1 infections in humans.[4] Topologically, Vpu consists of a single transmembrane helix encompassing residues 8−25 and a larger cytoplasmic region with two amphipathic in-plane helices.[5,6] The recently determined atomic resolution structure of the trans-membrane domain shows the presence of a kink at residue 17.[5] Vpu has been shown to play two important roles in the lifecycle of HIV-1, and these activities are associated with the distinct structural domains of the protein that are outlined in Figure 1.[6−8]

One of the primary functions of Vpu is to accelerate the degradation of the CD4 receptor in the endoplasmic reticulum (ER) of infected cells.[9] The first step in the HIV-1 infectious process involves the virus binding to the CD4 receptor on the surface of the target cell which is mediated by gp120 on the surface of HIV. To facilitate efficient replication, retroviruses commonly remove their respective cellular receptors after infection. This not only prevents the cell from becoming infected by additional virus particles but also sequesters both viral and cellular proteins in the ER. In the case of HIV-1, the precursor form of gp120 (gp160) binds to CD4 and effectively traps both proteins in the ER, which prevents gp160 from being processed in post-ER compartments. Vpu plays an important role in removing CD4 from the cell by binding to the CD4 receptor in the ER via contacts between the cytoplasmic domains of both proteins.[10] Phosphorylation of two serine residues in the cytoplasmic domain of Vpu is essential for the subsequent CD4 degradation and liberates the trapped gp160, allowing it

† Gryphon Therapeutics.
§ Section of Neurobiology, Division of Biological Sciences, University of California.
‡ Department of Chemistry and Biochemistry, University of California.

(1) Strebel, K.; Klimkait, T.; Martin, M. A. *Science* **1988**, *241*, 1221−1223.
(2) Cohen, E.; Terwilliger, E. F.; Sordroski, J. G.; Haseltine, W. A. *Nature (London)* **1988**, *334*, 532−534.
(3) Matsuda, Z.; Chou, M. J.; Matsuda, M.; Huang, J. H.; Chen, Y. M.; Redfield; R. M.; K. E., M.; Lee, T. H. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6968−6972.
(4) Bour, S.; Strebel, K. *Microb. Infect.* **2003**, *5*, 1029−1039.
(5) Park, S.; Mrse, A.; Nevzorov, A.; Mesleh, A.; Oblatt-Montal, M.; Montal, M.; Opella, S. *J. Mol. Biol.* **2003**, *333*, 409−424.
(6) Marassi, F. M.; Ma, C.; Gratkowski, H.; Straus, S. K.; Strebel, K.; Oblatt-Montal, M.; Montal, M.; Opella, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14336−14341.
(7) Schubert, U.; Ferrer-Montiel, A. V.; Oblatt-Montal, M.; Henklein, P.; Strebel, K.; Montal, M. *FEBS Lett.* **1996**, *398*, 12−18.
(8) Schubert, U.; Bour, S.; Ferrer-Montiel, A. V.; Montal, M.; Maldarell, F.; Strebel, K. *J. Virol.* **1996**, *70*, 809−819.
(9) Willey, R. L.; Maldarelli, F.; Martin, M. A.; Strebel, K. *J. Virol.* **1992**, *66*, 7193−7200.
(10) Bour, S.; Schubert, U.; Strebel, K. *J. Virol.* **1995**, *69*, 1510−1520.

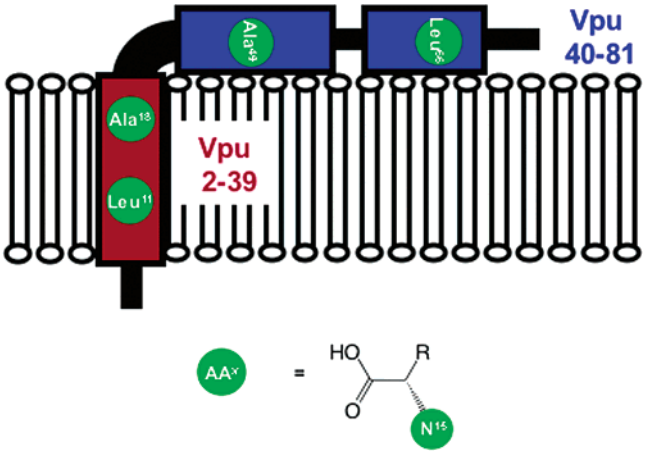

Figure 1. Schematic of the structure of HIV-Vpu indicating its topology relative to the membrane, domain structure, and the relative location of the isotopic labels employed in the NMR study. Vpu consists of an N-terminal, membrane spanning, helical domain and two C-terminal helical domains that lie parallel to the membrane plane.

to be processed to gp120 and subsequently migrate to the cell surface to become incorporated into new virus particles.[11−13] In addition to its effect on CD4, Vpu has a second major function. Vpu has been shown to facilitate the release of virus particles from infected cells, a process that is phosphorylation independent, occurs in post-ER compartments, and involves the transmembrane domain.[14] In the absence of Vpu, the viruses remain associated with the host cell and eventually become cytotoxic.[15] It has been suggested that this second function of Vpu is related to its ability to oligomerize and form ion-channels[7] similar to the influenza virus M2 ion-channel protein. Thus, Vpu not only increases the levels of envelope protein that reach the cell surface but also increases the secretion of virions that have incorporated this protein.

Understanding the different biological activities of Vpu requires detailed information about its structure and function. Unfortunately this is proving to be a particularly difficult task to perform on membrane proteins such as Vpu. The presence of lipids that surround membrane proteins in their native states often interferes with the preparation of samples for biophysical structure determination methods such as solution NMR spectroscopy or X-ray crystallography. By contrast, solid-state NMR of aligned samples is capable of providing atomic resolution structural constraints on membrane proteins in the presence of lipids and it is now possible to determine the three-dimensional structures of membrane proteins with atomic resolution using uniformly isotopically labeled samples by NMR spectroscopy.[5,16−18] However, many NMR and other spectroscopic studies [19] have been limited by the requirement for site-specific labeling. A number of recombinant-protein-expression based strategies have been devised to selectively label a protein (reviewed in ref 20), but they suffer from the limitation that solely a particular type of amino acid (e.g., alanine) or segment of the amino acid skeleton becomes labeled.[21,22] It would therefore be useful for the structure−function studies of membrane proteins to have a methodology that provides access to site-specifically labeled membrane proteins in sufficient quantities (milligrams) for NMR and other spectroscopic techniques, as well as crystallization trials.

The classic approach to the site-specific labeling of membrane proteins on the large scale needed for biophysical and spectroscopic studies has been to produce and label a selected peptide fragment derived from a region of interest in the target protein by SPPS (solid-phase peptide synthesis)-based methods. This approach has been applied to both the transmembrane domain[19,23] and the cytoplasmic domain[24−27] of Vpu to generate site-specifically labeled domain peptides for NMR, CD, and FTIR dichroism experiments. However, the chemical synthesis of full-length Vpu by SPPS has not been successful because this technique is typically limited to polypeptides between 25 and 50 amino acids in length due to the accumulation of side products after repeated amino acid couplings.[28]

To address this shortcoming of traditional peptide synthesis, we have recently introduced the total chemical synthesis of membrane proteins through chemical ligation methodologies. The basis of this approach is the covalent linkage of unprotected peptide segments using chemoselective ligation chemistries such as native chemical ligation.[29,30] Previously, the 97-residue membrane spanning M2 protein of *the influenza virus* was synthesized using native chemical ligation of peptide segments and assembled into native tetramers upon reconstitution into dodecylphosphocholine (DPC).[31] We have now applied modified ligation chemistries to synthesize full-length Vpu. We have demonstrated that the synthesized protein could be reconstituted into lipid bilayers in a physiologically relevant structure, since the structure and activity of the synthetic and expressed forms of the protein were found to be comparable. We have thus developed a methodology that allows for the preparation of full-length membrane proteins with site-specific labels in sufficient quantities for structural studies.

Experimental Section

Boc-protected amino acids were from Midwest Biotech (Fishers, IN). Trifluoroacetic acid (TFA) was obtained from Halocarbon (River Edge, NJ). *N*,*N*-Diisopropylethylamine was obtained from Applied Biosystems

(11) Schubert, U.; Strebel, K. *J. Virol.* **1994**, *68*, 2260−2271.
(12) Schubert, U.; Henklein, P.; Boldyreff, B.; Wingender, E.; Strebel, K.; Porstmann, T. *J. Mol. Biol.* **1994**, *236*, 16−25.
(13) Margottin, F.; Bour, S. P.; Durand, H.; Selig, L.; Benichou, S.; Richard, V.; Thomas, D.; Strebel, K.; Benarous, R. *Mol. Cell.* **1998**, *1*, 565−574.
(14) Strebel, K.; Klimkait, T.; Maldarelli, F.; Martin, M. A. *J. Virol.* **1989**, *63*, 3784−3791.
(15) Klimkait, T.; Strebel, K.; Hoggan, M. D.; Martin, M. A.; Orenstein, J. M. *J. Virol.* **1990**, *64*, 621−629.
(16) Opella, S. J.; Marassi, F. M.; Gesell, J. J.; Valente, A. P.; Kim, Y.; Oblatt-Montal, M.; Montal, M. *Nature Struct. Biol.* **1999**, *6*, 374−379.
(17) Marassi, F.; Opella, S. *Protein Sci.* **2003**, *12*, 403−411.
(18) Zeri, A.; Mesleh, M.; Nevzorov, A.; Opella, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6458−6463.
(19) Kukol, A.; Arkin, I. T. *Biophys. J.* **1999**, *77*, 1594−1601.
(20) Goto, N. K.; Kay, L. E. *Curr. Opin. Struct. Biol.* **2000**, *10*, 585−592.
(21) Hong, M. *J. Magn. Reson.* **1999**, *139*, 389−401.
(22) Hong, M.; Jakes, K. *J. Biomol. NMR* **1999**, *14*, 71−74.
(23) Wray, V.; Kinder, R.; Federau, T.; Henklein, P.; Bechinger, B.; Schubert, U. *Biochemistry* **1999**, *38*, 5272−5282.
(24) Wray, V.; Federau, T.; Henklein, P.; Klabunde, S.; Kunert, O.; Schomburg, D.; Schubert, U. *Int. J. Pept. Protein Res.* **1995**, *45*, 35−43.
(25) Henklein, P.; Schubert, U.; Kunert, O.; Klabunde, S.; Wray, V.; Kloppel, K. D.; Kiess, M.; Portsmann, T.; Schomburg, D. *Pept. Res.* **1993**, *6*, 79−87.
(26) Federau, T.; Schubert, U.; Flossdorf, J.; Henklein, P.; Schomburg, D.; Wray, V. *Int. J. Pept. Protein Res.* **1996**, *47*, 297−310.
(27) Henklein, P.; Kinder, R.; Schubert, U.; Bechinger, B. *FEBS Lett.* **2000**, *482*, 220−224.
(28) Henklein, P.; Sohr, R.; Brudel, M.; Pipkorn, R.; Schubert, U. Proceedings of the 23rd European Peptide Symposium, Braga, Portugal, Sept. 4−10, 1994; Maia, H. L. S., Ed.; ESCOM: Leiden, 1995.
(29) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779.
(30) Kochendoerfer, G. G. *Curr. Opin. Drug Discov. Devel.* **2001**, *4*, 205−214.
(31) Kochendoerfer, G. G.; Salom, D.; Lear, J. D.; Wilk-Orescan, R.; Kent, S. B. H.; DeGrado, W. F. *Biochemistry* **1999**, *38*, 11905−11913.

(Foster City, CA). HBTU (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) was obtained from Spectrum (Gardena, CA). Acetonitrile, dimethylformamide, and dichloromethane were obtained from Burdick & Jackson (Gardena, CA). ^{15}N-Labeled amino acids Boc-^{15}*N*-leucine-OH and Boc-^{15}*N*-alanine-OH were purchased from Cambridge Isotope Laboratories. Dihexanoylphosphatidylcholine D40 (DHPC) was obtained from Cambridge Isotope Laboratories (Andover, MA). Other chemicals were from Sigma (St. Louis, MO) or Aldrich (Milwaukee, WI) and were used as received. Peptide segments and ligation products were purified by preparative gradient RP-HPLC on a Rainin dual-pump high-pressure mixing system with 214 nm UV detection using a Vydac C-4 preparative column (10 μm particle size, 2.2 cm $\times$ 25 cm) and analytical RP-HPLC was performed on a Vydac C-18 analytical column (5 μm particle size, 0.46 cm $\times$ 15 cm), using a Hewlett-Packard Model 1100 quaternary pump high-pressure mixing system with 214 and 280 nm UV detection. Recombinant protein was purified on a Waters (Milford MA) 600 HPLC system using a Waters Deltapak C-4 preparative column (15 μm particle size, 0.78 $\times$ 30 cm).

Peptide Synthesis. The sequence of the ^{15}N-labeled N-terminal Vpu segment 1 (amino acid residues 2−39) is QPIQIAIVA LVVAIIIAIV VWSIVIIEYR KILRQRKID and the sequence of the ^{15}N-labeled C-terminal Vpu segment 2 (amino acid residues 40−81) is C LIDRLIERAE DSGNESEGEI SALVELGVEL GHHAPWDVDD L. This sequence corresponds to the sequence of Vpu from the HIV-1 isolate HTLVIII$_B$. ^{15}N-Labeled amino acids were incorporated into positions Leu^{11}, Ala^{18}, Ala^{49}, and Leu^{66}. The mutation R40C was introduced to allow for native chemical ligation at this site and was chosen since Vpu from a different strain (from the HIV-1 Eli isolate, Swiss Prot access code P05925) has a cysteine residue in the homologous position. All methionines were mutated to leucines to allow for cyanogen bromide cleavage during the recombinant production of the peptide (see below). This mutation has previously been shown to not alter Vpu function.[32] The sequence of synthetic, unlabeled Vpu(1−81) was identical to the sequence of Vpu (2−81) with the following changes: Methionine 1, which is deleted in Vpu (2−81), was reintroduced, amino acid residue 2 was a (wild-type) glutamate, and methionines at positions 66 and 70 were the wild-type residues.

Vpu segment 1 was synthesized at a 0.2 mmol scale on a thioester-generating resin,[33] using a custom-modified Applied Biosystems 433A peptide synthesizer by the in situ neutralization protocol for machine-assisted Boc (*tert*-butoxycarbonyl) chemistry.[34] The C-terminal aspartate residue of the α-thioester peptide was protected with an Asp(OFm) side-chain protecting group. Vpu segment 2 was synthesized analogously on a $-OCH_2-$Pam resin. Regular side-chain protecting groups were Arg(Tos), Asp(OChx), Asn(Xan), Cys(pMeBzl), Glu(OChx), His(Dnp), Lys-(2ClZ), Ser(Bzl), Trp(Hoc), and Tyr(2BrZ). No special coupling conditions were employed for ^{15}N-labeled amino acids. The peptides were deprotected and simultaneously cleaved from the resin support using HF/*p*-cresol according to standard Boc chemistry procedures.[34] Segment 1 was purified by preparative RP-HPLC with a linear gradient of 50 to 75% Buffer B (60% 2-propanol, 30% acetonitrile, 10% water containing 0.08% TFA) versus 0.1% aqueous TFA in 80 min. Segment 2 was purified with a linear gradient of 20−40% Buffer C (acetonitrile containing 0.1% TFA) versus 0.1% aqueous TFA in 80 min. Electrospray ionization mass spectra (ESI-MS) of the peptide products were obtained using a PE-Sciex API-1 quadrupole ion-spray mass spectrometer. Peptide masses were calculated from all the observed protonation states and peptide mass spectra were reconstructed using the MACSPEC software (PE-Sciex, Thornhill, ON, Canada). Theoretical masses were calculated using the MACPROMASS software (Terri Lee, City of Hope).

Chemical Protein Synthesis. A 50% molar excess of the purified, unprotected Vpu(40−81) peptide ("segment 2") was added to an aqueous solution of the purified α-thioester peptide Vpu (2−39) ("segment 1") (2 mM) containing 0.2 M sodium hydrogenphosphate, pH 7.5, 8 M urea, and 17 mg/mL of DPC. The ligation mixture was stirred for 5 h at room temperature and the reaction was monitored by analytical RP-HPLC and ESI-MS. After the reaction was complete, the reaction mixture was subsequently treated with an equal volume of a solution of acetonitrile/piperidine/β-mercaptoethanol (1/1/1 by volume) for 30 min to remove the Trp(Formyl) protecting group, the Asp(oFm) protecting group, and any residual His(Dnp) protecting groups. Reactants and products were separated by preparative RP-HPLC with a linear gradient of 55−75% Buffer B in water. Fractions containing full-length Vpu were identified by ESI-MS, pooled, and lyophilized. Unlabeled Vpu(1−81) was synthesized analogously.

Expression and Purification of Recombinant Vpu. Vpu(2−81) was expressed and purified as outlined previously.[6,32] Briefly, the protein was expressed as a Trp leader fusion protein in *Escherichia coli* to overcome the toxicity normally associated with expressing a membrane protein,[35] because the leader sequence targeted the fusion protein to inclusion bodies. After purification of the fusion protein using an N-terminal His tag, Vpu was cleaved from the leader sequence using cyanogen bromide. To facilitate this cleavage, all the methionine residues were mutated to leucine and the biological activity of the mutant was confirmed. Vpu was purified from the leader sequence using reverse phase chromatography with a linear gradient of 80% buffer I (90% water, 10% acetonitrile, 0.1% TFA) to 100% buffer II (10% water, 90% acetonitrile, 0.1% TFA).

Solution NMR Samples. The lyophilized protein samples were dissolved in 175 mM Dihexanoylphosphocholine (DHPC) at a concentration of 0.4 mM, pH 4.0. Fast HSQC spectra were acquired on a Bruker DMX 600 spectrometer at 50 °C using double INEPT transfer and State-TPPI for quadrature detection.[36]

Solid-State NMR spectra. All samples were prepared using the following protocol. The lyophilized protein (1 mg) was dissolved in 125 μL of trifluoroethanol (TFE). Then 50 mg of dioleoyl phosphatidylcholine (DOPC)/dioleoylphosphatidylglycerol (DOPG) at a mole ratio of 95:5 was dissolved in 125 μL of chloroform and added to the protein/TFE solution. The solution was subsequently spotted on 20 $\times$ 1 cm^2 glass cover slips and allowed to evaporate for 2 h. The cover slips were then placed under high vacuum for 24 h to remove any residual organic solvent. The plates were then stacked and incubated for 24 h at 42 °C in a desiccator with 98% RH that was achieved with a saturated solution of ammonium phosphate. After hydration, the sample was wrapped in Parafilm and sealed in a polyethylene bag. Spectra were acquired on a home-built NMR spectrometer with a ^{1}H resonance frequency of 700 MHz at 5 °C by spin-lock cross-polarization with 1 ms mix time as previously described.[6] ^{15}N chemical shifts were referenced to liquid ammonium at 0 ppm.

Single-Channel Recordings in Planar Lipid Bilayers. Lipid bilayers were assembled by apposition of two monolayers spread from a lipid solution in hexane as described.[6] The lipids were diphytanoylphosphatidylethanolamine and diphytanoylphosphatidylcholine (Avanti Polar Lipids, AL) at a 4:1 ratio in hexane (5 mg/mL). The aqueous subphase was composed of 0.5 M NaCl, and 5 mM Hepes (pH 7.4) or 0.5 M KCl, and 5 mM Hepes (pH 7.4), respectively. Purified polypeptides were dissolved in TFE at 0.01 mg/mL and added to the aqueous subphase after bilayer formation. Bilayer reconstitution experiments were performed at 24 °C.

Single-channel currents were recorded in the range of −200 to 200 mV. Acquisition and analysis of single-channel currents were performed

(32) Ma, C.; Marassi, F. M.; Jones, D. H.; Straus, S. K.; Bour, S.; Strebel, K.; Schubert, U.; Oblatt-Montal, M.; Montal, M.; Opella, S. J. *Protein Sci.* **2002**, *11*, 546−557.
(33) Camarero, J.; Muir, T. *Curr. Protocols Protein Sci.* **1999**, *18.4*, 1−21.
(34) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.
(35) Guinea, R.; Carrasco, L. *FEBS Lett.* **1994**, *343*, 242−246.
(36) Mori, S.; Abeygunawardana, C.; Johnson, M. O.; van Zijl, P. C. *J. Magn. Reson. B* **1995**, *108*, 94−98.

as described.[6] Recordings were filtered at 1 kHz and digitized at 0.1 ms per point by using an Axon TL-1 interface (Axon Instruments, Foster City, CA). Data processing was performed with pClamp 5.5 (Axon Instruments). The illustrated channel current recordings are representative of the most frequently observed conductances under the specified experimental conditions. Single-channel conductance was calculated from Gaussian fits to current histograms, and the channel open and closed lifetimes were calculated from exponential fits to probability density functions by using data from segments of continuous recordings lasting longer than 30 s and with > 300 events (mean ± SEM). Openings shorter than 0.3 ms were ignored. The data reported include statistical analysis of > 10 000 single channel openings.

Results

Total Chemical Synthesis. Full-length Vpu with the sequence of HIV-1 isolate $HTLVIII_B$ was synthesized from two peptide segments using the native chemical ligation strategy in combination with a lipid matrix-forming component. ^{15}N-Labeled amino acids were introduced at specific sites (Leu^{11}, Ala^{18}, Ala^{49}, and Leu^{66}) during the peptide segment assembly by coupling ^{15}N-labeled Boc-Leu-OH and Boc-Ala-OH, respectively. These positions were chosen on the basis of their locations in the membrane-spanning and cytoplasmic domains as determined from previous NMR studies (Figure 1). A cysteine residue needed for the native chemical ligation reaction was introduced to split the full-length sequence into two segments of comparable length. The particular mutation site Cys^{40} was also chosen because cysteine was found at this location in sequences from other HIV isolates and because this site also allowed for separation of the protein into the membrane-spanning domain and the cytoplasmic domain peptides.

In a pilot study, we produced 10 mg of unlabeled full-length polypeptide Vpu(1−81). Both peptide segments constituting the full-length Vpu peptide were readily assembled using optimized in situ neutralization Boc chemistry.[34] Pure peptides could be recovered at a 10−20% yield after cleavage and purification. A combination of denaturant (urea) and DPC was added to the ligation buffer to increase solubility of the segments during the ligation reaction. As a result, the native chemical ligation reaction of the peptides was extremely rapid and complete as demonstrated by the analytical data monitoring the progress of the reaction to form Vpu(1−81) that is presented in Figure 2. The top trace shows an RP-HPLC chromatogram at $t = 0$ and a reconstruct of the ESI-MS charge states of the two reactant segments, demonstrating high purity of the reactants. Only segment 1 is observed in the RP-HPLC trace, since Vpu segment 2 elutes in the void volume in the very hydrophobic HPLC gradients required to keep segment 1 in solution and is thus not detected in the chromatogram. After 3 h of reaction, the reactant peak at ∼18 min disappears and a new peak dominates at ∼14 min retention time. The earlier retention time is consistent with ligation of the more hydrophilic extramembraneous domain to the hydrophobic membrane-spanning domain. ESI-MS analysis (see inset) correlates this peak with the expected full-length reaction product (Vpu(1−81), experimental MW (molecular weight) 9136 ± 2 Da, theoretical MW 9136 Da; average isotope composition). Integration and comparison of the HPLC peaks of reactant and product after 5 h detected at 214 nm show that the reaction is >95% complete. After an additional protecting group removal step, reactants and products were separated by preparative reversed-phase HPLC. Fractions

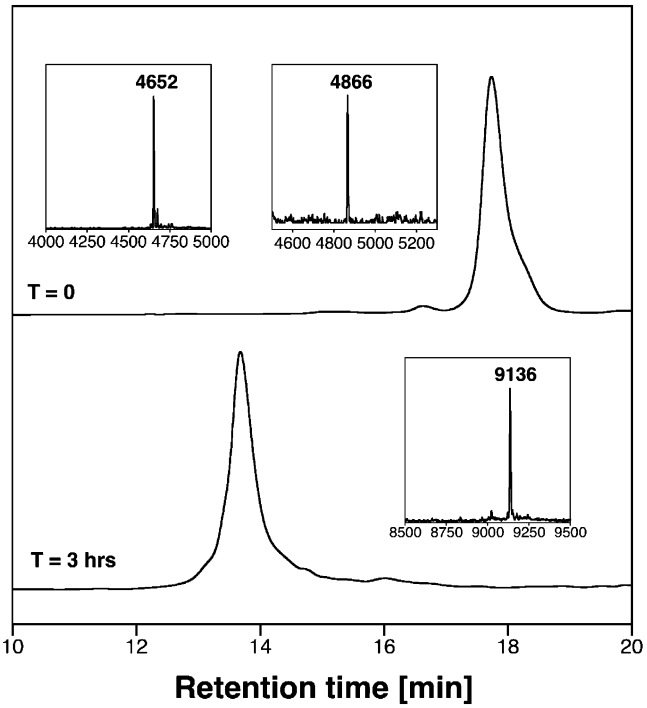

Figure 2. Progress of the ligation reaction: analytical RP HPLC traces (detected at 280 nm) monitoring the ligation reaction. Top panel: RP-HPLC chromatogram of the reaction mixture at $t = 0$. Only one major peak at ∼18 min retention time is observed in the reversed-phase HPLC. The mass of the eluting peptide is identified by ESI-MS analysis and correlated to the segment 1 thioester peptide Vpu(1−39). The more hydrophilic segment 2 is not retained at the high concentration of organic during loading and elutes in the void volume under these conditions. Inset: ESI-MS mass reconstruct spectra of the reactant peptides Vpu(1−39) (theoretical MW = 4865; average isotope composition) and Vpu(40−81) (theoretical MW = 4653; average isotope composition). Bottom panel: Equivalent RP-HPLC chromatogram after overnight reaction. The inset shows the ESI-MS spectrum and mass reconstruct of the purified ligation product Vpu(1−81) (theoretical MW = 9136 Da) (calculated MW = 9136; average isotope composition).

containing the expected ligation product were identified by electrospray mass spectrometry, pooled, and lyophilized.

We then proceeded to synthesize the target peptide Vpu(2−81) with ^{15}N-labels at positions Leu^{11}, Ala^{18}, Ala^{49}, and Leu^{66} using the same methods, as well as Vpu segment 2 (amino acid residues 40−81) with ^{15}N-labels at positions Ala^{49} and Leu^{66}. Analytical data demonstrating the identity and purity of full-length ^{15}N-labeled Vpu polypeptide (amino acid residues 2−81) obtained by the total chemical synthesis is shown in Figure 3. The reversed-phase HPLC chromatogram reveals a single symmetric peak for the desired product, indicating high product purity. The ESI-MS spectrum shows the desired mass including the isotopically enriched atoms (Vpu(2−81), experimental MW 8959 ± 2 Da, theoretical MW 8953 Da ($(N^{15})_4$ + average isotope composition). The additional peak in the RP-HPLC chromatogram at a later retention time is most likely due to a low level of pyroglutamate formation at the N-terminal glutamine site as suggested by a mass loss of −18 Da in the ESI-MS spectrum. After completion of these syntheses, the polypeptides were reconstituted into micelles and lipid bilayers for further studies.

Solution NMR Studies. The secondary and tertiary structure of Vpu is conveniently examined in phospholipid micelles of dihexanoylphosphocholine (DHPC) since the relatively rapid reorientation of such micelles enables the application of solution

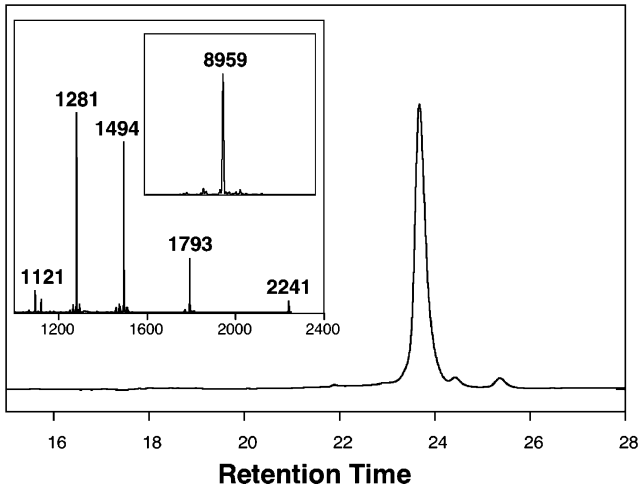

Figure 3. Analytical characterization of ^{15}N-labeled HIV-Vpu(2−81): RP-HPLC chromatogram of the purified ligation product Vpu(2−81). The inset shows the ESI-MS spectrum and the mass reconstruct from the charge states of the purified ligation product Vpu(2−81) (theoretical MW 8957 Da; average isotope composition including four ^{15}N-labeled alanine and leucine residues).

NMR methods. Figure 4 presents two-dimensional ^{1}H/^{15}N heteronuclear single quantum correlation (HSQC) spectra of recombinant Vpu(2−81), synthetic Vpu(2−81), and synthetic Vpu(40−81) in DHPC micelles. The resonances in the two-dimensional spectra are characterized by the ^{1}H and ^{15}N chemical shift frequencies associated with each ^{15}N-labeled site. These spectra are notable because the correlation peaks from the amide sites have very similar characteristics in all three samples. The spectra of synthetic Vpu(2−81) and Vpu(40−81) have only four and two resonances, respectively, which is consistent with the number of ^{15}N-labeled amino acids incorporated during the synthesis of the polypeptides. These resonances occur at frequencies very similar to those observed for the corresponding resonances in the spectra of recombinant Vpu(2−81) and are assigned accordingly. Since chemical shift frequencies are sensitive monitors of the local environment, these spectroscopic data suggest that all the proteins have very similar structures in DHPC micelles. The fact that the structure of the cytoplasmic domain is not significantly affected by the removal of the transmembrane domain is a function of the modular nature of membrane proteins and is consistent with previous observations on Vpu.[6]

Solid-State NMR Studies. Solution NMR relies on the rapid reorientation of the molecule to average the chemical shift to its isotropic value and the dipolar couplings to zero. Aligning a sample with respect to the magnetic field for solid-state NMR studies not only provides an alternative mechanism for obtaining high-resolution spectra but also enables the measurement of orientationally dependent chemical shift and heteronuclear dipolar coupling frequencies. The solid-state NMR of aligned samples[37] relies on the insertion of the protein in lipid bilayers and specifically aligning them in the magnetic field of the spectrometer. Under these conditions, the proteins are immobile on the NMR time scale. An ^{15}N chemical shift frequency near 200 ppm reflects an N−H bond orientation parallel to the magnetic field (transmembrane orientation) and a frequency near 70 ppm reflects an N−H bond oriented perpendicular to the field (orientation in the membrane plane). Figure 5 (left panel) shows a typical spectrum of recombinant Vpu(2−81) obtained in fluid bilayers of DOPC. Similar to previous spectra,[5,32] the Vpu resonances are segregated into two spectral regions: there is significant intensity in a peak around 200 ppm that arises from the N−H bonds in the transmembrane helix that is approximately perpendicular to the membrane surface and in a peak near 70 ppm that is consistent with N−H bonds that are in a helix lying parallel to the membrane surface. These results support a model that Vpu has a single transmembrane helix and two amphipathic helices that lie on the surface of the membrane.[6] The difference in intensities between these two regions of the spectra is the result of the presence of about 20 residues in the transmembrane domain, and about 50 in the cytoplasmic domain on the surface of the lipid bilayer.

The solid-state NMR spectrum of ^{15}N-labeled, synthetic Vpu(2−81) (Figure 5, center panel) is consistent with the results on uniformly labeled samples,[6] with one peak near 200 ppm and another near 70 ppm. For the synthetic protein, there are two labels in the transmembrane domain and two labels in the cytoplasmic region and the two peaks have comparable intensity. Thus, the overlapped resonances near 200 ppm are due to L11 and A18 and similarly the intensity around 70 ppm arises from Ala^{49} and Leu^{66}. These results indicate that the protein adopts a unique structure and orientation in the bilayer, since otherwise broad powder pattern line shapes would be observed. Since the peaks of the synthetic and recombinant samples occur in the same frequency region, the two proteins have similar structures and topologies with the helical transmembrane domain approximately perpendicular to the bilayer surface and the cytoplasmic region lying parallel to the surface of the membrane. The right panel of Figure 5 shows the spectrum obtained for synthetic Vpu(40−81) in fluid bilayers of DOPC. This spectrum consists of two peaks near 70 ppm. Since each peak arises from a single ^{15}N-labeled site, the difference in intensities probably reflects spectroscopic parameters, such as difference in cross-polarization efficiency during data acquisition. The resonance frequencies for these peaks are consistent with A49 and L66 in an α-helix on the surface of the membrane. In summary, the overlap of the ^{15}N resonances in the solid-state NMR spectra between recombinant and synthetic Vpu suggests that all possess a similar orientation and structure.

Functional Characterization. The ion-channel properties of several different recombinant constructs of Vpu have previously been characterized by the reconstitution of these molecules into planar lipid bilayers that were formed at the end of a patch pipet by the apposition of two monolayers.[5,6,32,38] This work not only confirmed that the transmembrane domain is responsible for mediating the ion flow across the lipid bilayer [7] but also indicated that the cytoplasmic region modulated the open and closed states of the channel.[38] Figure 6 shows representative channel current recordings of synthetic full-length Vpu in lipid bilayers and corresponding cumulative current histograms. Channel activity normally occurs in bursts with rapid transitions between the open and closed states. The most frequent channels of synthetic Vpu(2−81) displayed a conductance of 30 pS in both 0.5 M NaCl, 5 mM Hepes pH 7.4 (left column) and 0.5 M KCl, 5 mM Hepes pH 7.4 (center column) similar to that of recombinant Vpu under equivalent conditions.[6,38] In addition

(37) Opella, S.; Stewart, P.; Valentine, K. *Q. Rev. Biophys.* **1987**, *19*, 7−49.

(38) Montal, M. *FEBS Lett.* **2003**, *552*, 47−53.

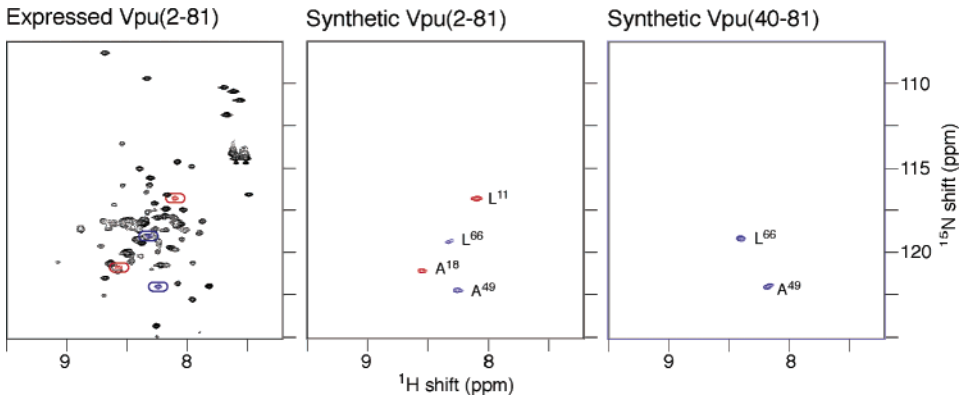

Figure 4. Two-dimensional heteronuclear single quantum correlation (HSQC) spectra of the three Vpu constructs in dihexanoyl phosphatidylcholine micelles. Peaks corresponding to the TM domain residues Leu[11] and Ala[18] are circled in red, and peaks corresponding to the cytoplasmic domain are circled in blue. Left panel: Expressed, uniformly ^{15}N-labeled recombinant Vpu(2−81). Center panel: Synthetic Vpu(2−81) with site-specific ^{15}N labels in positions Leu[11], Ala[18], Ala[49], and Leu[66]. Right panel: Synthetic Vpu(40−81) with site-specific ^{15}N labels in positions Ala[49] and Leu[66]. The correlation peaks found for the synthetic constructs match very closely the assigned correlation peaks observed in the homogeneously labeled samples.

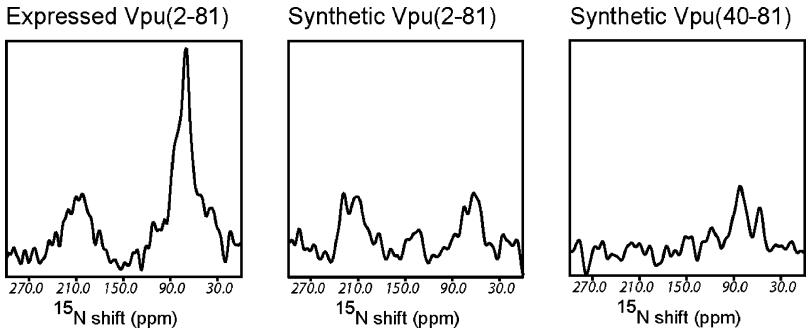

Figure 5. One-dimensional solid-state ^{15}N NMR spectra of the three Vpu constructs obtained at 4 °C in oriented lipid bilayers: Left panel: Expressed uniformly ^{15}N-labeled recombinant Vpu(2−81). Center panel: Synthetic Vpu(2−81) with site-specific ^{15}N labels in positions Leu[11], Ala[18], Ala[49], and Leu[66]. Right panel: Synthetic Vpu(40−81) with site-specific ^{15}N labels in positions Ala[49] and Leu[66]. The locations of the resonances for the site-specifically incorporated ^{15}N labels correspond to the expected shifts for residues in the transmembrane domain (Leu[11], Ala[18]) and cytoplasmic domain (Ala[49] and Leu[66]) (see text).

to the primary conductance, a second discrete channel with a conductance of 20 pS occurred with a high frequency (right column). Channels of 10 pS conductances were also detected, however with much lower occurrence. In ∼30% of experiments, no channel activity was detected. This pattern is typical of channel forming molecules that oligomerize and clearly demonstrate that synthetic full-length Vpu assembles in lipid bilayers into homooligomeric complexes and exhibits ion channel activity comparable to that of the recombinant protein.

Discussion

Membrane proteins are an abundant and important class of biomolecules, since virtually all cellular processes either occur at the surface of a membrane or have components that pass through a lipid bilayer. A significant impediment to biophysical and structural studies of membrane proteins remains in the requirement for large quantities of functional protein. Overexpression of integral membrane proteins in bacteria often leads to toxicity at high concentration and resultant lysis of the *E. coli* bacterial host,[35] whereas expression in mammalian cells typically results in low protein yields. While there are a number of methods to overexpress membrane proteins,[32,39,40] the majority of membrane protein structures have been determined from material from naturally abundant sources.[41]

Recent advances in the assembly of peptides on solid support[34] and in the applicability of chemoselective ligation chemistries (reviewed in ref 30) now make feasible the synthesis of membrane proteins by solely chemical means as an alternative to their recombinant production. As demonstrated previously [31] and in this report, chemical protein synthesis can be used to prepare milligram amounts of highly pure and homogeneous polypeptides that can be refolded into native membrane proteins. This methodology also facilitates the incorporation of noncoded amino acids that may introduce non-native function, incorporation of spectroscopic reporter groups such as isotope labels, spin labels, or fluorescent labels, and the site-specific introduction of posttranslational modifications such as glycosylation and phosphorylation under complete control. The latter posttranslational modification is of particular interest in the case of HIV-Vpu, since phosphorylation[42] of Ser[52] and Ser[56] plays an important role in its biological function.

In the context of NMR spectroscopy, chemical synthesis is highly complementary to traditional recombinant-based protein production. It has been applied successfully for the incorporation of ^{13}C-labels in the catalytically essential Asp[25] of HIV-1 protease [43] and of multiple labels into Pro[114] into RNase A[44]

(39) Grisshammer, R.; Tate, C. G. *Q. Rev. Biophys.* **1995**, *28*, 315−422.
(40) Jones, D. H.; Barber, K. R.; VanDerLoo, E. W.; Grant, C. W. *Biochemistry* **1998**, *37*, 16780−16787.
(41) Tate, C. G. *FEBS Lett.* **2001**, *504*, 94−98.
(42) Schubert, U.; Schneider, T.; Henklein, P.; Hoffmann, K.; Berthold, E.; Hauser, H.; Pauli, G.; Porstmann, T. *Eur. J. Biochem.* **1992**, *204*, 875−883.

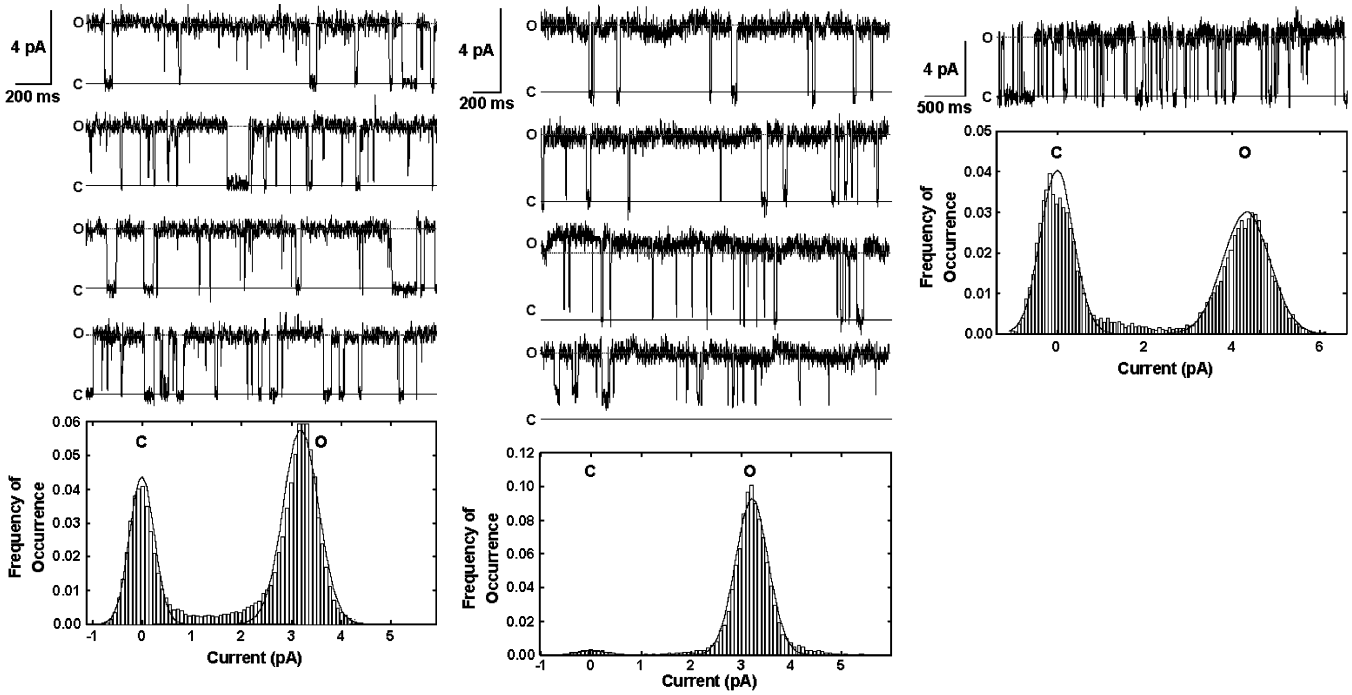

Figure 6. Single-channel recordings from synthetic full-length Vpu(1−81) reconstituted into lipid bilayers. Left column, top panel: Single-channel currents from synthetic Vpu(1−81) in 0.5M NaCl and 5 mM Hepes buffer (pH 7.0) recorded at 100 mV. Short segments of continuous recordings were selected to illustrate the occurrence of the dominant 30 pS channel. Upward deflections indicate channel opening; the solid lines labeled **c** indicate the current for the closed state and dotted lines labeled **o** mark the current of the open state. Bottom panel: Corresponding cumulative current histograms and Gaussian fits for the primary conductance generated from continuous segments of recordings lasting several minutes; **c** and **o** denote closed and open states. Center column, top panel: Single-channel recordings from synthetic full-length Vpu(1−81) in 0.5M KCl and 5 mM Hepes buffer (pH 7.0) at 100 mV. Short segments of continuous recordings were selected to illustrate the occurrence of the dominant 30 pS channel. Bottom panel: Corresponding cumulative current histograms and Gaussian fits for the primary conductances. Other conditions as for left column. Right column, top panel: Single channel currents recorded at 200 mV in 0.5M NaCl and 5 mM Hepes buffer (pH 7.0). In addition to the dominant 30 pS conductance shown on the left column, the right column displays openings with a conductance of 20 pS which are frequently observed. Bottom panel: Corresponding cumulative current histograms and Gaussian fits for the primary conductances. Other conditions as for left column.

for subsequent NMR spectroscopic studies. Total chemical synthesis has the clear advantage of allowing for site-specific labeling of proteins and may thus prove very useful for resonance assignment, as well as for structural fingerprinting of membrane proteins. For instance, using only few isotopic labels, one may test a hypothesis about the topology of a membrane protein by selectively labeling individual predicted domains and determining their orientation relative to the lipid bilayer by solid-state NMR spectroscopy. By contrast, it is significantly less expensive to uniformly ^{15}N, ^{13}C, and ^{2}H label proteins using traditional expression-based methods, mainly due to the prohibitive cost of isotope labeled amino acid starting materials employed in the synthesis. Bacterial expression thus represents a more size-independent and economical method to produce large quantities of material that is homogeneously labeled for NMR.

One of the main perceived problems in the chemical synthesis of membrane proteins is the need to fold the denatured polypeptide chain resulting from the synthesis into its native three-dimensional structure and proper oligomeric state. Whereas the folding of soluble proteins from their denatured state is reasonably well understood and has become routine, the mechanisms and driving forces that govern the folding of membrane proteins are far less understood. According to the two state model of membrane protein folding proposed by Engelman, Popot, and colleagues (reviewed in ref 45), helical membrane proteins may initially insert into lipid bilayers concomitant with secondary structure formation. This process is most likely driven by hydrophobic effects and hydrogen bond formation, even though the exact thermodynamics are still under investigation.[46] These helical regions subsequently may associate by forming specific contacts and ultimately generate the proper tertiary and quaternary structure.[45,47,48] This model would provide a possible mechanism for the folding of synthetic helical membrane proteins.

In this report we present both structural and functional evidence that full-length Vpu has indeed folded into the functional, correct three-dimensional structure. Comparison of the HSQC spectra of recombinant and synthetic Vpu and its C-terminal fragment demonstrate identical chemical shifts for the site-specific labels of the synthetic protein within the experimental error (based on a previous assignment of the peaks of the recombinant protein). Therefore, based on the HSQC experiment, micelle-refolded, full-length Vpu has a three-dimensional structure that is very similar to the recombinantly produced protein. The solid-state NMR data for the site-specifically labeled protein also indicates correct orientation of the synthetic protein when reconstituted into artificial DOPC/

(43) Smith, R.; Brereton, I.; Chai, R.; Kent, S. *Nature Struct. Biol.* **1996**, *3*, 946−950.
(44) Nilsson, B.; Hondal, R.; Soellner, M.; Raines, R. *J. Am. Chem. Soc.* **2003**, *125*, 5268−5269
(45) Popot, J. L.; Engelman, D. M. *Annu. Rev. Biochem.* **2000**, *69*, 881−922.
(46) White, S. H.; Wimley, W. C. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319−365.
(47) Dawson, J.; Weinger, J.; Engelman, D. *J. Mol. Biol.* **2002**, *316*, 799−805.
(48) DeGrado, W.; Gratkowski, H.; Lear, J. *Protein Sci.* **2003**, 647−665.

DOPG lipid bilayers. In summary, based on these structural studies, synthetic full-length Vpu is practically equivalent to its recombinantly produced homolog.

The observed ion channel activity of the synthetic protein is indistinguishable from the ion channel activity of recombinantly produced protein. Vpu protein is known to form an ion-channel and this characteristic provides a useful functional assay for the synthetic protein. Vpu was originally identified as a potential channel based on the sequence similarity of its transmembrane domain to the influenza ion-channel M2.[7,49] While it is clear that Vpu possesses channel activity, it is not known how this function is related to virus release.[7,49,50] Regardless of the mechanisms behind Vpu function, the channel activity found after reconstitution of both synthetic and recombinant Vpu was comparable, which demonstrates that the protein adopts a physiologically relevant conformation. Since the channel conductance is similar between both forms of Vpu studied, the transmembrane helices must be associating to form an oligomeric complex of similar pore size, which in turn indicates that the helices assemble into comparable oligomeric states. Our study therefore demonstrates for the first time wild-type like function for a synthetic [31] or semisynthetic [51] full-length ion channel protein.

In summary, our findings show that chemical protein synthesis can be employed to gain access to full-length membrane proteins in sufficient quantities for biophysical and functional studies and that these synthetic peptides can be folded into functional membrane proteins that have activities comparable to their native and recombinantly expressed counterparts.

Acknowledgment. This work was supported in part by grants GM62532 (GK), GM 56538 (SO and M.M.), GM066978 (SO), GM 49711 (M.M.) and the Biomedical Technology Resource for Solid-state NMR of proteins supported by P41EB002031 from the National Institutes of Health. D.J. was supported by postdoctoral fellowship 70561-31-RF from the amFAR Foundation. We thank Heather Porter and E. Neil Cagle (Gryphon Therapeutics) for expert assistance in peptide synthesis and purification and Christian Becker and Jose Santos for helpful discussions.

JA038985I

(49) Lamb, R. A.; Pinto, L. H. *Virology* **1997**, *229*, 1−11.
(50) Ewart, G. D.; Sutherland, T.; Gage, P. W.; Cox, G. B. *J. Virol.* **1996**, *70*, 7108−7115.
(51) Valiyaveetil, F. I.; MacKinnon, R.; Muir, T. W. *J. Am. Chem. Soc.* **2002**, *124*, 9113−9120.