A carboxy-terminal affinity tag for the purification and mass spectrometric characterization of integral membrane proteins

Julie P. Wong^{‡,§}, Emmanuelle Reboul[§], Robert S. Molday^{†,§,▲}, and Juergen Kast^{‡,||}

[‡]Biomedical Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

[†]Centre for Macular Research, University of British Columbia, Vancouver, British Columbia, Canada

[§]Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada

^{||}Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada

[▲]Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, British Columbia, Canada

Abstract

G-protein-coupled receptors (GPCRs) and other, structurally and functionally related membrane proteins represent particularly attractive targets for drug discovery. Integral membrane proteins are often difficult to purify from native contexts, and lack of sufficient quantities hampers subsequent structural and functional proteomic studies. We describe here the implementation of an optimized enrichment strategy involving a membrane protein-compatible 1D4 affinity tag that is derived from the carboxy-terminal nine amino residues of bovine rhodopsin, and its corresponding tag-specific, high-affinity monoclonal antibody. Expressing two GPCRs as well as two related ATP binding cassette (ABC) transporters in their functional forms in human cell lines, we have shown that a single detergent and wash condition can be employed for the purification of all said membrane proteins. Subsequent in-gel digestion with trypsin and mass spectrometric peptide analysis resulted in high sequence coverage for the ABC transporters ABCA1-1D4 and ABCA4-1D4. In contrast, digestion by various enzymatic combinations was necessary to obtain the best sequence coverage for affinity-enriched GPCRs CXCR4-1D4 and CCR5-1D4 in an annotated spectrum library, and to identify the N-glycosylation sites for CXCR4. Our results demonstrate that the 1D4-tag enrichment strategy is a versatile tool for the characterization of integral membrane proteins that can be employed for functional proteomic studies.

Introduction

G-protein-coupled receptors (GPCRs) represent the largest class of membrane proteins in the human genome. They are important in the control of fundamental processes such as

vision, olfaction, hormone signaling, stress responses as well as nervous system functions. Controlled, coordinated expression, localization and signaling are essential for normal physiology as malfunction leads to disease. Due to these critical implications, GPCRs are popular targets for drug discovery.

Despite extensive research in structural proteomics, only limited information is available on the three-dimensional structure of GPCRs, as these seven trans-membrane domain (7-TM) receptors are particularly difficult to recover in sufficient quantities from native contexts [1]. *E. Coli* and yeast overexpression systems are commonly employed to improve GPCR quantity, but generally do not produce proteins with accurate post-translational modifications or folding [2, 3]. Alternatively, enrichment approaches may be employed to address low expression levels, while maintaining physiologically relevant contexts as well as resolve experimental artifacts produced by heterologous expression systems. Affinity tagging is a general purification strategy for proteins that cannot be isolated conveniently via specific intrinsic properties. However, selecting an appropriate affinity purification system for a membrane protein of interest is non-trivial and may require extensive detergent screening and optimization [4].

The possibility of enriching for membrane proteins would be beneficial in functional proteomics, as it can be combined with sensitive detection techniques such as mass spectrometry to improve analysis in physiologically relevant cellular contexts [5]. In particular, targeted approaches such as Multiple Reaction Monitoring (MRM) and proteotypic peptide searches [6–9] that improve sensitivity through selective detection of pre-determined signature peptides and fragment ions are developing into key validation tools in proteomics. However, for targeted approaches to work at their full potential, experimental information is needed. Conversely, relying on predicted theoretical protein sequences is often experimentally irrelevant and inefficient. Populating data repositories with detailed and accurate peptide as well as associated fragmentation patterns can overcome this limitation, but these patterns are often underrepresented for GPCRs due to challenges associated with protein hydrophobicity, solubility and enzyme accessibility.

We describe here a versatile enrichment strategy involving a membrane protein compatible 1D4 affinity tag derived from bovine rhodopsin that shows high affinity to a tag-specific monoclonal antibody (MacKenzie et al. 1984; Hodges et al. (1988)). We demonstrate that the 1D4 enrichment strategy is applicable to GPCRs and related membrane proteins such as ATP binding cassette (ABC) transporters. We were able to achieve high sequence coverage from affinity-enriched CXCR4-1D4, CCR5-1D4, ABCA1-1D4 and ABCA4-1D4 that were expressed in physiologically relevant cell lines and functional contexts, digested by various enzymatic combinations, and characterized by mass spectrometry. From this analysis, we obtained the best sequence coverage for CXCR4 and CCR5 in an annotated spectrum library, and identified the N-glycosylation sites for CXCR4. Taken together, our results highlight the benefits of enrichment strategies by describing a membrane protein-specific affinity purification approach for functional proteomics studies.

Material and Methods

Construction of 1D4 tagged receptor constructs and site directed mutagenesis

Restriction endonuclease recognition sites and the C-terminal 1D4 tag (TETSQVAPA) were introduced into the primers used in the polymerase chain reaction (PCR) with wild type CXCR4 in pCDNA 3.1+ as the template. The PCR reactions were carried out for 35 cycles with 60 °C annealing temperature with *Pfu* polymerase (Fermentas, Glen Burnie, MD), forward and reverse primer sequences are as follows: *cggaattcgccgccatggaggggatcagtata* and *cgcctcgagttaggcaggcgccacttggtgtctctgtctggagtga* respectively. The 1D4 sequence is underlined. The PCR product was ligated into EcoRI and XhoI in the pCDNA 3.0+ multiple cloning site. The CXCR4 and CCR5 templates used were generous gifts from Dr. Bernhard Moser (Theodor-Kocher Institute; Bern, Switzerland) and the ABCA1-1D4 and ABCA4-1D4 constructs were generated as previously described [10].

N-linked glycosylation mutants were generated by site directed mutagenesis following the QuikChange kit (Stratagene; La Jolla, CA) with the CXCR4-1D4 pCDNA 3.0+ construct as the template. The following primers were used to mutate each of the asparagine residues to alanine: N11A (forward primer *agtatatacacttcagatgcctacaccgaggaatgggc*) and N176A (forward primer *cccgaattcatctttgccgcccgtcagtgaggcagatgac*). The sites of N A are underlined. The double mutant N11A/N176A was created by using the N11A construct as a template and mutating site N176 with the N176A primers with QuikChange kit.

Cell culture and transfection

Human 293T cells (ATCC, Manassas, VA) were cultured in Dulbecco's modified eagle media (DMEM) (Invitrogen, Carlsbad CA) and 10% fetal bovine serum (FBS). A confluent 20 mm dish of 293T cells was seeded 1:10 dilution 12 hours prior to calcium phosphate mediated transfection. 20 µg of CXCR4-1D4/pCDNA 3.0, CCR5-1D4/pCDNA 3.0, ABCA1-1D4/pCEP4 or ABCA4-1D4/pCEP4 were used to transfect one 20 mm dish of cells. 12 hours post-transfection, the cells were gently rinsed with phosphate buffered saline (PBS), media was replaced and cells were incubated at 37 °C in CO₂ overnight before harvest.

Cell lysis

Transiently transfected 293T cells in one 100 mm × 20 mm dish of transiently transfected 293T cells were solubilized in 1ml of 2% Triton X-100 in TBS for 1 hour at 4 °C with stirring and 1x complete protease inhibitors (Roche, Mannheim, Germany). Insoluble material and nuclei were pelleted at 200,000 × g TLA.3 Beckman rotor for 15 min. The resultant solubilized fraction were subjected to affinity purification.

Membrane fragment preparation

ABCA1-1D4 and ABCA4-1D4 transfected cells were resuspended and washed twice in a 10 mM Tris Buffer, pH 7.5, supplemented with 0.5 mM EDTA. Pellets were then resuspended in the same buffer with protease inhibitors, lysed for 1 h on ice vortexing occasionally, and homogenized with a glass homogenizer. This mixture was loaded on top of a gradient made with 5 % sucrose and 60 % sucrose in a 20 mM Tris buffer pH 7.4 supplemented in 150 mM

NaCl, 1 mM MgCl₂, 1 mM CaCl₂ and 0.1 mM EDTA, and centrifuged for 30 min at 24,000 rpm in a SW27.1 swinging bucket rotor. Membranes banded on top of the 60% sucrose fraction were harvested, washed one time, and centrifuged for 15 min at 40,000 rpm. Membranes were then stored in the same buffer, supplemented with protease inhibitors and 10 % glycerol, at -30 °C prior to solubilization. Membrane fragments prepared from ten plates, approximately 500 million transiently transfected cells, were solubilized in 1 ml of 2 % Triton X-100/TBS with protease inhibitors for 1 h with stirring at 4 °C. The insoluble material was removed by centrifugation at 100,000 × g for 10 min.

Immunoprecipitation and SDS-polyacrylamide gel electrophoresis

The purified rho 1D4 monoclonal antibody was purchased from UILO-UBC (<http://www.flintbox.com>). The cell lysate or solubilized sample was added to 50 µl of 1D4 monoclonal antibody (mAb) conjugated Sepharose prepared as previously described [11] with complete protease inhibitors (Roche, Mannheim, Germany), and rotated for 1 hour at 4 °C. Beads were washed ten times with 0.1 % Triton X-100/TBS. Tagged receptors were eluted twice with 100 µl 0.5 mg/ml 1D4 peptide in 0.1 % Triton X-100, and one final time with 1 % SDS/TBS at room temperature with shaking. SDS sample loading buffer (1 % SDS, 10 % glycerol, 0.1 M TrisHCl pH 6.8, bromophenol blue tracking dye) was added to samples and loaded on a 10 % SDS polyacrylamide gel with a 5 % stacking gel. Corresponding western blots were performed with these samples on polyvinylidene fluoride (PVDF) membranes (Millipore, Bedford, MA) for antibody staining and LI-COR Odyssey infrared imaging (Lincoln, NB).

Concanavalin A western blot probing

Lanes containing N11A CXCR4-1D4, N176A CXCR4-1D4, N11/176A CXCR4-1D4 and CXCR4-1D4 samples were transferred onto PVDF and analyzed by lectin binding with concanavalin A (ConA). The membrane was blocked at 37 °C in 5 % bovine serum albumin (BSA) in PBS with 0.01 % Tween-20 (PBST) and then bound with biotinyl succinylated ConA (Vector laboratories, Burlingame, CA) in blocking buffer. After 3 PBST washes, streptavidin 680 IR dye (LI-COR, Lincoln, NB) in 5 % BSA, 0.02 % SDS in TBS was bound and the blot was visualized on the LI-COR infrared imager.

In-gel digestion

Coomassie stained gel bands were washed three times 15 min each with 25 mM ammonium bicarbonate/50 % acetonitrile followed by reduction with 1.5 mg/ml dithiothreitol (DTT) at 50 °C for 30 min and alkylation with 10 mg/ml iodoacetamide (IAA) at room temperature for 20 min. Gel bands were rehydrated with 37.5 ng/ml sequencing grade trypsin and/or chymotrypsin (Roche, Mannheim, Germany) with 5–10 mM CaCl₂ in 50 mM ammonium bicarbonate (ABC) on ice for 30 min. Excess enzyme was removed and 15 µl of additional digestion buffer was added to prevent gel drying during digestion. The gel bands were incubated at 37 °C with shaking at 450 rpm for 18 h. Digests were extracted with 1:1 50 mM ammonium bicarbonate/acetonitrile and final extraction with 1:1 5 % formic acid/acetonitrile. Extracts were pooled and lyophilized to dryness and stored at -80°C until analyzed.

For double digests, the initial enzymes or chemicals employed were removed and the remaining gel band was dried with one volume of acetonitrile and speed vacuum centrifugation prior to rehydration with the second enzyme and subsequently extracted as described above. For cyanogen bromide and trypsin double digestion, the cyanogen bromide (CNBr) solution was prepared in a chemical fumehood by dissolving one crystal of CNBr (Aldrich, Milwaukee, Wisconsin) in 100 μ l of 50 % TFA. 20 μ l of CNBr solution was used to rehydrate the gel bands, which were kept in the dark overnight.

LC-MS/MS conditions

Proteolytically digested and/or CNBr-cleaved peptides were base- and then acid-extracted from the gel pieces, pooled and lyophilized. Samples were reconstituted in 6 μ l of 0.1 % formic acid (FA) and analyzed on a Qstar XL LC-MS/MS (Applied Biosystems, Foster City, CA). A PepMap C18, 3 μ m particle size and 100Å pore size column from LC Packings (Amsterdam, Netherlands) was used for peptide separation. Solvents A and B contained 5 % and 80 % acetonitrile, respectively, in water with 0.1 % formic acid. LC conditions started at 2 % solvent B with a gradient to 30 % B over 60 min, to 80 % B at 63 min, which was held for ten minutes before returning to 2 % B.

Sequence database search

The Mascot.dll script (version 1.6b21) in the Analyst QS 1.1 control software was used to create a peak list from the raw data. The MS/MS peaks recovered were all those with > 1 % base peak intensity. MS/MS spectra were grouped when precursors were within 0.1Da, and within 10 cycles. All extracted peaks were centroided at 50 % of peak height. The collected peak lists were then submitted to either Mascot or the X!Tandem algorithm for peptide sequence database search. Versions of the databases used for searching were the gpmDB, SwissProt 44.5, 45.0, 49.1, 49.3. Two databases were employed to confirm peptide/protein identifications. The search parameters entered were as follows: precursor ion mass tolerance 0.15 Da; fragment ion mass tolerance 0.5 Da; 3 missed cleavages; carbamidomethyl (C) fixed modification and oxidation (M) variable modification.

Confocal microscopy

Transfected monkey COS-7 cells were grown onto polylysine coated cover slips and then fixed with 4 % paraformaldehyde (Fisher; Fair Lawn, New Jersey) for 20 min at room temperature. Cells were blocked with normal goat serum and stained with 1:50 dilution of Rho1D4 antibody. The cells were labeled with anti-mouse Alexi 488 secondary antibody and subsequently counterstained for nuclei with with 4', 6-diamidino-2-phenylindole (DAPI).

Calcium Flux assays

2 million CXCR4 and CXCR4-1D4 transfected cells were harvested and loaded with Fluor-3AM fluorescent dye and kept in the dark at 37°C until ready to assay. 1 million cells were resuspended in 2.5 ml 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 140 mM NaCl, 10 mM glucose, 1.8 mM CaCl₂, 1 mM MgCl₂ and 3 mM KCl, pH 7.3. Cells were stirred in a Horiba Jobin FL3-22 TAU (Edison, New Jersey)

spectrofluorimeter at 506 nM excitation and 526 nM emission wavelengths. After approximately 20 s baseline stabilization, 100 nM SDF-1 was added to the cells and resultant fluorescence emission and calcium flux was recorded for 60 s time-based acquisition.

ABCA4 and ABCA4-1D4 reconstitution of proteins in liposomes

The lipids used for liposome formation (1:1 mixture of dioleoylphosphatidylethanolamine and brain polar lipids) were reconstituted in 25 mM NaHepes pH 7.5, 140 mM NaCl, 3 mM MgCl₂ and 10 % glycerol. The mixture was sonicated for 15 min at room temperature. The lipid stock (9 µl) was mixed with 15 % N-octylglucoside (6 µl) and purified ABCA4 or ABCA4-1D4 (100 µL) by slow addition. The resultant mixture was incubated 30 min on ice. Subsequently, detergent was removed from this mixture by passing through equilibrated Extracti-gel D resin (Pierce, Rockford, IL) and the flow through was collected. The reconstitution buffer consisted of 25 mM NaHepes pH 7.5, 140 mM NaCl, 10 % glycerol, 3 mM MgCl₂, 1 mM EDTA and 1 mM DTT.

ATPase assays

ATPase assays were performed in triplicate for each point determination using purified protein reconstituted into liposomes on the same day. The reaction was initiated with the addition of 1 µL of a mixture of 0.5 mM ATP, 5–10 µCi [α -³²P]-ATP in reconstitution buffer to the prepared liposomes and incubated at 37°C for 30 min. An aliquot of each reaction mixture was collected after 0, 5, 10, 20 and 30 min and spotted on polyethyleneimine cellulose (Aldrich, Milwaukee, WI). The hydrolyzed ATP was separated from the unhydrolyzed nucleotide by thin layer chromatography (TLC) in 0.5 M LiCl/1 M formic acid. The TLC plate was air dried and exposed to a phosphor screen for 3 hrs. The ABCA4 or ABCA4-1D4 content of the liposomes was assessed by Coomassie Blue gel and densitometry.

Results

C-terminally tagged receptors demonstrate accurate cellular localization

COS-7 cells were transiently transfected with CXCR4-1D4, CCR5-1D4, ABCA1-1D4 and ABCA4-1D4 constructs and labeled with the rho 1D4 monoclonal antibody to visualize receptor expression and localization by confocal microscopy. CXCR4-1D4, CCR5-1D4 and ABCA1-1D4 expression was observed along the plasma membrane (Figure 1A–C). ABCA4-1D4 was observed in vesicles in the cytosolic compartment of the transfected COS-7 cells (Figure 1D), mimicking the cellular localization of endogenous ABCA4 in the intradiscal compartments of outer segment disks in rod cells [12] [Illing et al. 1997]. Ectopically expressed bovine rhodopsin was used as a positive control for the 1D4 monoclonal antibody (Figure 1E).

The 1D4 tag does not alter receptor function

Murine B300-19 suspension pre-B cells, commonly used to study GPCR activity and signal transduction, were transfected with either CXCR4 or CXCR4-1D4 construct to determine the effect of the 1D4 tag on protein function. The cells were loaded with Fluo-3AM calcium

binding fluorescent dye indicator to monitor phospholipase gamma signal transduction and intracellular calcium mobilization. Receptor activation upon binding to stromal cell derived factor 1 (SDF-1) resulted in increased fluorescence signal, reflecting increased intracellular calcium levels and Fluoro-3AM binding. Both CXCR4-1D4 and CXCR4 showed similar fluorescence signal response times, with maximum fluorescence being achieved after 40 s following chemokine addition at 18 s (Figure 2A, 2B). Addition of SDF-1 to a final concentration of 100nM resulted in 450,000 cps and 350,000 cps change in fluorescence for CXCR4 and CXCR4-1D4 transfected cells, respectively. When the same volume of PBS was introduced at 18 s, baseline fluorescence was observed at 40 s and beyond (Figure 2C, 2D). The activity of ABCA4-1D4 was measured with ATP hydrolysis assays. The ATPase activity for ABCA4-1D4 followed the same increasing trend over 30 min as the positive control ABCA4. The activity after 30 min was 265 ± 3.65 nmol and 248 ± 47.35 nmol ATP/mg purified protein for ABCA4 and ABCA4-1D4, respectively (Figure 2E).

The 1D4 tag achieves high recovery for membrane proteins

Protein recovery from the 1D4 affinity purification was evaluated by performing densitometry on western blots of the purified receptors. To determine the peptide concentration necessary for maximum receptor recovery, the 1D4 peptide epitope was titrated for CXCR4-1D4 and ABCA1-1D4 elution from the immunoaffinity column, with the buffer and detergent concentrations kept constant. Optimum peptide concentration for receptor elution was 0.25 mg/ml for both receptors (Figure 3A, 3B). At concentrations beyond 0.25 mg/ml, the dose response curve started to plateau, while lower concentrations resulted in significantly reduced receptor recovery, making 0.25 mg/ml the preferred concentration, which is lower than what has been reported elsewhere [13]. Similarly, complete removal of Triton X-100 from the peptide elution also reduced receptor recovery, whereas as little as 0.05–0.1 % Triton X-100 detergent improved receptor yield (Figure 3C, 3D).

Applying the resultant standard 1D4 enrichment protocol to different membrane proteins resulted in recoveries of approximately 45–50 % of the receptor. To further increase yield, the remaining receptor was recovered from the immunoaffinity beads by a final non-specific elution with 1 % (v/v) SDS. From a single purification of 50 million cells transiently expressing CXCR4-1D4, approximately 500 ng of protein from peptide elution and another 500 ng from the final SDS elution was recovered, corresponding to approximately 90 % total protein recovery. On average, 1000 ± 20 ng and 300 ± 50 ng of CXCR4-1D4 and CCR5-1D4 was recovered from 1D4 purification of 100 million transiently transfected 293T cells, calculated from three independent experiments (data not shown).

The 1D4 tag enrichment is specific

The efficiency and specificity of this single-step enrichment system was determined by SDS-PAGE using equal protein amounts before and after purification. 1D4 tagged receptors were effectively separated from other cellular proteins, as shown in the reduction of bands in the eluate relative to the pre-column lane of the whole cell lysate (Figure 4). All of the receptors were enriched with minor aggregation and protein smearing on the Coomassie stained gel and Western blots. CXCR4-1D4 was observed as a broad band extending from ~40–50 kDa

and a lower molecular weight band at ~30 kDa. Two bands were visible for CCR5-1D4, three bands for ABCA1-1D4, and four bands for ABCA4-1D4. The identity of all of these bands was confirmed by *in situ* digestion and mass spectrometry as discussed below.

1D4 enrichment increases sequence coverage in mass spectrometry

To establish peptide maps of each receptor, CXCR4-1D4, CCR5-1D4, ABCA1-1D4 and ABCA4-1D4 were in-gel digested and subjected to mass spectrometric analysis after 1D4 enrichment, whereas due to its natural abundance, electrophoresis alone was adequate for separation of rhodopsin from rod outer segments of bovine retina. In each case, the gel bands excised for digestion and mass spectrometry are indicated by the arrows in Figure 4. Tryptic digestion of the ABC transporters identified 24 and six unique peptides for ABCA1 and ABCA4, respectively (Table 1). In contrast, tryptic digestion of CXCR4, CCR5 and rhodopsin only yielded one, two and three peptides, respectively.

To increase sequence coverage of the GPCRs, digestion with enzyme combinations was performed. The greatest number of unique peptides was obtained from a chymotrypsin/trypsin double digestion for CXCR4 and rhodopsin, which yielded four and six unique peptides, respectively. CNBr/trypsin digestion yielded six unique peptides for CCR5 (Table 1). For human CXCR4, these enzymatic combinations achieved the largest sequence coverage of 101 submitted entries in the Global Proteome Machine database (GPMDB) and a top score of -142, followed by -42 for the next entry (see Supplementary data II). No other entries were available for human CCR5 in this database, making this data set the first unique set of proteotypic peptides for this receptor (see Supplementary data II).

1D4 enrichment facilitates the detection of CXCR4 post-translational modifications

Computational predictions for N-linked glycosylation based on the presence of NX(S/T) consensus sequences revealed two putative glycosylation sites at positions 11 and 176 in human CXCR4. Subsequent comparison to the list of peptides obtained for this protein after 1D4 enrichment identified the presence of an unmodified asparagine at position 176 of the peptide AN₁₇₆VSEADDRY after a chymotrypsin and trypsin sequential double digestion (Table 1). In contrast, no peptide containing N11 could be detected by mass spectrometry. To confirm that failed detection of N11 glycosylation was due to the absence of the unmodified peptide, and that residue N11 was in fact glycosylated as indicated in previous reports [14, 15], the two predicted sites of post-translational modification were examined with site-directed mutagenesis as well as western blotting.

N11A, N176A and N11A/N176A CXCR4-1D4 mutant constructs were transiently transfected into human 293T cells and analyzed by Western blotting. When position 11 was mutated, the protein migrated at a lower molecular weight (Figure 5A lane 2). Mutation at position 176 did not change the migration relative to the tagged wild type receptor (Figure 5B, lanes 3 and 1). The double mutant N11A/N176A migrated at the same molecular weight as the N11A mutant (Figure 5A, lane 4). Duplicate lanes of the purified N A mutant samples run on the same polyacrylamide gel were transferred onto a PVDF membrane, which was probed with succinyl-biotinylated concanavalin A (ConA) to detect mannose-containing glycosylation sites on CXCR4 (Figure 5B). A sharp ConA positive band

migrating at approximately 40kDa was observed for CXCR4-1D4 and the N176A mutant, which corresponded to a minor, low intensity band for CXCR4-1D4 in the 1D4 probed blot (Figure 5B, lanes 1 and 3). However, this band was absent for the N11A and the N11A/N176A double mutants (Figure 5B, lanes 2 and 4).

Discussion

The 1D4 tag represents the C-terminal nine amino acids on the G-protein coupled receptor (GPCR) rhodopsin from *Bos taurus*. We have shown here that affinity enrichment using this tag is a viable option for studying other structurally and functionally similar GPCRs and other membrane proteins. Employing the 1D4 tag did not affect the localization and function of CXCR4, CCR5, ABCA1 and ABCA4. Each protein localized according to its natural context. CXCR4-1D4 stimulation with SDF-1 in B300-19 suspension pre-B cells induced calcium flux levels with response times and absolute change in fluorescence comparable to wild type CXCR4. The variation in fluorescence intensity between CXCR4 and CXCR4-1D4 upon SDF-1 addition is reflective of the differences in quantity of receptors expressed in the cell line. Similarly, the 1D4 tag did not affect the ATPase activity of a representative ABC transporter, ABCA4, as both receptor forms showed similar ATP hydrolysis rates.

Efficient removal of interfering proteins and high protein recovery is a key aspect of any enrichment strategy. While this has been achieved for some membrane proteins by individualizing the purification conditions, widely applicable universal protocols are clearly advantageous. The 1D4 enrichment procedure requires one detergent (Triton X-100) and wash condition for protein solubilization and purification of both groups of membrane proteins in a single step. The peptide competitive elution we have employed is specific, gentle and non-denaturing, minimizes receptor loss through protein precipitation and aggregation, which may occur with heating or lowering pH, and reduces non-specific protein and antibody leaching from the affinity column and non-specific background.

The resulting efficient enrichment of adequate quantities of membrane proteins has a direct impact on a number of functional proteomics applications. In proteomic studies involving complex protein mixtures, mass spectrometric analysis typically selects peptides with specific biophysical properties or high abundance for fragmentation [7, 19]. As less abundant proteins in the mixture are less likely to be sequenced, multi-pass membrane proteins, and GPCRs in particular, are often under-represented in such samples. Concentrating the membrane protein by enrichment can resolve some of these low expression issues. While *in silico* trypsin digestion of CXCR4 indicated that several peptides would have m/z ratios suitable for MS analysis, in-gel tryptic digestion of the enriched CXCR4 revealed only two peptides by tandem mass spectrometry, matching the limited entries in GPMDB (<http://gpmdb.thegpm.org>). Similarly, our entry for CCR5 was the only one available in GPMDB. This proved that the lack of proteotypic peptides after digestion is due to biophysical properties rather than protein abundance, as they were still not detected after enrichment. Accessibility of cleavage sites, low digestion rates, peptide size and/or hydrophobicity are possible reasons leading to this outcome.

To address limited sequence coverage upon trypsin digestion, 1D4 enrichment was used to comprehensively study enzyme combinations that increase the number of peptides sequenced by mass spectrometry for each of the GPCRs. This resulted in dramatically higher sequence coverage and generated the top entries for both GPCRs in GPMDB. As populating data repositories with accurate biochemical and mass spectral data on “proteotypic peptides” increases confidence in the correlation of peptide-protein pairs and facilitates the design of targeted detection methods such as MRM assays [18], other difficult to study multi-pass membrane proteins may also benefit from a similar strategy.

Combining increased sequence coverage with the ability to express functional mammalian receptors in a mammalian system enables the comprehensive analysis of post-translational modifications while resolving artifacts observed from heterologous expression. Of the two putative N-linked glycosylation sites on CXCR4, tandem mass spectrometry, site-directed mutagenesis and lectin binding revealed that N11 was the only one modified, whereas N176 was not glycosylated. As most GPCRs are predominantly glycosylated at the second extracellular loop (ECL) – approximately 66% of receptors with at least one glycosylation occur at ECL2, 14% ECL1 and 20% ECL3, and not on the N-terminus [16] – our findings are remarkable as they confirm previous reports of CXCR4’s unusual glycosylation pattern, which may help in discriminating these putative post-translational modification sites reported to be involved in HIV-1 infectivity and tropism [17].

Determining the peptides that are suitable for detection and digestion conditions required to recover these peptides could also improve GPCR biomarker identification. For example, human CXCR4 has been implicated in diseases such as AIDS and cancer metastasis. How the CXCR4 heterogeneity contributes to function or pathology is not well understood. Using CXCR4 as a biomarker for disease diagnosis or profiling will not be feasible unless the corresponding post-translational modifications on the receptor are characterized in the healthy and disease conditions. Our work will therefore be useful for accurate protein identification and relative quantification in such profiles.

In addition to opening new avenues in functional proteomics, the 1D4 enrichment may also aid structural proteomics studies by identifying constant, soluble regions on the receptor for high-resolution analysis. Its specificity and mild elution conditions could be highly beneficial to identify novel binding partners of membrane proteins, and facilitate ligand-receptor or small molecule-receptor fingerprinting or footprinting studies by chemical or photo-crosslinking. As demonstrated here, using the 1D4 tag for optimized universal enrichment may thus have important implications in various areas of proteomic research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by operating and infrastructure grants from the Natural Sciences and Engineering Research Council of Canada, the Michael Smith Foundation for Health Research, and a Canadian Institutes for Health Research grant (MT 5822). JPW was supported by a CIHR doctoral research fellowship. The authors would like to thank Dr. Ron Beavis for lending his expertise in this work and Shujun Lin for her technical assistance with

the mass spectrometry. We are also grateful to Laurie Molday and Celene Grayson for their advice and suggestions as well as Jason Rogalski and Dr. Geraldine Walsh for assisting with the reviewing of this manuscript. The authors would like to dedicate this manuscript in memory of Dr. Ian Clark-Lewis.

The abbreviations used are

ABC	ATP binding cassette
conA	concanavalin A
CNBr	cyanogen bromide
DAPI	4', 6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle media
ECL	enhanced chemiluminescence
DTT	dithiothreitol
GPCR	G protein-coupled receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
mAbs	monoclonal antibodies
MRM	multiple reaction monitoring
MS	mass spectrometry
PBS	phosphate buffered saline
PVDF	polyvinylidene fluoride
ROS	rod outer segments
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SDF-1	stromal cell derived factor-1
TBS	tris buffered saline

References

1. Luche S, Santoni V, RT. Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics*. 2003; 3:249–253. [PubMed: 12627377]
2. Logsdon CD. The Influence of the cellular context on receptor function: a necessary consideration for physiological interpretations of receptor expression studies. *Life Sciences*. 1999; 64(6/7):369–374. [PubMed: 10069498]
3. Molloy MP, et al. Extraction of *Escherichia coli* proteins with organic solvents prior to two-dimensional electrophoresis. *Electrophoresis*. 1999; 20:701–704. [PubMed: 10344235]
4. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiol Biotechnology*. 2003; 60:523–533.
5. Wu CC, Yates Jr III. The application of mass spectrometry to membrane proteomics. *Nature Biotechnology*. 2003; 21:262–267.

6. Mirzaei H, et al. Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. *Molecular and Cellular Proteomics*. 2007; 7(4):813–23. [PubMed: 18089551]
7. Mallick P, et al. Computational prediction of proteotypic peptides for quantitative proteomics. *Nature Biotechnology*. 2007; 25:125–31.
8. Kuster B, Schirle M, Mallick P, Aebersold R. Scoring proteome with proteotypic peptide probes. *Nature Reviews/Molecular Cell Biology*. 2005; 6:577583.
9. Craig R, Cortens JC, Beavis RC. The use of proteotypic peptide libraries for protein identification. *Rapid Communications in Mass Spectrometry*. 2005; 19:1844–1850. [PubMed: 15945033]
10. Ahn J, et al. Functional Interactions between the two halves of the photoreceptor-specific ATP binding cassette protein ABCR (ABCA4). *The Journal of Biological Chemistry*. 2003; 278:39600–39607. [PubMed: 12888572]
11. Weber BH, et al. Inactivation of the murine X-linked juvenile retinoschisis gene, *Rslh*, suggests a role of retinoschisin in retinal cell layer organization and synaptic structure. *Proc Natl Acad Sci USA*. 2002; 99:6222–6227. [PubMed: 11983912]
12. Ahn J, Molday RS. Purification and characterization of ABCR from bovine rod outer segments. *Methods in Enzymology*. 2000; 315:864–79. [PubMed: 10736746]
13. Takayama H, et al. High-level expression, single-step immunoaffinity purification and characterization of human tetraspanin membrane protein CD81. *PLoS ONE*. 2008; 3(6):e2314. [PubMed: 18523555]
14. Chabot DJ, et al. N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates. *J Virology*. 2000; 74(9):4404–4413. [PubMed: 10756055]
15. Wang J, et al. N-linked glycosylation in the CXCR4-N-terminus inhibits binding to HIV-1 envelope glycoproteins. *Virology*. 2004; 324(1):140–150. [PubMed: 15183061]
16. Lanctot PM, et al. Importance of N-glycosylation positioning for cell-surface expression, targeting, affinity and quality control of the human AT1 receptor. *Biochem J*. 2005; 390(1):367–376. [PubMed: 15869468]
17. Huskens D, et al. The role of N-glycosylation sites on the CXCR4 receptor for CXCL-12 binding and signaling and X4 HIV-1 viral infectivity. *Virology*. 2007; 363:280–287. [PubMed: 17331556]
18. Lange V, et al. Targeted Quantitative Analysis of *Streptococcus pyogenes* Virulence Factors by Multiple Reaction Monitoring. *Molecular Cellular Proteomics*. 2008; 7(8):1489–1500. [PubMed: 18408245]
19. Kuster B, et al. Scoring proteomes with proteotypic peptide probes. *Nature Reviews/Molecular Cell Biology*. 2005; 6:577–583. [PubMed: 15957003]
20. MacKenzie D, Arendt A, Hargrave P, McDowell J, Molday RS. Localization of the binding sites for carboxyl-terminal specific anti-rhodopsin monoclonal antibodies using synthetic peptides. *Biochemistry*. 1984; 23:6544–6549. [PubMed: 6529569]
21. Hodges RS, Heaton RJ, Parker JM, Molday LL, Molday RS. Antigen-Antibody Interactions. Synthetic Peptides Define Linear Antigenic Determinants Recognized by Monoclonal Antibodies Directed to the Cytoplasmic Carboxy Terminus of Rhodopsin. *J Biol Chem*. 1988; 263:11768–11775. [PubMed: 2457026]
22. Illing M, Molday LL, Molday RS. The Rim Protein of Retinal Rod Outer Segments is a Member of the ABC Transporter Superfamily. *J Biol Chem*. 1997; 272:10303–10310. [PubMed: 9092582]

Figure 1.

1D4 C-terminally tagged receptors express and demonstrate accurate localization. Confocal microscopy slices of COS-7 cells transiently transfected with A) CXCR4-1D4, B) CCR5-1D4, C) ABCA1-1D4, D) ABCA4-1D4 and E) rhodopsin. Cells were fixed with paraformaldehyde and stained with mouse 1D4 mAb, anti-mouse alexi 488 and DAPI nuclear staining (blue)

Figure 2. CXCR4-1D4 and ABCA4-1D4 are functional and induce calcium mobilization upon SDF-1 binding and ATP hydrolysis, respectively. Calcium mobilization and PLC γ activation was detected in A) CXCR4 and B) CXCR4-1D4 expressing B300-19 cells loaded with Fluo-3AM calcium indicator fluorescent dye. Corresponding controls with equivalent volumes of PBS alone were added to C) CXCR4 and D) CXCR4-1D4 cells as a negative control. E) Purified ABCA4 and ABCA4-1D4 were reconstituted in brain polar liposomes for ATPase activity assay. ATPase activity was monitored with the release of radiolabelled ADP from [α - 32 P]-ATP over the course of 30 min at 37°C. The average and standard deviation from 3 experiments are shown.

Figure 3. Determining the optimum 1D4 peptide and detergent concentration for 1D4-tagged receptor recovery. Integrated band intensities were determined from densitometry measurements of 1D4 probed western blots of purified A) CXCR4-1D4 and B) ABCA1-1D4 plotted against peptide concentration, and purified C) CXCR4-1D4 and D) ABCA1-1D4 plotted against detergent concentration in the 0.5mg/ml 1D4 peptide in TBS for competitive elution from 1D4 conjugated Sepharose.

Figure 4.

C-terminally tagged 1D4 sepharose purified receptors expressed in 293T cells. Immunoblots and coomassie stained gels show receptor enrichment post immunoaffinity capture and peptide specific elution and a final elution with 1 % SDS to recover remaining receptor. Ectopically expressed A) bovine rhodopsin was used as a control for the purification since rhodopsin was the natural antigen in which the antibody was raised against. Receptor recovery and expression varies amongst the different receptors B) CXCR4-1D4, C) CCR5-1D4, ABCA1-1D4 and ABCA4-1D4.

Figure 5. CXCR4 post-translational modifications, concanavalin A western blotting and site-directed mutagenesis. A) anti-1D4 and B) Concanavalin A western blot of N A CXCR4-1D4 mutants.

Table 1

Sequence coverage for membrane proteins employing different enzymatic combinations for *in-gel* digestion. Refer to Supplementary data I for respective links to database matches.

Receptor and accession #	Enzymatic combination	Unique peptides	score	Sequence coverage (%)	Peptides
CXCR4 P61073	Trypsin	2	191	7.4% (26/352)	134 ⁴ YLAIVHATNSQRPR ¹⁴⁸ 309 ⁷ TSQAQHALTSVSR ³²²
	Chymotrypsin/trypsin	4	216	16% (55/352)	134 ⁴ YLAIVHATNSQRPR ¹⁴⁸ 171 ¹ TIPDFIFANVSEADDR ¹⁸³ 271 ¹ QCCEFENTVHK ²⁸² 307 ⁷ FKTSAQHALTSVSR ³²²
CCR5 P51681	Trypsin	1	80	3% (12/352)	322 ² CCSIFQQEAPER ³³⁴
	CNBr/trypsin	6	522	21% (75/352)	49 ¹ LVILINCKR ⁶⁰ 126 ⁶ YLAIVHAFVFAK ¹³⁸ 210 ⁶ VICYSGILKTLR ²²³ 287 ⁷ THCCINPIYAFVGEK ³⁰³ 303 ⁷ FRNYLLVFFQK ³¹⁴ 322 ² CCSIFQQEAPER ³³⁴
Rhodopsin P02699	Trypsin	3	159	11.5% (40/348)	135 ⁵ YVVCCKPMSNFR ¹⁴⁷ 231 ¹ EAAQQQESATTQK ²⁴⁵ 296 ⁷ TSAVYNPVIYIMINK ³¹¹
	CNBr/trypsin	3	195	11% (38/348)	126 ⁶ SLVVLAIER ¹³⁵ 296 ⁷ TSAVYNPVIYIMINK ³¹¹ 322 ² NPLGDDDEASTTVSK ³³⁹
	Chymotrypsin/trypsin	6	151	21% (72/348)	21 ¹ SPFEAPQYY ³⁰ 146 ⁶ RFGENHAIMGVAF ¹⁵⁹ 228 ⁸ TVKEAAQQQESATTQ ²⁴⁴ 276 ¹ IFTHQGSDFGPIFMTPAFF ²⁹⁴ 325 ² NPLGDDDEASTTVSK ³³⁹

Receptor and accession #	Enzymatic combination	Unique peptides	score	Sequence coverage (%)	Peptides
ABCA1 095477	Trypsin	24	2037	13% (299/2261)	230 ^S NNMDILKPILR ²⁴¹ 320 ^S LNWYEDNNYK ³³¹ 382 ^I LYTPDTPATR ³⁹³ 500 ^F MECVNLNK ⁵⁰⁹ 862 ^S YWFGEESDEK ⁸⁷³ 896 ^L GVSIQNVLK ⁹⁰⁶ 1039 ^K LSVALAFVGGSK ¹⁰⁵² 1052 ^V VILDEPTAGVDPYSR ¹⁰⁶⁸ 1082 ^T ILSTHHMDEADVLGDR ¹¹⁰⁰ 1108 ^L CCVGSSFLK ¹¹¹⁹ 1119 ^N QLGTGYLTLVK ¹¹³² 1133 ^D VESLSSCR ¹¹⁴⁵ 1250 ^V AEESGVDAETSDGTLPAR ¹²⁶⁹ 1303 ^E TDLLSGMDGK ¹³¹⁴ 1323 ^L TQQQFVALLWK ¹³³⁵ 1490 ^Q NTADILQDLTGR ¹⁵⁰³ 1839 ^F VSPLSWDLYGR ¹⁸⁵¹ 1884 ^L SPLNDEDEDVR ¹⁸⁹⁶ 1901 ^L DGGGGNDILEIK ¹⁹¹⁵ 2023 ^V GEWAIR ²⁰³⁰ 2081 ^F LWNCALSVVK ²⁰⁹² 2134 ^F GDGYTTVVR ²¹⁴⁴ 2145 ^I AGSNPDLKPVQDFFGLAFPGSVLK ²¹⁶⁹ 2189 ^I FSILSQSK ²¹⁹⁸
					992 ^D IETSLDAVR ¹⁰⁰² 1556 ^Y GGISIGK ¹⁵⁶⁵ 1911 ^E PIVDEDDVAAER ¹⁹²⁵ 1945 ^T YPGTSSPAVDR ¹⁹⁵⁷ 1983 ^M LGTDTTITSGDATVAGK ²⁰⁰¹ 2049 ^V ANWSIK ²⁰⁵⁶
ABCA4 P78363	Trypsin	6	2601	3% (70/2273)	