

# Biophysical and structural investigation of bacterially expressed and engineered CCR5, a G protein-coupled receptor

Maciej Wiktor · Sébastien Morin · Hans-Jürgen Sass · Fabian Kebbel · Stephan Grzesiek

Received: 10 July 2012 / Accepted: 22 November 2012 / Published online: 11 December 2012  
© Springer Science+Business Media Dordrecht 2012

**Abstract** The chemokine receptor CCR5 belongs to the class of G protein-coupled receptors. Besides its role in leukocyte trafficking, it is also the major HIV-1 coreceptor and hence a target for HIV-1 entry inhibitors. Here, we report *Escherichia coli* expression and a broad range of biophysical studies on *E. coli*-produced CCR5. After systematic screening and optimization, we obtained 10 mg of purified, detergent-solubilized, folded CCR5 from 1L culture in a triply isotope-labeled ( $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ ) minimal medium. Thus the material is suitable for NMR spectroscopic studies. The expected  $\alpha$ -helical secondary structure content is confirmed by circular dichroism spectroscopy. The solubilized CCR5 is monodisperse and homogeneous as judged by transmission electron microscopy. Interactions of CCR5 with its ligands, RANTES and MIP-1 $\beta$  were assessed by surface plasmon resonance yielding  $K_D$  values in the nanomolar range. Using size exclusion chromatography, stable monomeric CCR5 could be isolated. We show that cysteine residues affect both the yield and oligomer distribution of CCR5. HSQC spectra suggest that the transmembrane domains of CCR5 are in equilibrium

between several conformations. In addition we present a model of CCR5 based on the crystal structure of CXCR4 as a starting point for protein engineering.

**Keywords** CCR5 · G protein-coupled receptor · Expression in *E. coli* · NMR · Homology modeling

## Introduction

G protein-coupled receptors constitute a large protein superfamily found only in eukaryotes. About 4 % of the protein-coding human genome codes for  $\sim 800$  GPCRs (Takeda et al. 2002). Based on phylogenetic analysis human GPCRs cluster into 5 main families: rhodopsin, adhesion, frizzled/taste2, glutamate and secretin, which comprise 701, 24, 24, 15 and 15 members, respectively (Fredriksson et al. 2003). The diversity of the GPCR superfamily members is reflected in the variety of their ligand types. Photons, ions, odorants, nucleotides, fatty acids, amino acids, peptides and proteins are only some of the messages that GPCRs can transduce (Overington et al. 2006). As GPCRs regulate so many physiological processes such as vision, smell, behavior, mood, immune system, blood pressure, heart rate, digestion or homeostasis, they remain the most commonly drugged protein family (Bockaert and Pin 1999). About 40 % of prescribed pharmaceuticals target GPCRs (Filmore 2004).

The structure determination of membrane proteins is notoriously difficult due to the many obstacles impeding membrane protein sample preparation and subsequent structure determination. When this publication was written, the Protein Data Bank (Bernstein et al. 1977) contained about 86,000 entries, but only 364 unique membrane protein 3D structures ([http://blanco.biomol.uci.edu/Membrane\\_](http://blanco.biomol.uci.edu/Membrane_)

**Electronic supplementary material** The online version of this article (doi:10.1007/s10858-012-9688-4) contains supplementary material, which is available to authorized users.

M. Wiktor · S. Morin · H.-J. Sass · S. Grzesiek (✉)  
Focal Area Structural Biology and Biophysics, Biozentrum,  
University of Basel, Klingelbergstrasse 50/70, 4056 Basel,  
Switzerland  
e-mail: stephan.grzesiek@unibas.ch

F. Kebbel  
Center for Cellular Imaging and NanoAnalytics (C-CINA),  
Biozentrum, University of Basel, Mattenstrasse 26, 4058 Basel,  
Switzerland

[Proteins\\_xtal.html](#)) (White 2004). Solved GPCR structures are even sparser. Until now 16 unique GPCR structures have been solved by X-ray crystallography: the first being bovine rhodopsin (Palczewski et al. 2000) followed by  $\beta_2$ -adrenergic (Rasmussen et al. 2007; Cherezov et al. 2007),  $\beta_1$ -adrenergic (Warne et al. 2008), adenosine  $A_{2A}$  (Jaakola et al. 2008; Lebon et al. 2011), dopamine D3 (Chien et al. 2010), CXCR4 (Wu et al. 2010) and several others. To obtain high-resolution structural data the replacement of the intracellular (IC) loop three with T4 lysozyme (Rasmussen et al. 2007), thermostabilization (Warne et al. 2008) or stabilization by anti- or nanobodies (Rasmussen et al. 2007) proved to be successful strategies. Additionally, all crystallized GPCRs were bound to an agonist (Lebon et al. 2011), an inverse agonist (Cherezov et al. 2007; Kruse et al. 2012; Rasmussen et al. 2007) or most often to an antagonist (Chien et al. 2010; Haga et al. 2012; Hanson et al. 2012; Jaakola et al. 2008; Manglik et al. 2012; Shimamura et al. 2011; Warne et al. 2008; Wu et al. 2010, 2012; Granier et al. 2012). Although not GPCRs, prokaryotic sensory rhodopsin II (Gautier et al. 2008) and proteorhodopsin (Reckel et al. 2011) are examples of 7-TM domain proteins solved by solution NMR spectroscopy. Very recently the structure of *E. coli*-expressed and refolded CXCR1 has been determined in phospholipid bilayers using solid state NMR spectroscopy (Park et al. 2012).

CCR5 (CC chemokine receptor 5) belongs to the  $\gamma$ -group of the rhodopsin family of GPCRs. It is found in the plasma membrane of Th1 lymphocytes, macrophages, NK cells and immature dendritic cells and is involved in various infectious and inflammatory diseases as well as cancer (Balistreri et al. 2007). Since humans carrying the  $\Delta 32$  allele of the CCR5 gene, a 32-base pair deletion resulting in a premature stop codon in the extracellular (EC) loop 2 and a nonfunctional receptor, are healthy, the exact role of CCR5 is not completely understood.

The main interest in CCR5 is, however, a consequence of its involvement in AIDS. R5-tropic HIV-1 infection necessitates the sequential interaction of viral envelope glycoprotein gp120 with CD4 and CCR5 (Choe et al. 1998). Two copies of the CCR5- $\Delta 32$  allele confer nearly complete resistance to HIV-1 infection (Liu et al. 1996; Samson et al. 1996).  $\Delta 32$  occurs at 5–14 % frequency in European Caucasians but not in African, Native American, and East Asian populations (Stephens et al. 1998). This is hypothesized to be a result of pandemics that took place in Europe in medieval ages (Duncan et al. 2005). Successful strategies to block HIV-1 entry have been developed based on small-molecule inhibitors of CCR5 (Kondru et al. 2008) as well as derivatives of its natural chemokine ligand RANTES (Gaertner et al. 2008; Lederman et al. 2004; Lusso et al. 2011; Nardese et al. 2001).

High-resolution structural data would greatly improve the understanding of CCR5 function and the nature of its interaction with the chemokine ligands RANTES, MIP-1 $\alpha$ , and MIP-1 $\beta$ , as well as substantially enhance possibilities for anti-HIV-1 drug discovery. So far it has been very challenging to obtain sufficient amounts of this protein suitable for structural studies. Large-scale CCR5 expression at the yield of 1 mg/L was reported in insect cells (Nisius et al. 2008) where screening for mutants is time-consuming and isotope labeling is very costly and has not been achieved for deuterium.

Alternatively, 1–3 mg/L of CCR3 but only 0.1–0.3 mg/L of CCR5 was obtained from *E. coli* after fusing the N-terminus of the chemokine receptor to the C-terminus of thioredoxin (Ren et al. 2009). However, the described expression system relied on the usage of rich TB medium and ligand binding of the expressed chemokine receptors was not shown. Nevertheless, there is a growing number of various GPCRs functionally expressed in *E. coli* (Attrill et al. 2009; Baneres et al. 2003; Dodevski and Pluckthun 2011; Furukawa and Haga 2000; Krepkiy et al. 2006; Shibata et al. 2009; Weiss and Grishammer 2002), including the chemokine receptor CXCR1, which was expressed as a GST-CXCR1 fusion construct in  $^{15}\text{N}/^{13}\text{C}$ -labeled form at 5 mg/L and after reconstitution to proteoliposomes could bind IL-8 and activate G protein (Park et al. 2012).

Petrovskaya et al. have compared direct expression of 17 diverse GPCRs in *E. coli* to hybrid expression with the N-terminal fusion partners OmpF or Mistic (Petrovskaya et al. 2010). Interestingly, almost all GPCRs expressed in the presence of a fusion partner at >5 mg/L yield, but for most the expression was severely reduced in its absence. Thanks to a better access to isotope labeling bacterial or yeast expression systems are preferred for NMR, however, a significant progress has been recently made in isotope labeling in mammalian cells, which, unlike bacteria or yeast, provide possibilities to obtain human posttranslational modifications (Sastry et al. 2011).

Here we report a CCR5 production platform that yields up to 10 mg of purified protein per 1 L of bacterial culture. CCR5 is solubilized from *E. coli* without the requirement of refolding. As the expression conditions were optimized in minimal medium, triple isotope ( $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ ) labeling does not compromise the yield. In order to boost the expression, we fused the N-terminus of CCR5 to well expressing small proteins or signal sequences. A C-terminal 10His-tag and rigorous washing conditions yield over 90 % purity after a single IMAC purification step. The fusion partner can be readily and quantitatively cleaved off by thrombin and separated on a size exclusion column, where CCR5 monomers and dimers migrate as separate symmetric peaks. Both monomers and dimers are monodisperse and homogeneous as judged from electron micrographs. The expected  $\alpha$ -helical

secondary structure content is confirmed by circular dichroism (CD) spectroscopy. When solubilized in a DDM/CHAPS/CHS/DOPC mixture CCR5 interacts with RANTES, MIP-1 $\beta$  and 2D7 with nanomolar affinities. Recorded  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra suggest that the TM domains of CCR5 are in equilibrium between several conformations. We also show that the number of cysteine residues has a severe impact on both protein yield and oligomeric state. Following Hernanz-Falcon et al. (2004) two point mutations I52V and V150A were introduced to reduce the tendency of dimer formation, but no such reduction was observed. Our system establishes a high-yield platform for biophysical and structural studies on CCR5.

## Materials and methods

### Generation of expression constructs

Plasmids pET28F10 and pMT10H10 containing the CCR2b sequence fused to OmpF and Mistic were a generous gift from Prof. A. Arseniev (Russian Academy of Sciences, Moscow, Russia). Plasmid pCA528 was kindly provided by Prof. A. Spang (Biozentrum, Basel, Switzerland). pET vectors were obtained from Novagen. The *E. coli*-optimized CCR5 DNA sequence in the pQE-T7 vector was generated by GeneArt. The CCR5 gene was cloned using standard molecular biology techniques. Plasmid DNA was amplified with the QIAprep Spin Miniprep Kit (Qiagen). Point mutations were carried out using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). DNA sequences of the cloned constructs can be found in the supporting information Text S1.

### Protein expression

Freshly transformed Rosetta 2 (DE3) Competent Cells (Novagen) were transferred to 1–2 L of M9 medium after overnight growth on LB agar plates. The cultures were shaken in 5 L baffled flasks at 100 rpm at 37 °C until  $\text{OD}_{600} = 2.6$ – $2.8$ . The cultures were cooled down on ice with occasional shaking until the temperature dropped to 20–25 °C. CCR5 expression was induced with 1 mM IPTG and the cultures were shaken at 100 rpm at 20 °C. After harvesting, cells were pelleted and stored at  $-70$  °C.

For expression in  $\text{D}_2\text{O}$  transformed cells were grown on LB agar plates prepared in 50 %  $\text{D}_2\text{O}$ . 1–2 L cultures were preceded by 100 mL precultures grown until  $\text{OD}_{600} = 1$ . All compounds used in the preparation of M9 medium in  $\text{D}_2\text{O}$  (including trace elements, vitamins, antibiotics) were prepared in 99.8 %  $\text{D}_2\text{O}$ . Uniform  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeling was carried out using  $^{15}\text{NH}_4\text{Cl}$  (98 %  $^{15}\text{N}$ , 1 g/L), and [ $^1\text{H}$ / $^{13}\text{C}_6$ ]-D-glucose (99 %  $^{13}\text{C}$ , 4 g/L) as the sole nitrogen

and carbon sources, respectively. Hence, the labeling efficiency is expected as  $\sim 98$  % for  $^{15}\text{N}$  and  $\sim 99$  % for  $^{13}\text{C}$ . Judging from strong peaks in the HN(CO)CA spectrum, which showed no signs of  $^1\text{J}_{\text{CH}}$  splitting in the absence of  $^1\text{H}$  decoupling during  $^{13}\text{C}^\alpha$  evolution, the deuteration efficiency is estimated as  $>80$  %. This is in agreement with the data of Otten et al. (2010) and is consistent with bacterial metabolism (Gottschalk 1986). Details of the M9 medium composition can be found in Text S2.

Expression of wild-type CCR5 in insect cells was performed as described previously (Nisius et al. 2008).

### Membrane fraction preparation

Frozen *E. coli* cell pellet (1 g) was suspended in 6–8 mL of buffer A (20 mM HEPES pH 7, 150 mM NaCl, 10 % (v/v) glycerol) supplemented with 0.5 mM PMSF, 5 mM benzamidine and EDTA-free complete protease inhibitor cocktail (Roche). Cells were broken using a French press at 31,600 psi. Cell debris was removed by centrifugation at 6,600g for 15 min. The supernatant was centrifuged at 126,000g for 15 min, and the resulting pellet (from now on called membrane fraction) collected. After suspending in buffer A, a 20 % (w/v) solution of the membrane fraction was stored at  $-70$  °C.

The preparation of the insect cell membrane fraction was carried out as described previously (Nisius et al. 2008).

### Detergent screening

Frozen 20 % (w/v) solutions of the membrane fraction were thawed, diluted twice and supplemented with detergent to the final concentration of 2 %. Solubilization was carried out at RT for 2 h with 1,000 rpm shaking. Unsolubilized material was removed by centrifugation at 100,000g for 30 min. The clarified supernatant (2  $\mu\text{L}$ ) was loaded onto a Protran BA85 nitrocellulose membrane (Whatman) and dried at RT. Dot blots were blocked, labeled with anti-His-tag antibody, developed and quantified in the same way as western blots described below. Detergents were obtained from Anatrace with the exception of 1,2-diheptanoyl-sn-glycero-3-phosphocholine (DHPC, Avanti Polar Lipids).

### Protein purification

A frozen 20 % (w/v) solution of membrane fraction was thawed and supplemented to a final concentration of 0.5 M NaCl, 20 mM KCl, 10 mM  $\text{MgCl}_2$  and 2.5 % FosCholine-12 (FC-12). Protein solubilization was carried out at 4–8 °C for 1–2 h. Unsolubilized material was removed by centrifugation at 126,000g for 30 min. The clarified supernatant was supplemented with 35 mM imidazole and bound to Ni-NTA beads (Qiagen) for 2 h. The resin was

washed with 100 column volumes of buffer B (20 mM HEPES pH 7, 1 M NaCl, 60 mM imidazole, 10 % (v/v) glycerol, 0.1 % FC-12). The protein was then eluted with buffer C (20 mM HEPES pH 7, 150 mM NaCl, 0.4 M imidazole, 0.15 % FC-12). Protein-rich fractions were pooled and dialyzed against buffer D (20 mM Tris pH 8, 150 mM NaCl, 0.5 mM EDTA, 0.1 % FC-12). To cleave the fusion partner, 2 U of thrombin per 1 mg of purified protein was sufficient to complete the cleavage over 16 h at RT. The protein was concentrated using a 30 kDa molecular weight cut off (MWCO) concentrator and injected onto Superdex 200 10/300 GL (analytical run) or Superdex 200 26/60 HiLoad (preparative run) columns equilibrated in buffer E (20 Na<sub>2</sub>HPO<sub>4</sub> pH 7.4, 180 mM NaCl, 0.1 % FC-12).

#### Gel electrophoresis and western blotting

Protein samples for SDS-PAGE were mixed with 5× SDS loading buffer (312.5 mM Tris–HCl pH 6.8, 50 % (v/v) glycerol, 25 % β-mercaptoethanol, 10 % SDS, 0.0125 % bromophenol blue), incubated at 30 °C for 15 min and centrifuged at 17,000g for 5 min prior to loading on a 4–20 % gradient precast gel (Pierce). The electrophoresis was performed at 100 V constant voltage. Gels were stained using 0.25 % solution of Coomassie Brilliant blue R-250 (AppliChem) in 25 % isopropanol and 10 % acetic acid and destained in 10 % acetic acid.

For western blotting onto PVDF membrane (Bio-Rad), a Criterion Blotter (Bio-Rad) was used. The transfer was performed at 0.5 A constant current for 1 h in the transfer buffer (48 mM Tris–HCl pH 9.2, 39 mM glycine, 0.375 % SDS, 20 % methanol). The membrane was blocked with 3 % BSA in TBST buffer (10 mM Tris–HCl pH 8, 150 mM NaCl, 0.5 % Tween-20). Subsequently the membrane was incubated with mouse monoclonal HIS-1 anti-polyhistidine-peroxidase antibody (Sigma-Aldrich) at 1:6,000 dilution for 1 h. After washing 4 × 2 min with TBST buffer, the blot was developed using chemiluminescent HRP substrate (Roche). The signal was recorded using a BioMax XAR Film (Kodak) or using a LAS-4000 luminescent image analyzer (Fujifilm). The signal intensities were quantified using ImageJ 1.43r (Abramoff et al. 2004).

#### Transmission electron microscopy

For transmission electron microscopy (TEM) analysis 5 μL of 10 μg/mL protein solution was adsorbed on carbon-coated copper 200 mesh grids rendered hydrophilic by glow discharge in air during 20 s. The grids were washed in five drops of double distilled water and negatively stained with two drops of 2 % uranyl acetate. Electron

micrographs were recorded on a Philips CM10 instrument equipped with a LaB6 filament operating at an accelerating voltage of 80 kV. Images were recorded at nominal defocus values of 0.5 μm on a Veleta CCD camera at a nominal magnification of 130,000×, corresponding to a pixel size of 3.7 Å at the sample level.

#### CD spectroscopy

Circular dichroism spectra were recorded on 3–13 μM monomeric CCR5 fractions. Measurements were performed on a Chirascan CD spectrometer (Applied Photophysics) at 20 °C in 1 mm quartz Suprasil cuvettes (Hellma). Typically, spectra in a wavelength range of 195–260 nm spectra were recorded in triplicates and averaged. After baseline (buffer) subtraction, the mean residue molar ellipticity  $\Theta_{\text{MRM}}$  was calculated from the following equation  $\Theta_{\text{MRM}} = \Theta / (C \times n \times l)$ , where  $\Theta$  is the ellipticity (deg),  $C$  is the concentration (mol/L),  $n$  is the number of residues and  $l$  is the optical path length (cm). The relative  $\alpha$ -helical content  $\alpha_r$  was calculated as follows  $\alpha_r = (-\Theta_{\text{MRM},222 \text{ nm}} + 3,000) / 39,000$  (Morrow et al. 2000), where  $\Theta_{\text{MRM}}$  is given in units of deg × cm<sup>2</sup> × dmol<sup>-1</sup>.

#### Surface plasmon resonance

Surface Plasmon resonance (SPR) interaction assays were performed using a T100 Biacore instrument (GE Healthcare) at 20 °C. The setup consisted of a CM5 chip on which an antibody against the His-tag (Qiagen) was immobilized, using amine coupling chemistry. The antibody (4,000–10,000 RU) could capture ~2,000–5,000 RU of recombinant His-tagged CCR5, solubilized from membranes using a detergent mixture of 1 % DDM, 1 % CHAPS, 0.2 % CHS, and 1 mM DOPC at pH 7. Thioredoxin removal was performed on the chip using 5 U of thrombin injected in 300 μL over 60 min (5 μL/min). Experiments were performed in buffer F (20 mM HEPES pH 7.0, 150 mM NaCl, 0.1 % DDM, 0.1 % CHAPS, 0.02 % CHS, 50 nM DOPC, 0.1 mg/mL BSA) with a flow rate of 50 μL/min. Signals were processed with the Biacore T100 Evaluation Software using double referencing with both a reference channel and blank injections.

#### NMR

Several samples of FC-12-solubilized m11CCR5 (monomeric fraction) produced in isotope labeled M9 medium were concentrated in a 30 kDa MWCO Ultracel-30 K Amicon Ultra Centrifugal Filter (Millipore) to 100–200 μM (~2–3 % FC-12) and supplemented with 5 % D<sub>2</sub>O. All spectra were recorded in Shigemi tubes on a Bruker DRX800 spectrometer equipped with a triple resonance

Z-gradient TCI cryoprobe.  $^1\text{H}$ - $^{15}\text{N}$  TROSY (Pervushin et al. 1997) spectra were acquired under various buffer and temperature conditions (see text) as data matrices of  $63^*(^{15}\text{N}, t_1) \times 512^*(^1\text{H}^{\text{N}}, t_2)$  data points (where  $n^*$  refers to the number of complex points) with acquisition times of 25 ms ( $^{15}\text{N}$ ) and 40 ms ( $^1\text{H}^{\text{N}}$ ). Standard three-dimensional triple resonance TROSY spectra for backbone assignment (Salzmann et al. 1999) were recorded on a sample of 200  $\mu\text{M}$  uniformly  $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled CCR5 in  $\sim 3\%$  FC-12 Foscholine at 20 °C. Experimental times were HNCO: 5.5 days, HNCA: 2.7 days, HN(CO)CA: 2.7 days, HN(CA)CO: 6.4 days, and HNCACB: 7.3 days. All spectra were processed using NMRPipe (Delaglio et al. 1995).

### CCR5 model building

The core of CCR5 (residues 19–298) was built using the SWISS-MODEL server (<http://swissmodel.expasy.org/workspace>) with the crystal structure of CXCR4 (3ODU (Wu et al. 2010), 32 % sequence identity) as a template. At the C-terminus of CCR5, helix H8 modeled based on the rhodopsin structure 3C9L (Stenkamp 2008) was added using VMD 1.9 (Humphrey et al. 1996). In addition, the N-terminus of CCR5 (residues 1–18) and another part of the C-terminus including palmitoylated cysteines (residues 312–331) were added as an extended amino acid chain. Residues 332–352 were not included to reduce computational time. Finally, sulfate groups were added to Tyr10 and Tyr14 as well as palmitoyl groups to Cys321, Cys323 and Cys324.

After each manipulation step the structure was energy-minimized and relaxed by a short molecular dynamic simulation (MD) run using NAMD 2.7 (Phillips et al. 2005). For these MD runs the protein was embedded in a lipid bilayer of 137 POPC molecules, hydrated with 10,774 TIP3 water molecules and neutralized by adding  $\text{Na}^+$  and  $\text{Cl}^-$  ions. The final structure was embedded in a bilayer of 188 POPC molecules, hydrated with 20,781 TIP3 water molecules, relaxed with several short ( $\leq 1$  ns) equilibration steps and finally equilibrated with a 10 ns MD run.

## Results

### Protein expression

Even though many approaches are described in the literature, there is no universally applicable strategy to obtain a high yield GPCR expression system. The selection of expression vector, bacterial strain, culturing conditions, etc. remains largely empirical. To increase the chance of achieving high yield, we tested the expression of CCR5 cloned into several different T7-inducible vectors containing various N- and C-terminal fusion partners/tags. The

summary of tested constructs can be found in Table 1. As we intended to use the expression system also for isotope labeling, expression was carried out in M9 minimal medium supplemented with Hutner's trace elements (Hutner et al. 1950). To neutralize the codon bias in some of the constructs we used Rosetta 2 (DE3) cells carrying the pRARE plasmid encoding for rare tRNAs.

GPCR overexpression was assayed by western blot for each of the cloned constructs. The expression in pET-22b and pQE-T7 vectors, which provide no or only a very small fusion partner, was clearly lower than in the others. This suggests that CCR5 expression yield benefits from the N-terminal fusion partner. However, the type of the fusion partner seems of much less importance than expected (Table 1). Therefore, shortly after the preliminary screening, the work was restricted to the TrxA-CCR5 fusion construct, which was selected because of its high yield, purity and convenience of separation, and since it can be directly compared to the analogous expression system for chemokine receptors developed by Ren et al. (2009).

For every tested fusion construct, the yield was significantly higher at 20 °C than at 37 °C (Fig. 1a, b). A further decrease of the temperature to 12 °C or a decrease of IPTG concentration from 1 mM to 0.1 mM resulted in a lower yield (data not shown). The highest yield was achieved at 20 °C at 24–48 h after induction (Fig. 1a–d).

For further optimization of the protein construct, it was important to anticipate the sequence-specific position of the secondary structure elements. Initially the constructs were based on the two-dimensional topology predicted by Oppermann (2004). However, after the crystal structure of CXCR4 (Wu et al. 2010) became available, we generated a homology model based on the latter structure and the C-terminal helix H8 of rhodopsin (Stenkamp 2008) using state-of-the-art molecular dynamics energy minimization in explicit solvent of CCR5 embedded in a lipid bilayer. The result of the simulation is shown as a full structural model in Fig. 2 and the subsequently derived secondary structure topology in Fig. 3a.

Anticipating problems with the formation of intermolecular disulfide bridges we have systematically tested the role of all 12 cysteines by the truncation of the cysteine-containing C-terminus (after N306 or R319) and site-directed mutagenesis of the remaining 9 cysteines in other regions. In these regions, solvent-exposed cysteines were mutated to serines, whereas cysteines in the TM domains were replaced by alanines. The locations of the respective residues are highlighted in Figs. 2, 3a, and the naming convention of the various mutants is listed in Fig. 3b.

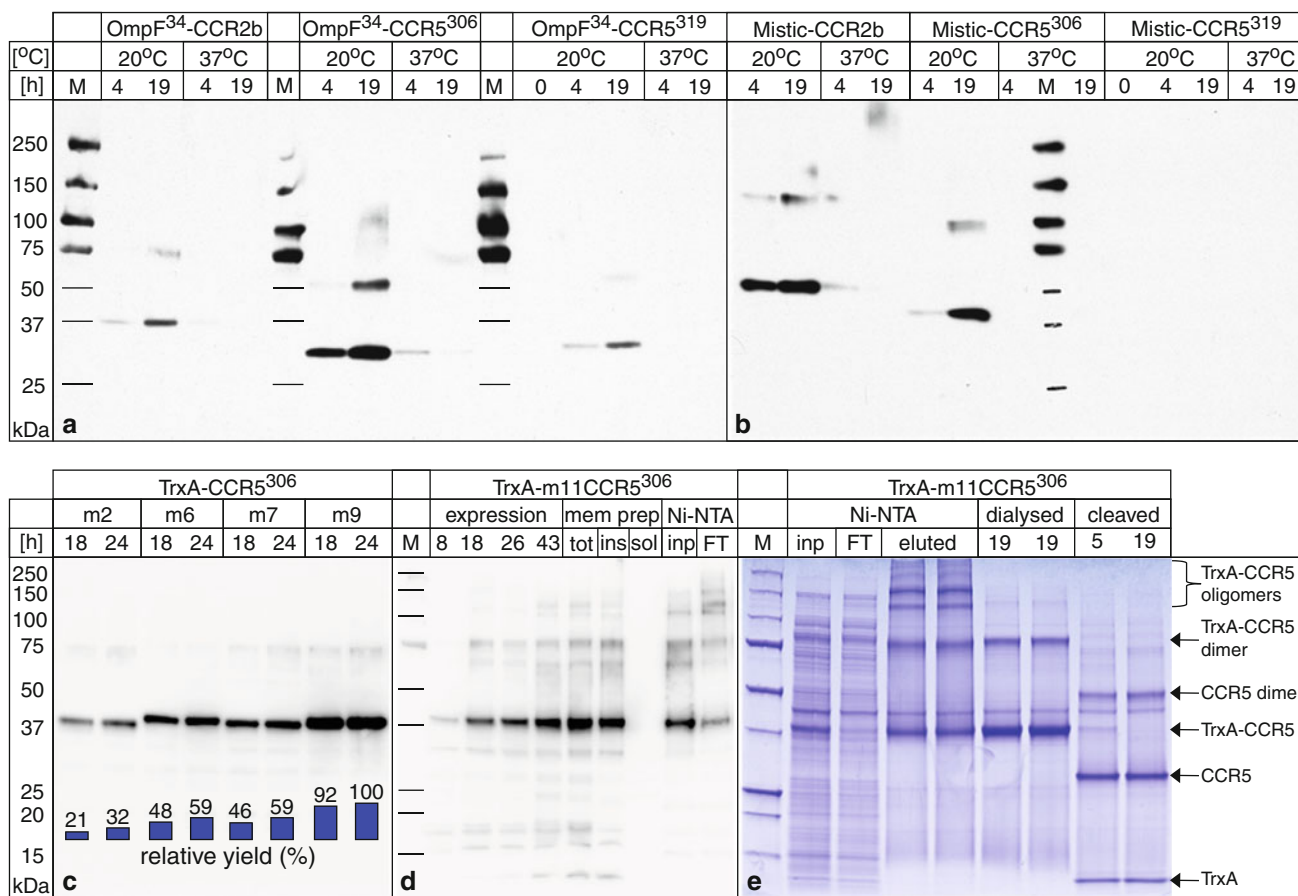
The expression of these cloned constructs was monitored by western blotting against the C-terminal His-tag. The signal from the shorter (1–306)  $\text{OmpF}^{34}$ -m2CCR5<sup>306</sup> and  $\text{Mistic-m2CCR5}^{306}$  constructs was stronger than from

**Table 1** Summary of GPCR constructs tested for expression

Vector	N-term. tag	Fusion partner	Cleavage site	GPCR <sup>b</sup>	cDNA	C-term. tag	Expression
pET28F10	–	OmpF (1–34/362) <sup>a</sup>	–	CCR2b [1]	<i>H. sapiens</i>	6His	+++
pMT10H10	–	Mistic (1–110/110)	Thrombin	CCR2b [5]	<i>H. sapiens</i>	10His	+++
pET-22b	–	pelB (1–22/374)	pelB	CCR5 [16]	<i>H. sapiens</i>	8His	+
pGEV2	–	GB1 (1–56/56)	Thrombin	CCR5 [15]	<i>H. sapiens</i>	8His	+++
pQE-T7	6His	–	TAGZyme	CCR5 [14]	<i>E. coli</i>	–	+
pET28F10	–	OmpF (1–34/362)	–	m7CCR5 <sup>306</sup> [4]	<i>E. coli</i>	6His	+++
pMT10H10	–	Mistic (1–110/110)	–	m7CCR5 <sup>306</sup> [8]	<i>E. coli</i>	10His	+++
pET-41a	–	GST (1–218/218)	–	m7CCR5 <sup>306</sup> [18]	<i>E. coli</i>	10His	+++
pCA528	6His	SUMO (1–98/101)	Ulp1	m7CCR5 <sup>306</sup> [17]	<i>E. coli</i>	10His	+++
pET-32b	–	TrxA (1–109/109)	Thrombin	m7CCR5 <sup>306</sup> [11]	<i>E. coli</i>	10His	+++

<sup>a</sup> Residues 1–34 from 362 total

<sup>b</sup> The most frequently used constructs are listed in this Table. However, other ones (differing with regards to cleavage site, cDNA and C-terminal tag) were also expressed. A comprehensive list of constructs with their DNA sequences is given in Text S1. The square brackets in Table 1 refer to the numbers in Text S1



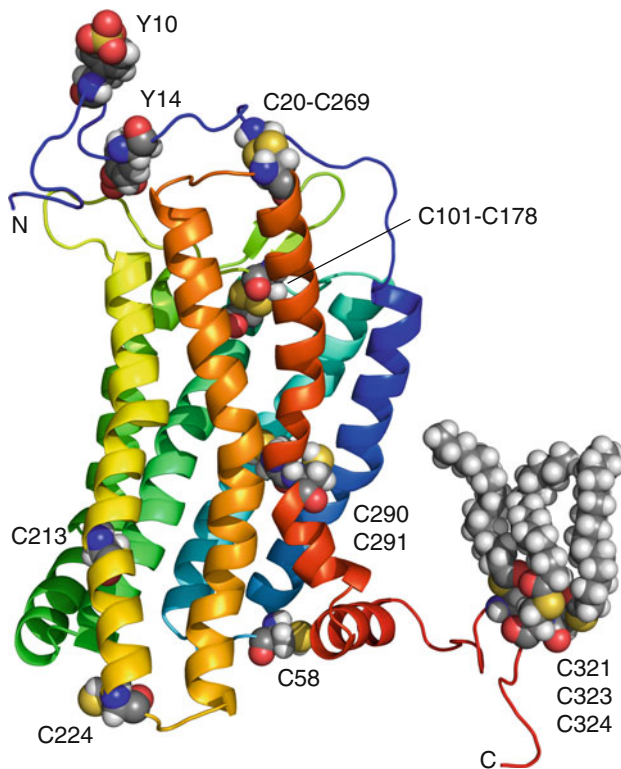
**Fig. 1** Summary of the expression and purification of various CCR5 constructs in *E. coli* monitored by western blot and SDS-PAGE. Comparison of the expression of longer (1–319) and shorter (1–306) versions of OmpF<sup>34</sup>-m2CCR5 (**a**) and Mistic-m2CCR5 (**b**) constructs at 20 °C and 37 °C. CCR2b constructs are used as a positive control. **c** Comparison of the expression of various Cys mutants of TrxA-CCR5<sup>306</sup>. **d** Expression, membrane preparation and binding to Ni-NTA of TrxA-m11CCR5<sup>306</sup>. Broken *E. coli* cells expressing CCR5

were centrifuged to remove cell debris. The remaining suspension (tot) was subsequently separated into insoluble membrane (ins) and soluble cytoplasmic (sol) fractions. CCR5 was found in the membrane fraction (ins) but not in the cytoplasmic fraction (sol). Solubilized membranes (inp) were loaded on Ni-NTA. **e** Purification of m11TrxA-CCR5<sup>306</sup>. After elution from Ni-NTA oligomerized m11TrxA-CCR5<sup>306</sup> was dialyzed and digested with thrombin

the longer (1–319) OmpF<sup>34</sup>-m2CCR5<sup>319</sup> and Mistic-m2CCR5<sup>319</sup> constructs (Fig. 1a, b). From this observation, we concluded that the shorter constructs were either expressing better or were more resistant to C-terminal degradation. Therefore, further work was limited to the shorter (1–306) CCR5 constructs (CCR5<sup>306</sup>). Within the latter, a negative correlation exists between the expression yield and the number of cysteine residues (Fig. 1c). Thus TrxA-m2CCR5<sup>306</sup> construct containing 7 cysteines (Fig. 3b) expressed worse than TrxA-m7CCR5<sup>306</sup> (4 Cys) or TrxA-m6CCR5<sup>306</sup> (3 Cys), and much worse than TrxA-m9CCR5<sup>306</sup> (0 Cys).

### Detergent screening

A good detergent for membrane protein studies should be able to solubilize the protein, keep it stable and functional in solution as well as allow structural studies. In order to explore the possible detergent space, we performed a systematic screen by solubilizing *E. coli* membrane fractions in various detergents at 2% (w/v) concentration. After removal of the unsolubilized material, the clarified solutions were dried on a nitrocellulose membrane and analyzed by dot blot using an anti-His antibody. The



**Fig. 2** Modeled 3D structure of CCR5 (residues 1–331) based on the CXCR4 structure (Wu et al. 2010). Sulfation of Tyr10 and Tyr14 as well as palmitoylated Cys321, Cys323 and Cys324 are depicted as spheres

chemiluminescent signal was quantified densitometrically and normalized to the maximum value (Fig. 4a).

The results indicate that OmpF<sup>34</sup>-m7CCR5<sup>306</sup> was efficiently solubilized by anionic (sodium dodecanoyl sarcosine and SDS) and zwitterionic detergents (FosCholines and dimethyl glycines) with aliphatic chains. The cationic trimethylammonium chlorides and the zwitterionic Anzergents were intermediate to moderate in their solubilization efficiency. Nonionic detergents (maltosides and Anapoes) turned out to solubilize OmpF<sup>34</sup>-m7CCR5<sup>306</sup> extremely poorly with the single exception of tetradecylmaltoside, which solubilized about a third as much as FosCholines.

These results on *E. coli* OmpF<sup>34</sup>-m7CCR5<sup>306</sup> are similar to a solubility screen carried out on wild-type CCR5 expressed in Sf21 cells (Fig. 4b). Analogous to *E. coli* CCR5, the insect cell protein was efficiently solubilized by sodium dodecanoyl sarcosine, SDS and FosCholines. Dimethyl glycines, Anzergents and trimethylammonium chlorides solubilized relatively worse and maltosides somewhat better, but still not very efficiently.

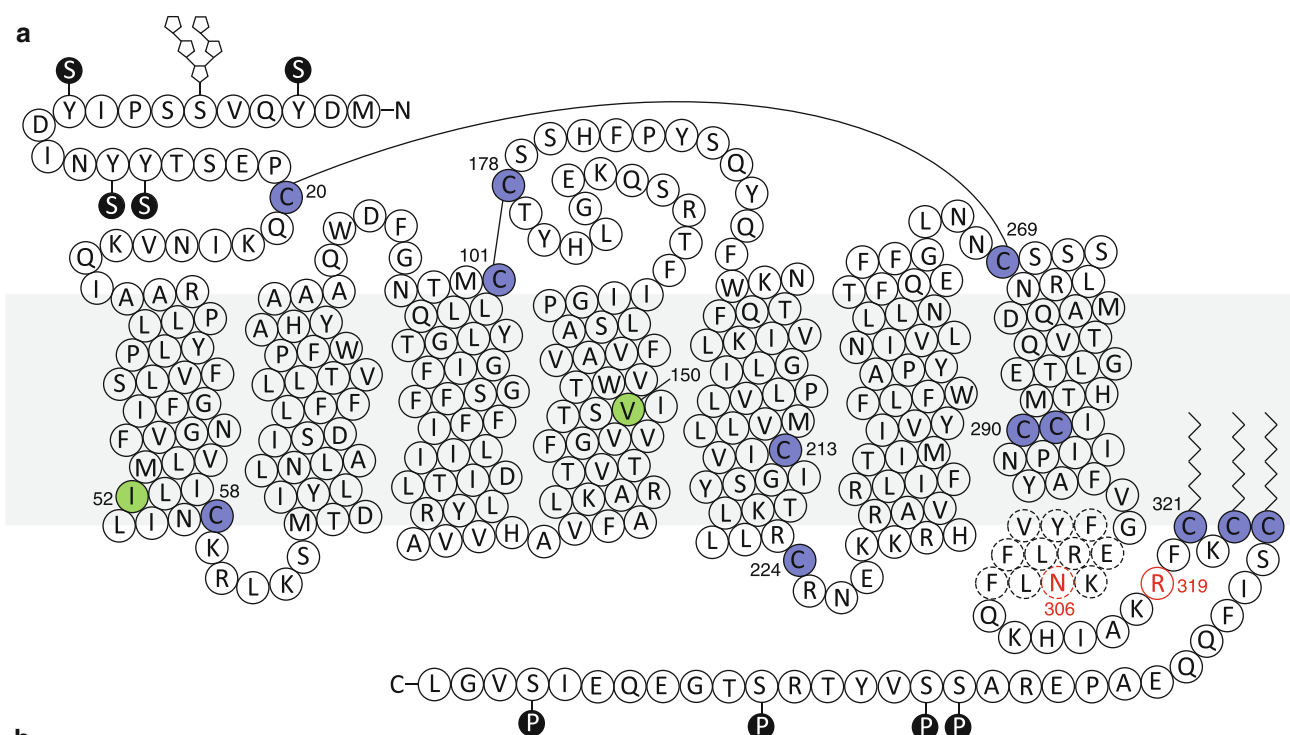
Due to its relatively mild character and lipid-like headgroup we picked FC-12 as the main working detergent. Even though FosCholines with longer hydrocarbon tails performed better, they are much less suitable for NMR due to their high aggregation number and lower solubility.

### Protein purification and identity confirmation

Considering a broad scope of applications we sought to establish a simple, robust and efficient purification scheme. Fractionation by centrifugation of the disrupted *E. coli* cells showed that the expressed TrxA-m11CCR5<sup>306</sup> was only present in the membrane fraction and the heavier cell debris fraction, but not in the soluble, cytoplasmic fraction (Fig. 1d). The isolated membrane fraction was readily solubilizable by a number of detergents (see Detergent screening section). Similarly, the receptor could also be solubilized from the cell debris. However, for most applications only the preparation from the lighter fraction was used.

The solubilized TrxA-m11CCR5<sup>306</sup> was purified in FC-12 using Ni-NTA chromatography resulting in up to 10 mg of ~90% pure (as estimated from SDS-PAGE) receptor per 1 L of *E. coli* culture (Fig. 1d, e). Interestingly, purification by Ni-NTA triggered TrxA-m11CCR5<sup>306</sup> oligomerization on SDS-PAGE, which was reversible by dialysis (Fig. 1e). The fusion partner was cleavable with thrombin (Fig. 1e). Other proteases were also tested (data not shown) including TEV and 3C protease with no (TEV) or partial success (3C).

Trials to solubilize the receptor in maltosides failed (data not shown). Some TrxA-m11CCR5<sup>306</sup> could be purified in tetradecylmaltoside but precipitated within few hours after elution from the Ni-NTA column. TrxA-



**Fig. 3** CCR5 topology and engineered mutations. **a** Membrane topology prediction of the human CCR5 according to the CXCR4 homology model (Fig. 2). The grey rectangle approximates the position of the membrane. EC (IC) space is at the top (bottom). The potential posttranslational modifications include sulfation of Y3, Y10, Y14 and Y15, phosphorylation of S336, S337, S342 and S349 (both marked as black circles), palmitoylation of C321, C323 and C324 as

well as glycosylation of S4. The positions of mutated residues are highlighted (C in blue, other in green). C-terminal truncations are marked with red circles and potential helix H8 with dashed lines. Disulphide bridges form between C20 and C269 and between C101 and C178. **b** Table summarizing the introduced point mutations of the listed CCR5 mutants

m11CCR5<sup>306</sup> solubilized in FC-12 followed by a detergent exchange to dodecylmaltoside on Ni-NTA also resulted in nearly complete protein precipitation.

The purified, cleaved m11CCR5<sup>306</sup> migrated on SDS-PAGE as a mixture of partially stable dimers at apparent MW of ~50 kDa and monomers at ~30 kDa (Fig. 1e). Both MW values are smaller than expected. This phenomenon is common for membrane proteins and can be caused by incomplete unfolding by SDS and/or by a larger relative amount of SDS bound as compared to the soluble protein standard. Besides monomers and dimers also higher order oligomers were often observed (Fig. 1e), especially after protein concentration.

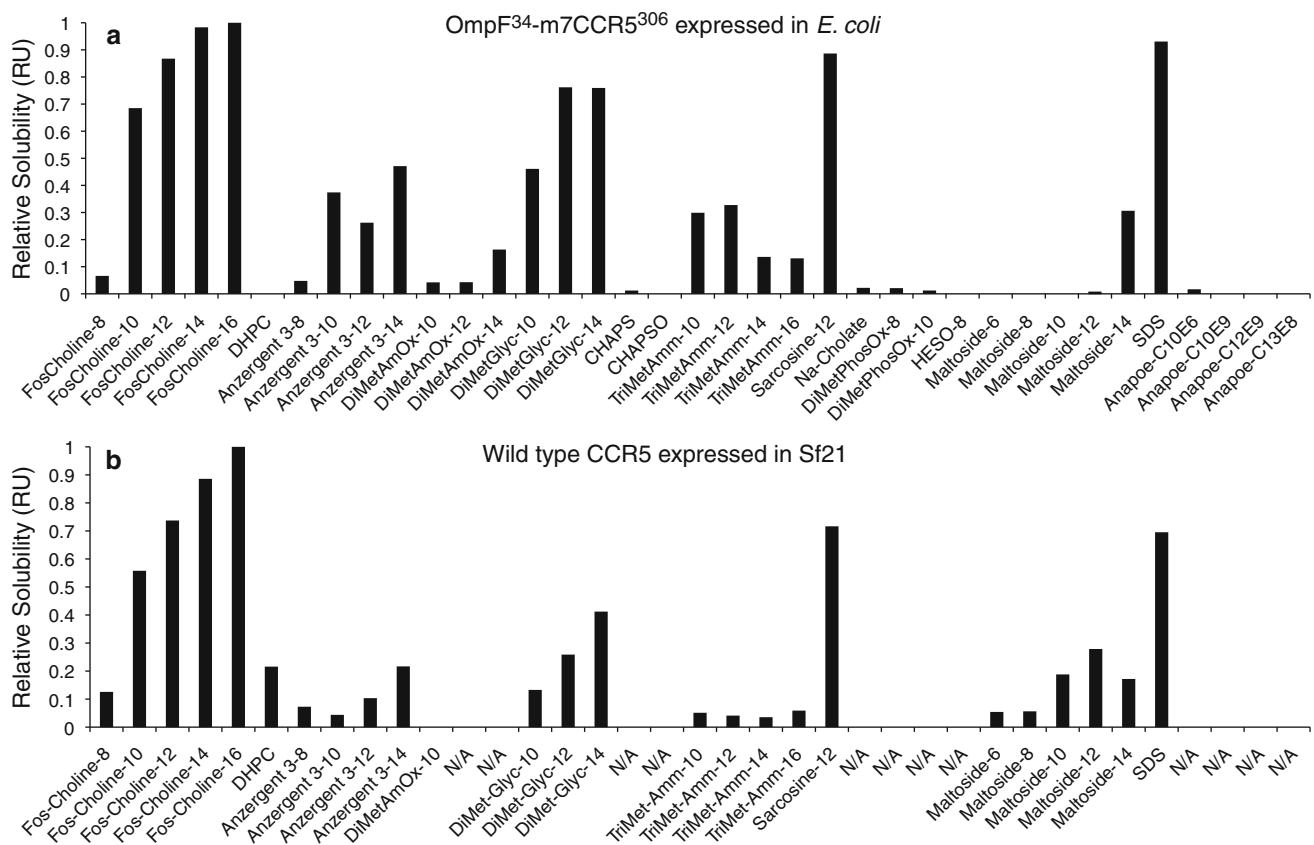
Discrete and sharp bands of CCR5 monomer and oligomers on the SDS-PAGE suggest that the primary structure of the protein is maintained (Figs. 1e and S1). The identity and integrity of the C-terminus of the expressed constructs were

confirmed by anti-His antibody western blotting (Fig. 1a–d). To further confirm the protein identity, trypsinized TrxAm7CCR5<sup>306</sup> and Mystic-m7CCR5<sup>306</sup> were analyzed by mass spectrometry. We were able to identify large stretches of fusion partners and the N-terminal fragment of the receptor in both monomer and oligomer (Fig. S1). Peptides from TM domains were not detectable, which suggests that the CCR5 core was resistant to proteolysis.

#### Characterization of CCR5 size distribution, stability and homogeneity

It is commonly observed that GPCRs form homo- and heterodimers as well as higher oligomeric structures. For both *E. coli* (Fig. 1) and insect cell expressed CCR5 (Nisius et al. 2008), besides monomers also oligomers are detected on SDS gels. The biological relevance of GPCR





**Fig. 4** Detergent screening for solubilization of OmpF<sup>34</sup>-m7CCR5<sup>306</sup> expressed in *E. coli* (a) and wild-type CCR5 expressed in Sf21 cells (b). Values were normalized against FC-16. DHPC, DiMetPhOx-10, *n*-decyl-*N,N*-dimethylamine-*N*-oxide; TriMetAmm-10, *N*-

dodecyltrimethylammonium chloride; Sarcosine-12, sodium dodecyl sarcosine; DiMetPhOx-8, dimethyloctylphosphine oxide; HESO-8, *N*-octyl-2-hydroxyethyl sulfoxide; Maltoside-6, *n*-hexyl-β-D-maltopyranoside

oligomerization is not clear. Since this heterogeneity also presents a problem for structural studies, the question of oligomerization was further investigated under non-reducing conditions using size exclusion chromatography.

After Ni-NTA purification and digestion by thrombin, cleaved TrxA-m11CCR5<sup>306</sup> was concentrated and injected onto a Superdex 200 column. The receptor migrated as a mixture of monomers, dimers and higher order oligomers (Fig. 5a). This observation is consistent with the results of the SDS-PAGE (Fig. 1e). Good protein separation was achieved on a 60 cm long size exclusion column. According to a column calibration with standard soluble proteins, the monomer and dimer peaks migrated similarly to particles of about  $95 \pm 3$  (SD) kDa and  $184 \pm 9$  kDa MW, respectively ( $N = 7$ ). This suggests that the monomeric (dimeric) protein micelle contains  $\sim 165$  ( $\sim 313$ ) FC-12 molecules. The ratio of monomer and dimer micelles depended on the stringency of Ni-NTA washing conditions, since higher imidazole concentrations depleted the monomeric fraction (data not shown). Apparently, this is due to the weaker binding of monomers to Ni-NTA.

Relative to the monomers and dimers, the fraction of higher order oligomers was much smaller.

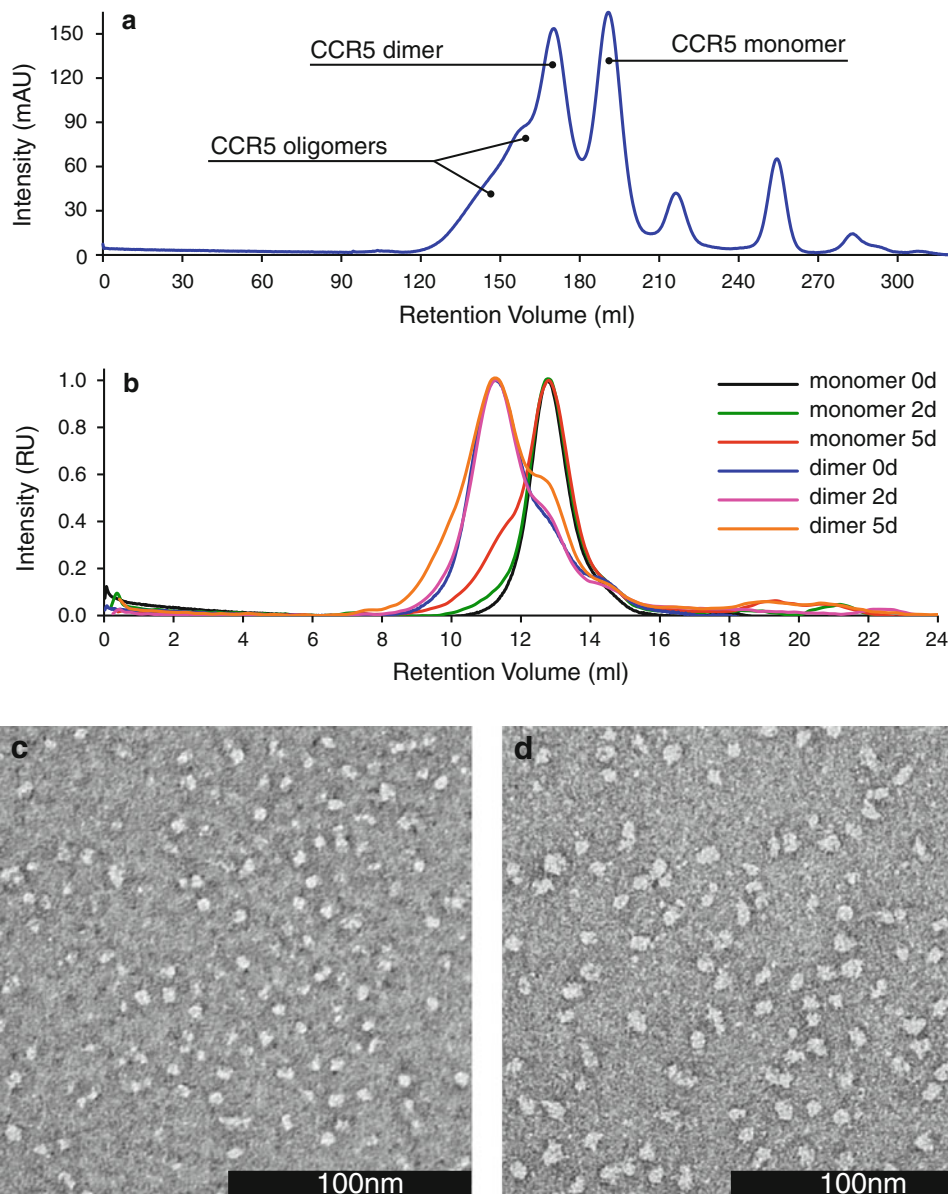
In order to assay the influence of disulphide formation on the quality of the preparation, several different cysteine-containing CCR5<sup>306</sup> mutants were compared to the cysteine-free mutant under non-reducing conditions by size exclusion chromatography (Fig. 6). The number of cysteines clearly correlates with enhanced oligomerization. The m2CCR5<sup>306</sup> mutant (7 Cys) formed the most oligomers, whereas m6CCR5<sup>306</sup> (3 Cys) and m7CCR5<sup>306</sup> (4 Cys) mutants were less oligomerized. Interestingly, the effect of EC Cys mutations (m6CCR5<sup>306</sup>) seems similar to the effect of TM Cys mutations (m7CCR5<sup>306</sup>), which suggests that both EC and TM Cys may mediate disulphide bond formation. The higher oligomer formation of the cysteine-containing mutants could be suppressed by the addition of a reducing agent (Fig. S2). Mutation of all Cys residues (m9CCR5<sup>306</sup> and m11CCR5<sup>306</sup>) resulted in a significant reduction of oligomerization, essentially rendering most of the protein monomeric or dimeric. Hence, it is likely that the remaining dimers and the residual higher oligomers are

stabilized by non-disulphide interactions, presumably between the TM domains.

As non-dimerizing CCR5 would be of advantage for structural studies, following the findings by Hernanz-Falcon et al. that point mutations I52V and V150A strongly reduce dimer formation in HEK-293 cells (Hernanz-Falcon et al. 2004), we tested these mutations in the m9CCR5<sup>306</sup> mutant, which does not contain cysteines that could lead to intermolecular disulphide bridges. In contrast to the *in vivo* findings (Hernanz-Falcon et al. 2004), these mutations did

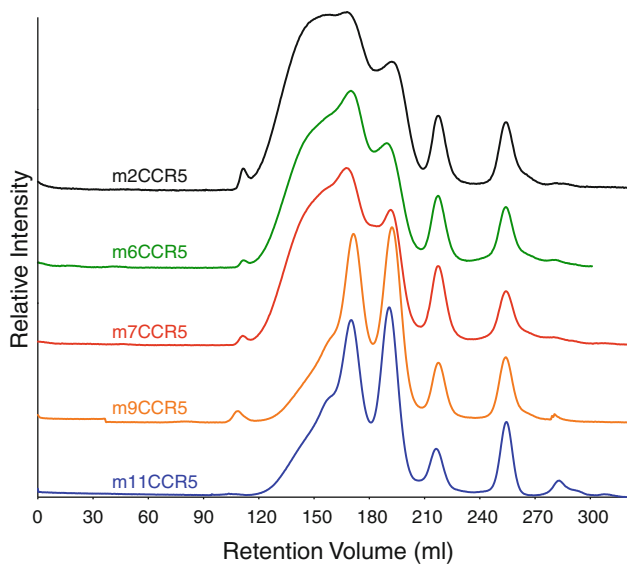
not reduce the dimerization propensity of the receptor (Fig. 6).

To assess the stability of m7CCR5<sup>306</sup> monomer and dimer preparations under reducing conditions, both fractions were concentrated to  $\sim 40 \mu\text{M}$  and incubated for 5 days at RT. After 2 days of incubation, almost no change in the size distribution was detected, whereas after 5 days only a small fraction of monomers interconverted to dimers and some of dimers fell apart to monomers or formed higher order oligomers (Fig. 5b). We tested a maximum monomer m7CCR5<sup>306</sup> concentration of

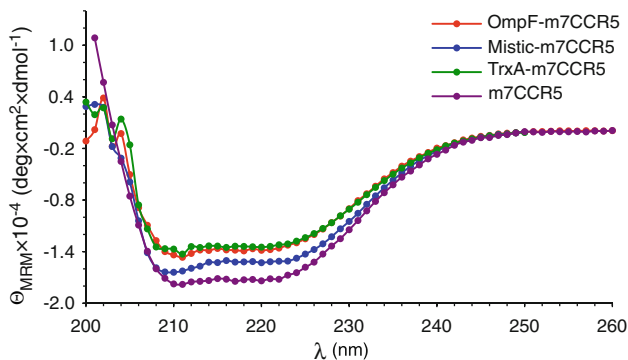


**Fig. 5** Monomers and dimers of m7CCR5<sup>306</sup> and m11CCR5<sup>306</sup>. **a** Size exclusion chromatography of m11CCR5<sup>306</sup> on a Superdex 200 HiLoad 26/60 column. The 60 cm long column enables isolation of monomers and dimers. **b** Stability test of m7CCR5<sup>306</sup> monomers and dimers. To prevent Cys oxidation 1 mM TCEP was included. Purified

monomers and dimers were concentrated separately to  $\sim 40 \mu\text{M}$  and re-run on a Superdex 200 10/300 GL column. For easier comparison all six chromatograms were scaled to one. Negative stain pictures of m7CCR5<sup>306</sup> monomers (**c**) and dimers (**d**)



**Fig. 6** Size exclusion chromatography of various CCR5<sup>306</sup> mutants demonstrates the impact of Cys residues on the oligomeric state of the purified receptor. No observable difference between m9CCR5<sup>306</sup> and m11CCR5<sup>306</sup> suggests that I52V and V150A mutations are not involved in CCR5 dimerization



**Fig. 7** CD spectra of OmpF<sup>34</sup>-m7CCR5<sup>306</sup>, Mistic-m7CCR5<sup>306</sup> and TrxA-m7CCR5<sup>306</sup> fusion constructs and of m7CCR5<sup>306</sup> (cleaved receptor without a fusion partner) normalized to the protein concentration. The data was recorded on 3–13 μM samples in 20 Na<sub>2</sub>HPO<sub>4</sub> pH 7.4, 180 mM NaCl, 0.1 % FC-12 at 20 °C on the monomeric receptor fractions isolated on size exclusion chromatography. Each spectrum shows characteristic features of α-helical secondary structure

137 μM, which also did not show any significant oligomerization after 4 days of incubation. Thus on the time scale of several days, both monomer and dimer preparations are very stable. The homogeneity of the monomeric and dimeric m7CCR5<sup>306</sup> was confirmed by negative stain TEM. Monomeric (Fig. 5c) and dimeric (Fig. 5d) particles had average diameters of ~6.6 and ~8.3 nm, respectively.

#### Characterization of CCR5 secondary structure

The secondary structure content of several m7CCR5<sup>306</sup> monomer preparations was assessed by CD. For all studied

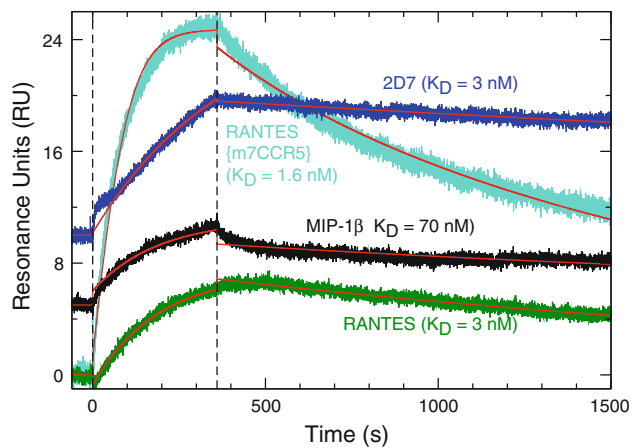
constructs we observed double minima at about 208 and 222 nm characteristic for α-helical proteins (Fig. 7). The helical content derived from the mean residue molar ellipticity  $\Theta_{\text{MRM},222 \text{ nm}}$  for Mistic-m7CCR5<sup>306</sup> (46 %) was slightly larger than for OmpF<sup>34</sup>-m7CCR5<sup>306</sup> (43 %) and TrxA-m7CCR5<sup>306</sup> (42 %). This can be explained by the fact that Mistic is a purely helical bundle and increases the  $\Theta_{\text{MRM}}$  of the whole fusion construct. This is not the case for the other fusion constructs, where the fusion partners contribute much less to  $\Theta_{\text{MRM}}$  due to their mixed α/β (TrxA-m7CCR5<sup>306</sup>) or likely β secondary structure (OmpF<sup>34</sup>-m7CCR5<sup>306</sup>). The 42 % α-helical content of TrxA-m7CCR5<sup>306</sup> is similar to the value of ~40 % obtained by Ren et al. for the thioredoxin-CCR3 fusion construct (Ren et al. 2009). For the m7CCR5<sup>306</sup> monomer, that is after removal of the fusion partner from TrxA-m7CCR5<sup>306</sup>, the CD signal was the strongest and indicated an α-helical content of 52 %. This is in a good agreement with the ~50 % helical content of a typical GPCR (Baneres et al. 2003, 2005).

To assess the thermal stability of the CCR5 preparation, the CD spectrum of TrxA-m7CCR5<sup>306</sup> was followed over the range from 5 to 95 °C in 5 °C increments (Fig. S3a). With increasing temperature the spectrum lost amplitude and its characteristic double minima. Decreasing the temperature from 95 °C back to 5 °C did not restore the initial shape and intensity, which indicates that denaturation was irreversible. The plot of the ellipticity at 222 nm against temperature (Fig. S3b) shows a very broad thermal transition between 20 and 80 °C. Low thermal stability is commonly observed for GPCRs. In the present case, this problem may be aggravated by a non-optimal membrane-mimicking detergent system, which lacks important lipids and the rigidity of the two-dimensional membrane, as well as the absence of stabilizing ligands.

#### Functional studies of CCR5

Due to the numerous differences in the expression machinery and the cellular environment, the production of functional GPCRs in heterologous systems is very challenging. To prove the proper folding and the functionality of our CCR5 preparation, we tested binding of several ligands to the receptor using SPR. High sensitivity, automation and high-throughput makes this method widely used in the GPCR field for screening ligands (Navratilova et al. 2011), solubilization (Navratilova et al. 2005) and crystallization (Navratilova et al. 2006) conditions.

For the SPR experiments, the receptor was solubilized in a DDM/CHAPS/CHS/DOPC mixture since a similar detergent/lipid composition was demonstrated to give best ligand binding activity for CCR5 and CXCR4 (Navratilova et al. 2005) as opposed to FC-12 where little binding could be detected. The protein was immobilized on the sensor chip via an anti-His-tag antibody. Subsequently, TrxA was



**Fig. 8** SPR binding assay of m7CCR5<sup>306</sup> (cyan) and m11CCR5<sup>306</sup> (other colors) solubilized in a DDM/CHAPS/CHS/DOPC mixture. The graph contains four overlaid independent runs, normalized to the amount of immobilized receptor and plotted to the same scale. Each run is composed of three phases separated by the dashed lines: equilibration, binding and dissociation. Data were fitted to a simple 1:1 binding model including a correction term for mass transport (red)

cut off from the N-terminus of the receptor by an injection of thrombin. As monitored by the decrease of the SPR signal, the cleavage efficiency was estimated to be typically about 70 % (Fig. S4).

Binding was assayed for the CCR5 chemokine ligands RANTES and MIP-1 $\beta$  as well as for the conformation-dependent antibody 2D7 (Fig. 8), which recognizes several residues from the second EC loop (Khurana et al. 2005). Each ligand showed fast binding and slow dissociation reactions.  $K_D$  values obtained from fitted  $k_{on}$  and  $k_{off}$  rates were all in the nanomolar range. m7CCR5<sup>306</sup> bound RANTES with a  $K_D$  of 1.6 nM. m11CCR5<sup>306</sup>, in which cysteines involved in disulphide bridge formation are absent, showed a two-fold decreased affinity ( $K_D = 3.1$  nM) and a three-fold decrease of the response amplitude. MIP-1 $\beta$  and 2D7 bound with 71 nM and 2.8 nM affinity, respectively.

For comparison, an identical experiment was performed with the wild-type CCR5 expressed in Sf21 cells. The obtained  $K_D$  values for RANTES, MIP-1 $\beta$ , 2D7 binding are 2.6, 200, and 0.1 nM, respectively. A summary of the performed experiments can be found in Table S1. Considering the differences in the protein constructs, the  $K_D$  values for *E. coli* and insect cell expressed CCR5 are in reasonable agreement. However, when refractive index amplitudes for ligand binding are normalized to the refractive index amplitudes of bound CCR5 (Table S1), it is evident that the amount of bound RANTES and MIP-1 $\beta$  is about 2–3 fold and of 2D7 about 15 fold reduced for *E. coli* m11CCR5<sup>306</sup>. We attribute this reduction to the lack of closed disulphide bridges at the extracellular side and the missing tyrosine sulfation in *E. coli*, which is important for chemokine binding (Bannert et al. 2001).

## NMR studies of CCR5

As opposed to crystal structures, which provide frozen snapshots of GPCR structures, NMR in principle can give simultaneous access to protein structure, dynamics and interactions. Thus it emerges as a promising method to rationalize GPCRs' function. However, due to the numerous challenges in the sample preparation, the success of NMR studies on GPCRs has been very limited so far.

To make our system suitable for NMR, the expression optimization was carried out directly in minimal medium. In this way isotope labeling does not compromise the final yield, which for detergent-solubilized, cleaved, monomeric m11CCR5<sup>306</sup> was 2 mg per 1L of cell culture in triply isotope-labeled (<sup>2</sup>H/<sup>15</sup>N/<sup>13</sup>C) minimal medium. For NMR measurements, samples were prepared from monomeric CCR5 fractions of the m11CCR5<sup>306</sup> mutant. To estimate the quality of the preparation <sup>1</sup>H-<sup>15</sup>N correlation spectra were recorded (Figs. 9 and S5). To optimize spectral quality, a variation of salt (0–180 mM NaCl), pH (4.2–7.4) and temperature (5–35 °C) was carried out. Optimal conditions were found at 20 °C, 0 mM NaCl and pH 4.2. Under these conditions, the spectra did not change over a period of few months. An increase in temperature to 35 °C gave only marginal improvement (data not shown). However, it had a destabilizing effect on the protein and caused a decrease of the NMR signal over time. The spectrum of m11CCR5<sup>306</sup> under optimal conditions (Fig. 9a) has a narrow dispersion, characteristic for an  $\alpha$ -helical protein. It contains on the order of 60–80 intense and narrow resonances that presumably correspond to flexible backbone amides in the N- and C-terminal tails and the interhelical loops. Furthermore, a background of many more broad resonances is observed that most likely correspond to protein core residues. The line broadening in this region may be related to intermediate conformational exchange and/or to the large size of the protein/detergent micelle.

An attempt was made to assign at least some of the better-resolved backbone resonances by three-dimensional triple resonance TROSY HNC(O), HN(CA)CO, HNCA, HN(CO)CA, and HNCACB experiments (Salzmann et al. 1999). Due to the low signal to noise ratio, unambiguous assignments could only be achieved for 21 residues within the CCR5 amino acid sequence. These are located at the N-terminus (M1-S7), in the loop between helix 6 and 7 (F264-S270), in helix 7 (L285-T288) as well as in the putative helix 8 (V300-E302). The secondary <sup>13</sup>C $^{\alpha}$ , <sup>13</sup>C' and <sup>13</sup>C $^{\beta}$  shifts for these residues are indicated in Fig. 9b. It is obvious that most residues have close to random coil shifts consistent with higher flexibility and concomitant higher resonance intensity. However, residues L285-T288 show moderately positive ( $\sim 1$ –2 ppm) and residues V300-E302 larger positive ( $\sim 2$ –3 ppm) <sup>13</sup>C $^{\alpha}$  and <sup>13</sup>C' secondary

shifts, which are consistent with a helical structure. Since besides the flexible N-terminus only residues in the region of helix 7 had a high enough signal to noise ratio for assignment, one may speculate that the region of helix 7 displays increased flexibility or more generally a different time scale of motional averaging. However, due to the highly limited assignment, this statement should be considered as very preliminary.

## Discussion

Due to its involvement in HIV-1 infection, CCR5 is a major target for structural biology and the pharmaceutical industry. Despite that expression and purification schemes have been described for numerous GPCRs, there is a lack of an efficient isotope labeling platform for CCR5. 1 mg/L expression of CCR5 was reported in insect cells (Nisius et al. 2008) where screening for mutants is time-consuming and isotope labeling very costly. On the other hand, so far no high-yield expression in isotope-labeled form has been reported for CCR5 in *E. coli* where these limitations are not present (Ren et al. 2009). Our goal was to develop methods that allow structural and biophysical characterization in particular by NMR for CCR5 and potentially other GPCRs. Here, we have achieved large overexpression of CCR5 by fusing small stable protein domains or signal sequences to its N-terminus.

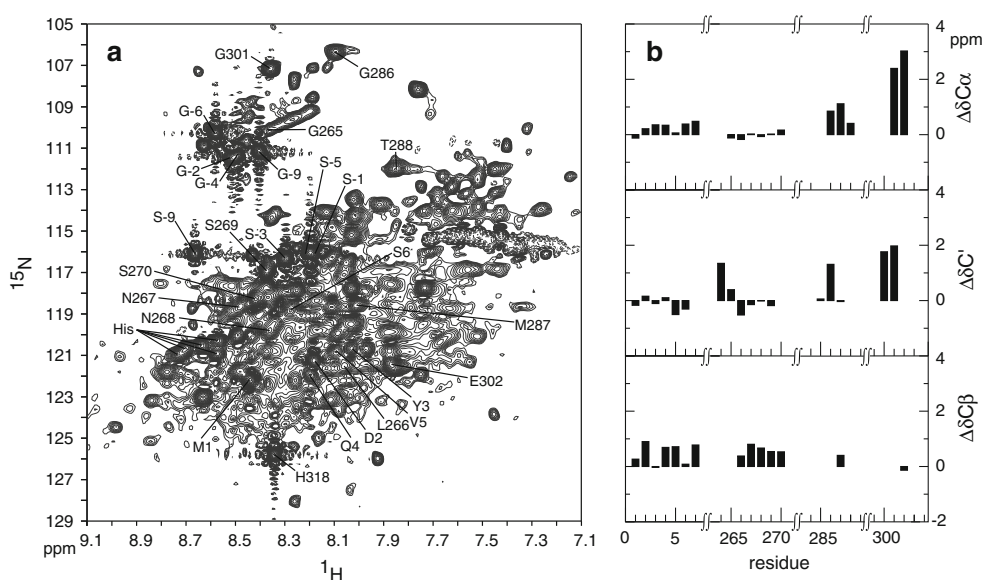
As the induction of CCR5 expression essentially arrested *E. coli* growth, increasing cell density proved to be a successful strategy to maximize the yield. The highest receptor overexpression was observed 24–48 h post induction at  $OD_{600} \sim 3$ . The induction at earlier or later phase of growth resulted in lower yields. Temperature had a dramatic effect on the expression level with the optimum

$\sim 20^\circ\text{C}$ . Variation of the CCR5 sequence also influenced the final yield. Thus the expression of the longer CCR5 constructs (1–319) seemed much lower than the expression of the shorter ones (1–306). The number of cysteines in the CCR5 sequence correlated negatively with the expression level. When all 9 Cys residues were mutated (TrxA-m9CCR5<sup>306</sup> and TrxA-m11CCR5<sup>306</sup>), the yield was highest, while it was lowest for TrxA-m2CCR5<sup>306</sup> (2 IC Cys mutated), i.e.  $\sim 1/3$  of TrxA-m11CCR5<sup>306</sup>.

A detergent screen revealed that charged detergents, especially anionic and zwitterionic were very efficient in  $\text{OmpF}^{34}$ -m7CCR5<sup>306</sup> solubilization. Nonionic detergents, with the exception of tetradecylmaltoside, which solubilized about  $\sim 1/3$  of available receptor, worked very poorly. A very similar solubility pattern was observed for the wild-type CCR5 receptor expressed in insect cells, i.e. there is good solubility in charged detergents and low solubility in nonionic detergents. In addition exchange trials from FC-12 to maltoside consistently failed for material from both expression systems. These observations suggest that the poor CCR5 solubility in maltoside detergents, which are widely used in GPCR research, is not unique to the receptor expressed in *E. coli* and therefore rather a consequence of the receptor's low stability than a problem specific to the bacterial expression. In this respect it should be noted that the homology to other chemokine receptors like CXCR4 and CXCR1, which have more favorable solubilization properties, is not very high, i.e. about 30 %. In particular, larger differences exist at the CCR5 C-terminus, which harbors 3 cysteine palmitoylation sites not present in CXCR1 and CXCR4.

The detergent screening results are in agreement with previous screens proposing FosCholines as promising candidates for CCR5 solubilization (Ren et al. 2009).

**Fig. 9** **a**  $^1\text{H}$ - $^{15}\text{N}$  TROSY spectrum of 112  $\mu\text{M}$  monomeric  $^2\text{H}/^{15}\text{N}$ -labeled m11CCR5<sup>306</sup> (5 mM sodium acetate pH 4.2, 5 %  $\text{D}_2\text{O}$ ,  $\sim 3$  % FC-12) recorded at  $20^\circ\text{C}$  on an 800 MHz spectrometer equipped with a cryoprobe with a total experimental time of 18 h. Assigned resonances are labeled. **b** Secondary  $^{13}\text{C}^\alpha$ ,  $^{13}\text{C}'$  and  $^{13}\text{C}^\beta$  shifts for residues in the CCR5 sequence, for which backbone assignments could be established



Unfortunately, a good surfactant for solubilization is not always also well suited for other purposes. For some applications, like the SPR functional assay, other detergents or detergent/lipid mixtures provide better receptor activity (Navratilova et al. 2011). Thus the search for an optimal detergent system or efficient detergent exchange protocols is still ongoing in our laboratory.

Protein oligomerization can severely decrease homogeneity of a sample and in this way compromise the quality of a sample for structural studies. In the case of CCR5 expressed in *E. coli*, the Cys residues, besides affecting the yield, also mediate oligomerization. Using size exclusion chromatography we have shown that the number of cysteines in CCR5<sup>306</sup> constructs correlates with the amount of oligomerized protein (Fig. 6). The fact that Cys-mediated oligomerization was also observed in the case of m7CCR5<sup>306</sup>, for which all but the EC Cys were mutated, may suggest that in our system, at least to some extent, EC disulphide bridges are not properly formed. On the other hand, the oligomerization of m6CCR5<sup>306</sup>, for which all but the TM cysteines C213, C290, C291 were mutated, implies that also TM Cys residues are reactive. This observation is consistent with the homology model (Fig. 2), where C213 and C291 are located on the surface of the CCR5 core and accessible for intermolecular disulphide formation.

When not jeopardized by intermolecular disulphide bridge formation, CCR5 forms a mixture of monomers, dimers and higher order oligomers. Due to their high stability, dimers and oligomers are also visible on SDS-PAGE. Both monomeric and dimeric species can be separated, concentrated and studied separately. The interconversion between monomers and dimers occurs after few days and goes both possible directions. As judged by TEM, both fractions are homogenous and monodisperse with a clear difference in size. Based on the retention volume, the size of monomers and dimers was estimated to be  $95 \pm 3$  (SD) kDa and  $184 \pm 9$  kDa, respectively (N = 7).

Based on computer modeling followed by the cross-linking of CCR5-transfected cells it was proposed that two point mutations together I52V and V150A yield a non-signaling, non-dimerizing mutant of CCR5 (Hernanz-Falcon et al. 2004). Such a non-dimerizing CCR5 mutant would be highly desirable for NMR studies. Unfortunately, the I52V and V150A mutants (m7CCR5<sup>306</sup> or m11CCR5<sup>306</sup>) did not exhibit significantly smaller propensity for dimerization in comparison to the non-mutated forms of CCR5 (Fig. 6). This is in line with results of co-immunoprecipitation and BRET experiments (Lemay et al. 2005) that contradict the impairment of CCR5 dimerization for these mutants. The involvement of these two residues in dimerization is further challenged by the recently published CXCR4 structure, which shows dimer interactions at unrelated surfaces, i.e. by helices V and VI (CXCR4 bound

to IT1t) or by the intracellular ends of helices III and IV (CXCR4 bound to CVX15) (Shimamura et al. 2011).

Due to its robustness, polyhistidine-tag chromatography is widely used as a first purification step. Using a 10His-tag we achieved strong binding and could apply more rigorous washing conditions without compromising the final yield. This resulted in  $\sim 10$  mg of purified TrxA-m11CCR5<sup>306</sup> from 1 L of *E. coli* culture. This is a considerable improvement over the previously described system, where  $\sim 0.3$  mg of Trx-hCCR5 per L was reported (Ren et al. 2009). Importantly, this yield is not compromised when isotope labeling including D<sub>2</sub>O is applied, which makes our system fully suitable for NMR studies. Out of 10 mg of TrxA-m11CCR5<sup>306</sup> oligomeric mixture it is possible to isolate 2 mg of cleaved monomeric m11CCR5<sup>306</sup>.

The quality of our preparations was assessed by CD, where all m7CCR5<sup>306</sup> constructs showed the characteristic features of an  $\alpha$ -helical secondary structure. NMR supports this observation as the HSQC spectrum of m11CCR5<sup>306</sup> exhibits, typical for  $\alpha$ -helical proteins, rather narrow peak dispersion ( $\sim 2$  ppm). Based on the circular dichroism data, we estimate that  $\alpha$ -helices constitute  $\sim 52\%$  of the sequence of m7CCR5<sup>306</sup> monomer which suggests that the receptor produced with our method has a correct secondary structure. The CD data indicate that the thermal stability of CCR5 is not very high. Some secondary structure is already lost at 5 °C but, as the amplitude of these changes is relatively small, it is difficult to judge their consequence on CCR5 structure and activity. Low thermal stability can be explained by several factors, most importantly suboptimal detergent system, lack of important lipids, absence of a ligand, nano/antibody or a small molecule drugs, that would stabilize CCR5.

The quality of the protein preparation was further validated by an SPR interaction assay. We observed nanomolar binding of RANTES to m7CCR5<sup>306</sup> ( $K_D = 1.6$  nM) and m11CCR5<sup>306</sup> (3.1 nM) and of MIP-1 $\beta$  to m11CCR5<sup>306</sup> (70 nM). These affinities are comparable to affinities of insect cell expressed CCR5 and within one order of magnitude to values observed in cellular binding assays (RANTES 0.38 nM, MIP-1 $\beta$  7.2 nM) (Bannert et al. 2001). The observed affinities may be affected by the lack of posttranslational modifications (tyrosine sulfation) in *E. coli*, which increase the affinity of CCR5 for chemokines (Bannert et al. 2001). Furthermore, compared to m7CCR5<sup>306</sup> RANTES affinity is weaker for m11CCR5<sup>306</sup>, which lacks the extracellular cysteines. This is consistent with the reported importance of the extracellular disulphide bridges for chemokine binding (Blanpain et al. 1999).

High-affinity (2.8 nM) binding of m11CCR5<sup>306</sup> was also observed for the 2D7 antibody, which is commonly used as a native conformation probe. Nevertheless, this affinity is considerably lower than for the insect cell CCR5 (0.1 nM).

In addition, the refractive signal amplitudes (Table S1) also suggest that the amount of bound ligand relative to m11CCR5<sup>306</sup> is about 2–3 times smaller for chemokines and about 15 times smaller for 2D7. Again the lack of tyrosine sulfation and extracellular disulphide bridges may be the reason for this finding. Thus further efforts are necessary to obtain higher activity, e.g. by proper refolding of disulphide bridges under controlled conditions.

Due to the substantial challenges in the preparation of isotope labeled samples, NMR spectra of GPCRs are very sparse in the literature. Therefore, only few HSQC spectra of <sup>15</sup>N labeled GPCRs have been reported, including the vasopressin V2 receptor (Tian et al. 2005), bovine rhodopsin (Werner et al. 2008), and the chemokine CXCR1 receptor (Park et al. 2011).

Here, we present a spectrum of uniformly <sup>15</sup>N-labeled m11CCR5<sup>306</sup>. Our initial HSQC spectrum had low dispersion and very broad lines besides for a number of apparently mobile terminal or loop residues. Similar observations have been made for other GPCRs (Park et al. 2011; Tian et al. 2005; Werner et al. 2008). However, the quality of the m11CCR5<sup>306</sup> spectrum could be improved by a decrease of the pH from 7.4 to 4.2 and the removal of salt, which reduced hydrogen exchange and increased the sensitivity of the measurement. Unfortunately, even with these improvements the quality of the spectra is still not sufficient for structural analysis and needs further improvement but presents a starting point in the NMR investigation of CCR5. Obviously, the key bottleneck is the severe line broadening, which may be the result of conformational heterogeneity of the TM domains and/or chemical exchange on an intermediate time scale in the microsecond to millisecond range. Therefore, CCR5 stabilization by locking it in a single conformation, may be an important step towards the improvement of the NMR spectra.

Due to major difficulties in sample preparation for structural studies, protein engineering is very common in the GPCR field. This process alters the protein sequence and may modify its native properties, but so far has been indispensable for gaining insights into the structure and function of this important class of proteins. Since a vast majority of GPCR structures were solved by X-ray crystallography, not surprisingly, alterations comprised stabilization (rigidification, fixation in selected conformations), removal of unstructured regions, introduction of a soluble domain into a loop, etc.

NMR spectroscopy on the other hand requires isotope labeling. However, so far efficient isotope labeling has not been possible for insect cells or natural tissues from which all except CXCR1 solved GPCRs were derived. Thus simple prokaryotic organisms, like *E. coli* are often the system of choice for an NMR spectroscopist, as they allow cost-effective isotope labeling in addition to fast access to

protein engineering. Due to the size limitations of NMR, the preparation of stable, monomeric and non-aggregating GPCRs is vital. Taking advantage of *E. coli*, we developed an efficient and robust CCR5 expression platform, which may find applications in biophysical, functional as well as structural characterizations of CCR5. We also believe that many of our observations have more general character and may be useful and applicable for other GPCRs.

**Acknowledgments** We thank Prof. A. Arseniev for providing the plasmids pET28F10 and pMT10H10 and Prof. A. Spang for the plasmid pCA528, Dr. Paul Jenö and Suzanne Moes for mass spectrometry analysis as well as Dr. Marcel Blommers, Dr. Lukasz Skora and Prof. Sebastian Hiller for stimulating discussions. This work was supported by the EU FP7 Combined Highly Active Anti-Retroviral Microbicides (CHAARM), SNF Grant 31-109,712 and SystemsX.ch (C-CINA).

## References

- Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image processing with imageJ. *Biophotonics Int* 11:36–42
- Attrill H, Harding PJ, Smith E, Ross S, Watts A (2009) Improved yield of a ligand-binding GPCR expressed in *E. coli* for structural studies. *Protein Expr Purif* 64:32–38
- Balistreri CR, Caruso C, Grimaldi MP, Listi F, Vasto S, Orlando V, Campagna AM, Lio D, Candore G (2007) CCR5 receptor: biologic and genetic implications in age-related diseases. *Ann NY Acad Sci* 1100:162–172
- Baneres JL, Martin A, Hullot P, Girard JP, Rossi JC, Parello J (2003) Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *J Mol Biol* 329:801–814
- Baneres JL, Mesnier D, Martin A, Joubert L, Dumuis A, Bockaert J (2005) Molecular characterization of a purified 5-HT4 receptor: a structural basis for drug efficacy. *J Biol Chem* 280:20253–20260
- Bannert N, Craig S, Farzan M, Sogah D, Santo NV, Choe H, Sodroski J (2001) Sialylated O-glycans and sulfated tyrosines in the NH2-terminal domain of CC chemokine receptor 5 contribute to high affinity binding of chemokines. *J Exp Med* 194:1661–1673
- Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The protein data bank: a computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542
- Blanpain C, Lee B, Vakili J, Doranz BJ, Govaerts C, Migeotte I, Sharron M, Dupriez V, Vassart G, Doms RW, Parmentier M (1999) Extracellular cysteines of CCR5 are required for chemokine binding, but dispensable for HIV-1 coreceptor activity. *J Biol Chem* 274:18902–18908
- Bockaert J, Pin JP (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J* 18:1723–1729
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318:1258–1265
- Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* 330:1091–1095

- Choe H, Martin KA, Farzan M, Sodroski J, Gerard NP, Gerard C (1998) Structural interactions between chemokine receptors, gp120 Env and CD4. *Semin Immunol* 10:249–257
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293
- Dodevski I, Pluckthun A (2011) Evolution of three human GPCRs for higher expression and stability. *J Mol Biol* 408:599–615
- Duncan SR, Scott S, Duncan CJ (2005) Reappraisal of the historical selective pressures for the CCR5-Delta32 mutation. *J Med Genet* 42:205–208
- Filmore D (2004) It's a GPCR world. *Mod Drug Discov* 7:24–28
- Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogue groups, and fingerprints. *Mol Pharmacol* 63:1256–1272
- Furukawa H, Haga T (2000) Expression of functional M2 muscarinic acetylcholine receptor in *Escherichia coli*. *J Biochem* 127:151–161
- Gaertner H, Cerini F, Escola JM, Kuenzi G, Melotti A, Offord R, Rossitto-Borlat I, Nedellec R, Salkowitz J, Gorochov G, Mosier D, Hartley O (2008) Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide. *Proc Natl Acad Sci USA* 105:17706–17711
- Gautier A, Kirkpatrick JP, Nietlispach D (2008) Solution-state NMR spectroscopy of a seven-helix transmembrane protein receptor: backbone assignment, secondary structure, and dynamics. *Angew Chem Int Ed Engl* 47:7297–7300
- Gottschalk G (1986) *Bacterial Metabolism*, 2nd edn. Springer, New York
- Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, Weis WI, Kobilka BK (2012) Structure of the delta-opioid receptor bound to naltrindole. *Nature* 485:400–404
- Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* 482:547–551
- Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, Sanna MG, Han GW, Kuhn P, Rosen H, Stevens RC (2012) Crystal structure of a lipid G protein-coupled receptor. *Science* 335:851–855
- Hernanz-Falcon P, Rodriguez-Frade JM, Serrano A, Juan D, del Sol A, Soriano SF, Roncal F, Gomez L, Valencia A, Martinez AC, Mellado M (2004) Identification of amino acid residues crucial for chemokine receptor dimerization. *Nat Immunol* 5:216–223
- Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14(33–8):27–28
- Hutner SH, Provasoli L, Schatz A, Haskins CP (1950) Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proc Am Philos Soc* 94:152–170
- Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* 322:1211–1217
- Khurana S, Kennedy M, King LR, Golding H (2005) Identification of a linear peptide recognized by monoclonal antibody 2D7 capable of generating CCR5-specific antibodies with human immunodeficiency virus-neutralizing activity. *J Virol* 79:6791–6800
- Kondru R, Zhang J, Ji C, Mirzadegan T, Rotstein D, Sankuratri S, Dioszegi M (2008) Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. *Mol Pharmacol* 73:789–800
- Krepkiy D, Wong K, Gawrisch K, Yeliseev A (2006) Bacterial expression of functional, biotinylated peripheral cannabinoid receptor CB2. *Protein Expr Purif* 49:60–70
- Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, Rosemond E, Green HF, Liu T, Chae PS, Dror RO, Shaw DE, Weis WI, Wess J, Kobilka BK (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* 482:552–556
- Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474:521–525
- Lederman MM, Veazey RS, Offord R, Mosier DE, Dufour J, Mefford M, Piatak M Jr, Lifson JD, Salkowitz JR, Rodriguez B, Blauvelt A, Hartley O (2004) Prevention of vaginal SHIV transmission in rhesus macaques through inhibition of CCR5. *Science* 306:485–487
- Lemay J, Marullo S, Jockers R, Alizon M, Brelot A (2005) On the dimerization of CCR5. *Nat Immunol* 6: 535; author reply 535–536
- Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86:367–377
- Lusso P, Vangelista L, Cimbri R, Secchi M, Sironi F, Longhi R, Faiella M, Maglio O, Pavone V (2011) Molecular engineering of RANTES peptide mimetics with potent anti-HIV-1 activity. *FASEB J* 25:1230–1243
- Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S (2012) Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* 485:321–326
- Morrow JA, Segall ML, Lund-Katz S, Phillips MC, Knapp M, Rupp B, Weisgraber KH (2000) Differences in stability among the human apolipoprotein E isoforms determined by the amino-terminal domain. *Biochemistry* 39:11657–11666
- Nardese V, Longhi R, Polo S, Sironi F, Arcelloni C, Paroni R, DeSantis C, Sarmientos P, Rizzi M, Bolognesi M, Pavone V, Lusso P (2001) Structural determinants of CCR5 recognition and HIV-1 blockade in RANTES. *Nat Struct Biol* 8:611–615
- Navratilova I, Sodroski J, Myszkowski DG (2005) Solubilization, stabilization, and purification of chemokine receptors using biosensor technology. *Anal Biochem* 339:271–281
- Navratilova I, Pancera M, Wyatt RT, Myszkowski DG (2006) A biosensor-based approach toward purification and crystallization of G protein-coupled receptors. *Anal Biochem* 353:278–283
- Navratilova I, Besnard J, Hopkins AL (2011) Screening for GPCR ligands using surface plasmon resonance. *ACS Med Chem Lett* 2:549–554
- Nisius L, Rogowski M, Vangelista L, Grzesiek S (2008) Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES. *Protein Expr Purif* 61:155–162
- Oppermann M (2004) Chemokine receptor CCR5: insights into structure, function, and regulation. *Cell Signal* 16:1201–1210
- Otten R, Chu B, Krewulak KD, Vogel HJ, Mulder FA (2010) Comprehensive and cost-effective NMR spectroscopy of methyl groups in large proteins. *J Am Chem Soc* 132:2952–2960
- Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5:993–996
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* 289:739–745
- Park SH, Casagrande F, Das BB, Albrecht L, Chu M, Opella SJ (2011) Local and global dynamics of the G protein-coupled receptor CXCR1. *Biochemistry* 50:2371–2380
- Park SH, Das BB, Casagrande F, Tian Y, Nothnagel HJ, Chu M, Kiefer H, Maier K, De Angelis AA, Marassi FM and Opella SJ



- (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature*. doi:10.1038/nature11580.
- Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94:12366–12371
- Petrovskaya LE, Shulga AA, Bocharova OV, Ermolyuk YS, Kryukova EA, Chupin VV, Blommers MJ, Arseniev AS, Kirpichnikov MP (2010) Expression of G-protein coupled receptors in *Escherichia coli* for structural studies. *Biochemistry (Mosc)* 75:881–891
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
- Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* 450:383–387
- Reckel S, Gottstein D, Stehle J, Lohr F, Verhoeven MK, Takeda M, Silvers R, Kainosho M, Glaubitz C, Wachtveitl J, Bernhard F, Schwalbe H, Guntert P, Dotsch V (2011) Solution NMR structure of proteorhodopsin. *Angew Chem Int Ed Engl* 50:11942–11946
- Ren H, Yu D, Ge B, Cook B, Xu Z, Zhang S (2009) High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS ONE* 4:e4509
- Salzmann M, Wider G, Pervushin K, Senn H, Wüthrich K (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J Am Chem Soc* 121:844–848
- Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyl-dermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382:722–725
- Sastry M, Xu L, Georgiev IS, Bewley CA, Nabel GJ, Kwong PD (2011) Mammalian production of an isotopically enriched outer domain of the HIV-1 gp120 glycoprotein for NMR spectroscopy. *J Biomol NMR* 50:197–207
- Shibata Y, White JF, Serrano-Vega MJ, Magnani F, Aloia AL, Grisshammer R, Tate CG (2009) Thermostabilization of the neurotensin receptor NTS1. *J Mol Biol* 390:262–277
- Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S (2011) Structure of the human histamine H1 receptor complex with doxepin. *Nature* 475:65–70
- Stenkamp RE (2008) Alternative models for two crystal structures of bovine rhodopsin. *Acta Crystallogr D Biol Crystallogr* D64: 902–904
- Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, Carrington M, Winkler C, Huttley GA, Allikmets R, Schriml L, Gerrard B, Malasky M, Ramos MD, Morlot S, Tzetis M, Oddoux C, di Giovine FS, Nasioulas G, Chandler D, Aseev M, Hanson M, Kalaydjieva L, Glavac D, Gasparini P, Kanavakis E, Claustres M, Kambouris M, Ostrer H, Duff G, Baranov V, Sibul H, Metspalu A, Goldman D, Martin N, Duffy D, Schmidtke J, Estivill X, O'Brien SJ, Dean M (1998) Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet* 62:1507–1515
- Takeda S, Kadowaki S, Haga T, Takaesu H, Mitaku S (2002) Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Lett* 520:97–101
- Tian C, Breyer RM, Kim HJ, Karra MD, Friedman DB, Karpay A, Sanders CR (2005) Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor. *J Am Chem Soc* 127:8010–8011
- Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454:486–491
- Weiss HM, Grisshammer R (2002) Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur J Biochem* 269:82–92
- Werner K, Richter C, Klein-Seetharaman J, Schwalbe H (2008) Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy. *J Biomol NMR* 40:49–53
- White SH (2004) The progress of membrane protein structure determination. *Protein Sci* 13:1948–1949
- Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330:1066–1071
- Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V, Stevens RC (2012) Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* 485:327–332