Comparison of Fragments Comprising the First Two Helices of the Human Y4 and the Yeast Ste2p G-Protein-Coupled Receptors

Xuan Shao,[†] Chao Zou,[†] Fred Naider,[‡] and Oliver Zerbe[†*]
[†]Institute of Organic Chemistry, University of Zurich, Zurich, Switzerland; and [‡]Department of Chemistry, College of Staten Island, City University of New York, Staten Island, New York

ABSTRACT Solution NMR techniques are used to determine the structure and the topology of micelle integration of a large fragment of the Y4 receptor, a human G-protein-coupled receptor, that contains the entire N-terminal domain plus the first two transmembrane (TM) segments. The structure calculations reveal that the putative TM helices are indeed helical to a large extent, but that interruptions of secondary structure occur close to internal polar or charged residues. This view is supported by ^{15}N relaxation data, amide-water exchange rates, and attenuations from micelle-integrating spin labels. No contacts between different helices are observed. This is in contrast to a similar TM1-TM2 fragment from the yeast Ste2p receptor for which locations of the secondary and the tertiary structure agreed well with the predictions from a homology model. The difference in structure is discussed in terms of principal biophysical properties of residues within central regions of the putative TM helices. Overall, using the biophysical scale of Wimley and White the TM regions of Ste2p display much more favorable free energies for membrane integration. Accordingly, the full secondary structure and the tertiary structure in TM1-TM2 of the Y4 receptor is likely to be formed only when tertiary contacts with other TM segments are created during folding of the receptor.

INTRODUCTION

The structural biology of G-protein-coupled receptors (GPCRs) has experienced dynamic progress during the last decade. Structures of a number of different GPCRs have been solved for both the ground state (1–7) and the activated state (8–12). This has resulted in a tremendous improvement in our understanding of the structural adaptations underlying the activation process of these receptors. Recently, the first reports on structures of the complex of a ligand, its GPCR, and the cognate G-protein appeared (13). All these structures were solved using x-ray crystallography and except for the rhodopsin structures all GPCRs required stabilizing mutations to enable crystallization and to render these receptors sufficiently thermostable for structural studies.

No structure of an entire GPCR has been solved from NMR solution data up to now, although the NMR studies of sensory rhodopsin II (14,15) and recent reports on the solution structure of proteorhodopsin (16) have indicated that these systems, in principle, are amenable to such analysis (17). To facilitate the assignment of entire GPCRs and to better understand their folding, and to optimize conditions for NMR measurements, we have decided to study large fragments of two different GPCRs, the human Y4 receptor, a class A GPCR targeted by the NPY family of neurohormones (18) and Ste2p, a class D GPCR from yeast targeted by the α-factor pheromone (19).

The two-stage model postulates that the secondary structure of helical integral membrane proteins spontaneously forms upon association of the protein with the membrane surface (20,21). Individual helices then insert into and diffuse within the membrane until the correct tertiary contacts are made and the 7-TM assembly is formed. Here, we address the following question: To what extent is the correct secondary or tertiary structure present in putative intermediates of GPCR folding? This question is prompted by the knowledge that the nascent chain of the GPCR is continuously released by the translocon from its N- to its C-terminus during membrane insertion. Accordingly, the segments comprising TM1 and TM2 may initiate the folding pathway. Correct insertion and complete folding of the TM1-TM2 domain may or may not depend on the formation of interhelical contacts and insertion into the membrane.

We have previously determined the structure of a segment comprising 19 residues and the first two transmembrane helices from the N-terminal domain of the Ste2p receptor (22), referred to here as TM1-TM2, in detergent micelles. Secondary structure was very well defined and long-range interhelical contacts were observed. Moreover, the location of the two TM helices in the primary structure perfectly agreed with predictions from a homology model based on the crystal structure of bovine rhodopsin (23). Previously, we have also reported on backbone assignments of a similar construct from the Y4 receptor (24). For an overview of the sequences of the investigated proteins see Fig. 1. Herein, we report on the assignment of the side-chain resonances of Y4_TM1-TM2, and present structural data on that protein. The structural properties and the membrane-insertion topology of Y4_TM1-TM2 are compared with those previously determined for TM1-TM2 from Ste2p (22) and related to thermodynamic data on amino acid partitioning into different membrane compartments. The implication of the data for TM helical membrane protein folding is discussed.

Submitted March 30, 2012, and accepted for publication July 12, 2012.

*Correspondence: oliver.zerbe@oci.uzh.ch

Chao Zou's present address is Novartis Pharma AG, Basel, Switzerland.

Editor: Francesca Marassi.

0006-3495/12/08/0817/10 $2.00

http://dx.doi.org/10.1016/j.bpj.2012.07.012

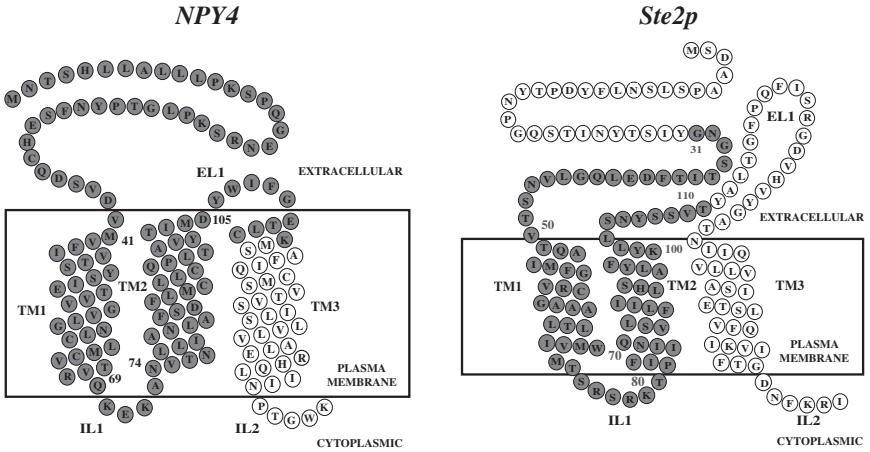

FIGURE 1 Snake plots (34) of the segments comprising the N-termini and the first three transmembrane helices of the Y4 (*left*) and Ste2p (*right*) receptors. Gray spheres highlight residues that are part of the proteins described in this work.

MATERIALS AND METHODS

Chemicals and solutions

$^{15}NH_4Cl$, ^{13}C-d_7-D-Glucose and D_2O were purchased from Spectra Stable Isotopes (Andover, MA). 1-palmitoyl-2- hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LPPG)/dodecyl-phosphocholine (DPC) were bought from Avanti Polar Lipids (Alabaster, AL). All other chemicals were obtained from Sigma-Aldrich (Buchs, Switzerland).

Biosynthesis of Y4_TM1-TM2

The cloning and expression of Y4_TM1-TM2 has been previously described by us (24). The gene encoding this fragment was derived from the cDNA of the human Y4 receptor (University of Missouri-Rolla, USA). The target protein was expressed in *Escherichia coli* BL21-AI cells and purified from inclusion bodies using a combination of Ni-affinity chromatography and C4-reverse-phase HPLC (24), resulting in comparably high expression levels with 6 mg of purified protein per liter of M9 culture in light water or 3 mg from heavy water. NMR samples were prepared as described (24). A deuterated d_{36}-LPPG and d_{38}-DPC detergents mixture were used for the side-chain assignment with ^{13}C-edited spectra. In our experience LPPG is clearly the detergent with better properties for multi-TM GPCR fragments usually resulting in spectra with better signal dispersion and more homogenous linewidths. However, DPC was required to solubilize TM1-TM2. Studies with the other 2-TM fragments, from the Y4 receptor that can be solubilized in LPPG/DPC mixtures and in pure LPPG, revealed that spectra and dynamical properties of the proteins were almost identical in the two systems (data not shown).

NMR spectroscopy

All samples were measured at 320 K on a Bruker AV700 spectrometer equipped with a triple-resonance cryoprobe. NMR samples contained ~0.3 mM protein in 40 mM phosphate buffer, pH = 6.0 with 40 mM DTT, 6% LPPG, and 1% DPC as described previously (24). Proton chemical shifts were referenced to the water line at 4.47 ppm at 320 K, from which the nitrogen and carbon scales were derived indirectly by using the conversion factors of 0.10132900 (^{15}N) and 0.25144954 (^{13}C).

Experiments used for backbone assignments are described in Zou et al. (24). Side-chain resonance assignment was accomplished using hCCH-TOCSY/COSY (25,26) in combination with [^{13}C,^{1}H]-HSQC and ^{13}C-resolved aliphatic/aromatic-NOESY experiments (27). Spectra for side-chain assignments required the use of d_{36}-LPPG and d_{38}-DPC to eliminate the strong residual signals from detergent that would otherwise obscure the region of C-α and methyl resonances. We noticed small changes in peak positions between spectra measured on deuterated and nondeuterated proteins as well as for deuterated and nondeuterated detergent. Therefore, Cα and Cβ chemical shifts obtained from the backbone assignments were initially adjusted by using hCCH-TOCSY or hCCH-COSY spectra. All chemical shifts were finally correlated to peak positions in the [^{15}N,^{1}H]- and [^{13}C,^{1}H]-HSQC spectra. Spectra were processed within the Bruker spectrometer software Topspin 2.1 and chemical shift assignments were performed using the software CARA (Keller 2004). Important parameters of the NMR experiments are summarized in the Supporting Material, Tables S1 and S2.

The extent of amide hydrogen exchange was estimated from the exchange crosspeak with water in the 70 ms ^{15}N-resolved NOESY spectrum. ^{15}N-relaxation data were recorded using proton-detected versions of the $^{15}N\{^{1}H\}$-steady-state NOE experiment (28), of the Carr-Purcell-Meiboom-Gill (R_2) and the inversion-recovery (R_1) experiments using a 0.4 mM ^{15}N-labeled sample under otherwise the same conditions of detergent, pH, and temperature. For R_1 and R_2 experiments recycle delays of 2.2 s and for the $^{15}N\{^{1}H\}$-NOE delays of 3.2 s were applied. 16 scans were accumulated for each increment in the R_1 or R_2 series and 64 scans for each increment of the $^{15}N\{^{1}H\}$-NOE. The R_1 series used the following relaxation delays: 0.010, 0.030, 0.050, 0.100, 0.200, 0.500, 1.000, and 2.000 s and the R_2 series the following delay settings: 0.017, 0.034, 0.068, 0.085, 0.119, 0.153, 0.187, 0.237, and 0.288 s. Peak volumes were integrated within the program SPSCAN that uses lineshape deconvolution of signals for proper integration of partially overlapping peaks and evaluated using home-written routines for least-squares fitting.

Studies of membrane insertion topology were performed using the micelle integrating spin label (2-(3-carboxypropyl)-4,4-dimethyl-2-tridecyl-3-oxazolidinyloxy (5-DSA) (Sigma). Samples for PRE measurements were prepared by addition of concentrated stock solutions of the paramagnetic substance to the micellar protein solution containing ^{15}N-labeled Y4_TM1_TM2. After addition the pH was readjusted to pH 6.0. The final concentrations of 5-DSA was 6 mM (29).

Structure calculation

Distance restraints were obtained from ^{15}N-resolved NOESY spectra recorded on ^{15}N,^{1}H- and ^{15}N,^{2}H-labeled Y4_TM1-TM2 samples with mixing times of 70 and 200 ms, respectively, and from 100 ms ^{13}C-resolved NOESY spectra. In general, ^{15}N- or ^{13}C-resolved NOESY spectra were recorded with nondeuterated and deuterated detergents, respectively. In contrast to the spectra recorded on Ste2p, we did not observe significant differences in linewidths in deuterated versus nondeuterated detergent mixtures. In addition, dihedral angle restraints obtained using the program TALOS+ (30), that uses chemical shifts of $^{1}H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$, and ^{15}N nuclei, were added. Peak lists for each spectrum were picked by the

program package UNIO'10 (31). The automatically picked peak lists were manually edited to remove artifactual peaks (e.g., t_1 noise) or to pick additional weak peaks. The integrated, nonassigned peak list from UNIO'10 was subsequently transferred to CYANA (32), which annotated the peak list in seven iterative cycles using the build-in macro noeassign. Again, the results from the automatic assignments were carefully checked and edited when necessary. The final CYANA calculation was performed with 100 randomized starting structures, and the 20 CYANA conformers with the lowest target function values were selected to represent the NMR ensemble.

Peak positions in the ^{13}C- and ^{15}N-resolved NOESY spectra differed for samples recorded in deuterated from those in nondeuterated detergent. To account for that deviation chemical shift positions were related to the corresponding NOESY spectrum both manually and automatically using the side-chain adaptation routine of UNIO'10.

RESULTS

Backbone and side-chain assignments

We reported on the backbone assignment of Y4_TM1-TM2, which is complete except for residues M1$^{N\text{-}ter}$, N2$^{N\text{-}ter}$, H5 $^{N\text{-}ter}$, and Q95$^{2.58}$, previously (BMRB 15921) (The numbering in the superscript follows the Ballesteros-Weinstein nomenclature (33) and is used in the GPCR-SSFE Database (34)). A representative [^{15}N,^{1}H]-TROSY spectrum is depicted in Fig. 2. Interresidue correlations were determined from triple-resonance spectra recorded on a ^{2}H,^{13}C,^{15}N-labeled sample, mostly relying on common Cα and Cβ resonances from HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH experiments (24). Side-chain assignments were made using hCCH-TOCSY and hCCH-COSY as well as ^{15}N- and ^{13}C-resolved NOESY spectra starting from the known Cα and Cβ chemical shifts. The assignment procedure for side-chain resonances is described in the Supporting Material (see also Fig. S1). At a late stage during the assignment the 2D [^{13}C, ^{1}H]-HSQC was scanned for unassigned spins to help locate thus far unassigned moieties. The aliphatic side-chain chemical shifts of all residues could be annotated except for the first residue M1$^{N\text{-}ter}$. Due to severe resonance overlap, assignments in the stretch comprising L9$^{N\text{-}ter}$, L10$^{N\text{-}ter}$, L11$^{N\text{-}ter}$, and W107^{EL1}, I108^{EL1}, F109^{EL1}, however, were ambiguous. Further side-chain resonances were annotated with the help of ^{15}N-(Gln, Asn) or ^{13}C-resolved (aromatic residues) NOESY spectra. In total 93.3% of the chemical shifts have been assigned and deposited in the BMRB database under accession code BMRB 18319.

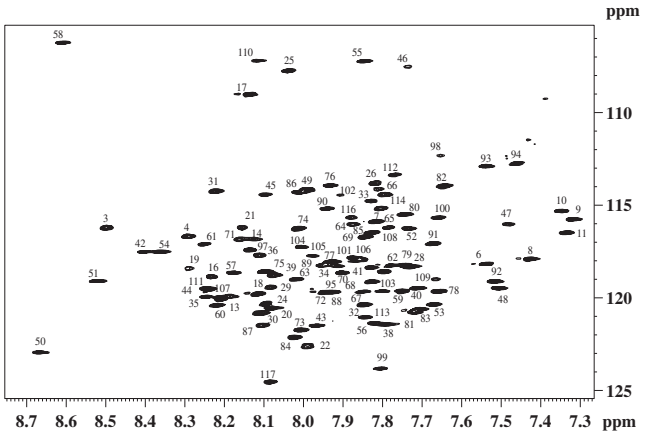

FIGURE 2 Assignments of the [^{15}N,^{1}H]-HSQC spectrum of Y4_TM1-TM2. Peaks from backbone moieties are numbered corresponding to the location of the corresponding residue in the Y4 receptor. The sample contained 0.5 mM protein, 1% (28 mM) DPC, 6% (118 mM) LPPG, 40 mM phosphate buffer, 40 mM DTT, pH = 6.0 at 47°C.

Structure calculation

The structure of Y4_TM1-TM2 was computed using the program UNIO'10 in combination with CYANA. In total 940 distance restraints were obtained from the ^{15}N- and ^{13}C-resolved NOESY spectra. Medium-range NOEs were almost exclusively observed in the putative TM-helical regions (see Fig. 3 *B*). Many of these NOEs, however, did not correspond to α,β (i,i+3) or α,HN contacts. The NOE-derived restraints were complemented by 100 torsion-angle restraints derived from backbone chemical shifts using the secondary structure prediction program TALOS+ (30). Helices were predicted to be located in the stretches S4 $^{N\text{-}ter}$-L9 $^{N\text{-}ter}$, F30 $^{N\text{-}ter}$-V45$^{1.36}$, I50$^{1.41}$-K70^{IL1}, A73 ^{IL1}-I80$^{2.43}$, F85$^{2.48}$-Q95$^{2.58}$, and L97$^{2.60}$-C114$^{3.25}$. TALOS+ predicts 74% of the 77 residues in the C-terminus of the fragment (the 2 TM helices plus the loops) to be helical, whereas the snake plot derived from a rhodopsin homology model contains 81.3% for the corresponding segment (see Fig. 1) (34). Interestingly, in both TM helices TALOS+ predicts the helices to be destabilized adjacent to the internal charged residues Glu-51$^{1.42}$ in TM1 and Asp-87$^{2.50}$ in TM2.

An overview of the restraints used for the structure calculation is depicted in Fig. 3. The difference of secondary chemical shifts, Δ(Cα-Cβ) (Fig. 3 *A*), allows the regions of secondary structure to be located. The chemical shift data indicate that most of Y4_TM1-TM2 contains secondary structure except for the N-terminal domain. However, as depicted in Fig. 3 *C*, not all helical parts manifested medium-range α,N(i,i+3) or α,β(i,i+3) contacts typically observed in stable helices. Surprisingly, such NOEs occurred only in the segments comprising residues N59$^{1.50}$-Q69$^{1.60}$ and N74$^{2.37}$-N82$^{2.45}$ in TM1 and TM2, respectively, and for the short helix α1 and α2 at the N-terminus. The segment comprising residues I50$^{1.41}$-V54$^{1.45}$ contains many other medium-range NOEs (Fig. 3 *B*) but helices in D87$^{2.50}$-Q95$^{2.58}$ and L97$^{2.60}$-K115$^{3.26}$ were mainly identified using the chemical shift restraints. The occurrence of medium-range NOE contacts, usually related to the helix stability, is not well-correlated to increased values for ΔCα-ΔCβ (Fig. 3 *A*). Although high values of the ^{15}N{^{1}H}-NOE or the ΔCα-ΔCβ chemical shift difference predict secondary structure to be most stable (e.g., for residues M41$^{1.32}$ to T46$^{1.37}$ or I50$^{1.41}$ to G55$^{1.46}$), due

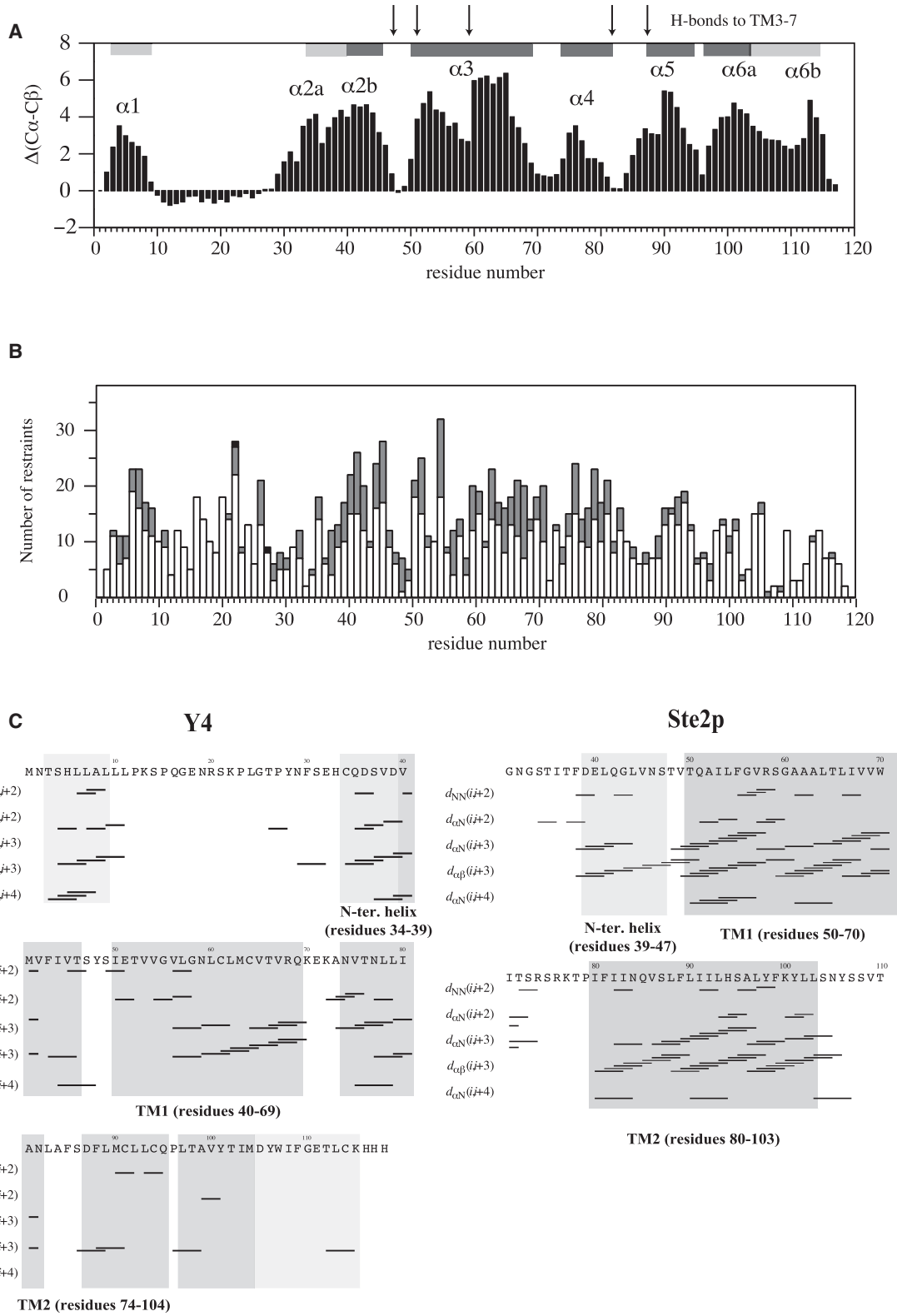

FIGURE 3 NMR parameters identifying secondary structures in Y4_TM1-TM2. (*A*) This panel displays the difference of the Cα and Cβ secondary chemical shifts. Arrows indicate the location of polar residues that form H-bonds with residues from TM3 to TM7, and dark and light gray bars indicate the location of helical regions based on the results of the structure calculations. Regions that are putative helices are labeled with an α and are numbered starting at the N-terminus. Helical regions are shaded in light gray for helices not embedded in the micelle or in dark gray for those of the putative TM1 or TM2 segments. (*B*) Number of NOE restraints for each residue: the white bars denote intraresidual and sequential restraints ($|i - j| \leq 1$), gray bars medium-range restraints ($1 < |i - j| < 5$), and black bars long-range restraints. (*C*). Comparison of NOEs identifying secondary structure in TM1-TM2 of the Y4 receptor (*left*) or the Ste2p (*right*).

to the increased linewidths for resonances for atoms in those segments NOE contacts may be impossible to identify. The increased linewidths result from the combined effects from the prolonged correlation time for moieties anchored in the micelle, resulting in broader and hence weaker peaks, and the exchange broadening due to rigid-body motions of the partially formed helices in the micelle.

The lowest energy structures displayed reasonably low values of the target function (see Table S3) that characterizes the extent of residual consistent violations of distance and dihedral angle restraints. The N-terminal tail contains two short helices denoted as α1 ($T3^{N\text{-}ter}$-$L9^{N\text{-}ter}$, root mean-square deviation (RMSD) 0.30 Å) and α2a ($C34^{N\text{-}ter}$-$D39^{N\text{-}ter}$, RMSD 0.39 Å) (see Fig. S2 *A*). The putative TM1 contains α2b ($V40^{N\text{-}ter}$-$S47^{1.38}$, RMSD 0.29 Å) and following a short turn formed by $S47^{1.38}$-$Y48^{1.39}$-$S49^{1.40}$, a hydrophobic helix α3 ($I50^{1.41}$-$Q69^{1.60}$, RMSD 1.25Å), which extends up to the putative location of the first intracellular loop (IC1) encompassing residues K70 ^{IL1} to A73 ^{IL1}. The helix α3 is slightly destabilized around the internal $G55^{1.46}$xx$G58^{1.49}$ motif. The putative TM2 is formed by the three helices α4 ($N74^{2.37}$-$N82^{2.45}$, RMSD 0.48Å) and α5 ($D87^{2.50}$-$Q95^{2.58}$, RMSD 0.53 Å), interrupted close to Ser-$86^{2.49}$/Asp-$87^{2.50}$, and α6a ($L97^{2.60}$-$M104^{2.67}$, RMSD 0.41Å). Helices α5 and α6a are joined via Pro-$96^{2.59}$. Finally, the hydrophobic helix α6a is directly connected to the amphiphilic helix α6b ($D105^{2.68}$-$K115^{3.26}$, 0.59 Å). A superposition of backbone atoms for all residues of the putative TM1 and TM2 results in RMSDs of 2.48 Å and 3.73 Å, respectively. The results from the structure calculation indicate that the additional motion introduced within TM2 by the polar residues $S86^{2.49}$ and $D87^{2.50}$ leads to a greater destabilization than the perturbation within TM1 by residues $G55^{1.46}$ and $G58^{1.49}$. Finally, we noticed that the amphiphilic helix α6b extends into the putative TM3 segment (Fig. 1, vide infra). Unfortunately, no unambiguous contact NOEs could be assigned between any residues of putative TM1 (α2a and α3) with those of putative TM2 (helix α4, α5, and α6a). Accordingly, no tertiary structure is present in the computed ensemble of conformers (see Fig. S2 *B*).

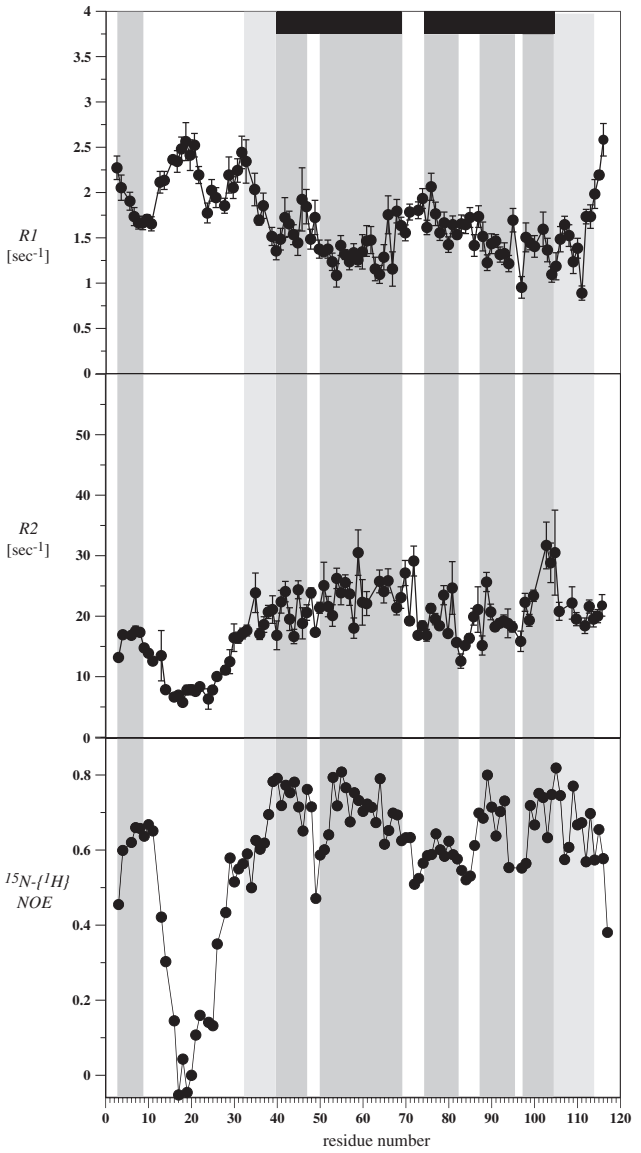

FIGURE 4 Dynamics of Y4_TM1-TM2: ^{15}N R1 (*top*) and R2 (*middle*) relaxation rates and ^{15}N{^{1}H}-NOE values (*bottom*). The measurements were made with a 0.5 mM sample of Y4_TM1-TM2 in 4:1 LPPG/DPC, 40 mM phosphate buffer, 40 mM DTT, pH = 6.0 at 47°C. Shading follows the rationale explained in Fig. 3. In addition, the putative locations of TM1 and TM2 (see Fig. 1) are indicated by black bars.

Backbone dynamics

The structure calculations of Y4_TM1-TM2 indicated that the TM helices are not fully formed, and that no stable inter-helical contacts are present. To confirm this result and ensure that it is not an artifact resulting from a lack of restraints we have probed for backbone dynamics by measuring ^{15}N relaxation at 600 MHz. The R1, R2, and ^{15}N{^{1}H}-NOE data are summarized in Fig. 4. The relaxation data all confirm that residues of the comparably long polypeptide N-tail comprising residues 10 to 32 possess increased flexibility. Interestingly, in comparison to TM2 residues of the putative TM1 display slightly lower R1 (1.43 ± 0.22 Hz for residues V40 $^{N\text{-}ter}$-$Q69^{1.60}$ vs. 1.50 ± 0.23 Hz for residues $N74^{2.37}$-$M104^{2.67}$) and slightly higher R2 rates (24.11 ± 7.3 Hz *vs.* 20.66 ± 6.0 Hz for the same residues) indicating the TM1 is slightly more rigid than TM2. This observation is supported by the larger number of medium-range NOEs for residues of TM1 compared to residues of TM2 (Fig. 3 *B*). Moreover, only three N-terminal and the last C-terminal residues are flexible indicating that both the N- as well as the C-terminus interact with the micelles or are part of stable secondary structure. Residues of the short loop centered at K70 ^{IL1} that connects TM1 to TM2 seem to have a modest increase in mobility compared to the residues in the contiguous helices. The lack of

significantly increased flexibility for these residues may either be an indication of transient interhelical contacts or may be because the two adjacent helices are both integrated into the micelle and therefore have restricted mobilities. We have also noticed increased R2 rates at the C-terminus of α6a that we attribute to conformational exchange effects.

In general, a good correlation is observed between the $\Delta(C\alpha\text{-}C\beta)$ values (Fig. 3 *A*) and the $^{15}N\{^{1}H\}$-NOE values (Fig. 4). The dynamics data for Y4_TM1-TM2 fully support the view that the protein is not composed of two stable TM helices that are well integrated into the micelles. Rather Y4_TM1-TM2 likely consists of a number of shorter TM helical stretches that are connected to each other via residues that either serve to break the helix or that prevent the helix from favorably integrating into the micelle interior (vide infra). The destabilization in both TM helices proximal to the internal polar residues E51$^{1.42}$ and T52$^{1.53}$ in TM1 or S86$^{2.49}$ and D87$^{2.50}$ in TM2 is clearly visible in the $^{15}N\{^{1}H\}$-NOE data, but less apparent in the R1 and R2 data.

Spin labels and solvent exchange

To probe for the topology of the protein in the detergent a micelle-integrating spin label, 5-DSA was used and attenuations in the [^{15}N,^{1}H]-HSQC spectra were observed (see Fig. 5). With 5-DSA, attenuation is strongest for amide moieties in close proximity to detergent headgroups. Attenuations of amide peaks due to 5-DSA are only consistently weak for residues of the long loop in the N-terminal domain. In contrast, attenuations for most residues in the putative transmembrane domain, with the exception of those proximal to residues V65$^{1.56}$ and I103$^{2.66}$, are >60%, suggesting that these are inside the micelle. Interestingly, attenuations of residues in the central region of the helices are smaller indicating that those segments are better integrated into the micelle and further away from the headgroups. Moreover, residues from the helix α4 and the adjacent residues (V75$^{2.38}$ to F85$^{2.48}$) are strongly attenuated, and we believe that this is because the helix α4 is mainly located in the interfacial region. The general noise observed in the attenuations and the lack of a clear trend in the paramagnetic relaxation may indicate that extensive motional averaging influenced the data.

Amide proton exchange with solvent presents an additional tool to probe for solvent protection that is either due to the presence of hydrogen bonds or to shielding from solvent (Fig. 5). A comparison of the data from Ste2p and the Y4 receptor is shown in the Supporting Material (Fig. S3). In both proteins, solvent exchange is reduced in the putative TM regions. However, although exchange within the TM regions is generally very slow for nearly all of these Ste2p residues, many residues in the TM segments of the Y4 receptor exhibit accelerated exchange, in particular in the regions connecting the short helices α2 and α3 as well as α4 and α5 and around the GxxG motif in α3. In

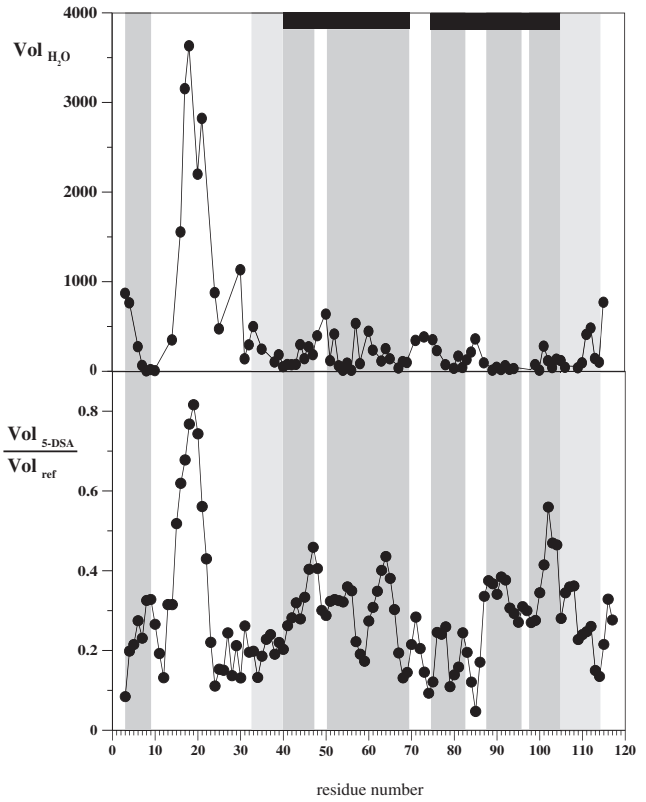

FIGURE 5 Amide proton deuterium exchange and spin label broadening of amides in Y4_TM1-TM2. Top: Integral of the NOESY exchange peak at the water frequency of Y4_TM1-TM2 (arbitrary scaling of *y* axis). Bottom: Attenuations of amide crosspeak volumes due to the presence of the spin label 5-DSA relative to crosspeak volumes Vol_{ref} in absence of the 5-DSA measured on the same sample. The putative locations of the TM helices are indicated by black bars.

both, the proton exchange is accelerated in the loop connecting the two TM segments, and fast exchange in Y4_TM1-TM2 additionally occurs for residues P12 $^{N\text{-}ter}$ to R20 $^{N\text{-}ter}$ of the long hydrophilic loop in the extracellular N-terminal domain. Overall, the solvent exchange data of Ste2p TM1-TM2 are fully compatible with the proposed helical turn structure (22). However, in the case of the Y4 receptor, the exchange data also indicate that the topology is different from the predictions depicted in Fig. 1. Clearly, solvent-exposed sites are visible within the Y4-TM helix regions, identifying instable helices that are interrupted at the positions of polar residues.

Probing for the importance of Asp-87$^{2.50}$

To probe for the influence of the potentially charged Asp-87$^{2.50}$ on micelle integration and formation of interhelical contacts, the mutant D87$^{2.50}$N-Y4_TM1-TM2 was produced in ^{15}N-labeled form. In general, the [^{15}N,^{1}H]-HSQC spectra were very similar to those of the wild-type protein. Large chemical shift differences are limited to the site of the mutation as well as to the C-terminal decapeptide

(see Fig. S4). Similarly, the attenuations from the micelle-integrating spin-label 5-doxylstearate display similar trends, and the $^{15}N\{^{1}H\}$-NOE indicates that the mutation failed to stabilize the TM segments significantly.

Comparison with TM1-TM2 of Ste2p

Comparison of the structures of Y4_TM1-TM2 and TM1_TM2 of Ste2p (PDB code: 2K9P) reveals a helix before TM1 in both proteins, which is located in the stretch $C34^{N\text{-}ter}$–$D39^{N\text{-}ter}$ in Y4_TM1-TM2. In Ste2p the helix is longer and spans D39-S47. In the Y4 receptor, we observe a significant change in the $^{15}N\{^{1}H\}$-NOE values from ~0.6 for residues ($F30^{N\text{-}ter}$-$Q35^{N\text{-}ter}$) to ~0.8 for residues ($S37^{N\text{-}ter}$-$T46^{1.37}$). Together with the fact that the segment comprising residues $C34^{N\text{-}ter}$ to $D39^{N\text{-}ter}$ is amphiphilic in nature, whereas residues from helix α2b are hydrophobic, we suggest that helix α2a and α2b form two different entities, and that α2b is actually part of TM1, even though we are not able to see a separation of the helices in the calculated structures.

The previous analysis leads us to conclude that TM1 in the Y4 receptor is longer by nine residues compared with TM1 of Ste2p. The extent to which the interruption between α2b and α3 is influenced by the presence of Glu-$51^{1.42}$ and Ser-$47^{1.38}$/$S49^{1.40}$ or the chosen detergent system LPPG/DPC is unclear presently. We note that the NMR structures of both sensory rhodopsin (15) and proteorhodopsin (16) displayed the correct helix topology indicating that the cause of the instability in Y4_TM1-TM2 is probably due to the uncompensated polar/charged residues in this short fragment. It was reported previously that the presence of a GxxxG motif in Ste2p TM1 results in a destabilization of the helix in the center of TM1 (35). A similar GxxG motif is observed in TM1 of the Y4, which likely is another destabilizing factor for this helical domain. Furthermore, although several polar residues such as Q85, S87, S95, K100 are present within TM2 of Ste2p, a complete set of medium-range NOEs is observed in that segment (see Fig. 3 *C*). Finally, TM2 in Ste2p is shifted by six residues compared with TM2 of Y4 (from $N74^{2.37}$-$Q95^{2.58}$ to $I80^{2.43}$-$I103^{2.66}$) because the connecting loop is much longer. Indeed, a prominent difference between the two proteins is the length of the IL1 loop (three residues in NY4 and six residues in Ste2p) that may play a crucial role in the folding of two transmembrane helices in a micellar system. The loop conformation was found to exert a large influence on interhelical contacts in a CFTR TM helical hairpin (36), especially when the hairpin is marginally stable. As discussed below, however, we believe that other factors, in particular the occurrence of uncompensated charged or polar residues, have an important (likely larger) influence.

DISCUSSION

In this work, we compare the structures of fragments corresponding to the first two transmembrane helices of yeast and a human GPCR and investigate how the very different content of polar/charged residues influences the biophysics of these polypeptides. The structure of the TM1-TM2 fragment from the yeast Ste2p receptor was determined in LPPG micelles by us previously (22), and manifested a tertiary structure similar to that predicted from a rhodopsin homology model (23). Herein, we report data for a similar fragment from the human Y4 receptor. The NMR spectrum of Y4_TM1-TM2 displayed reasonable linewidths and signal dispersion enabling us to achieve a rather high degree of resonance assignments for this 115-residue membrane protein. Our computed NMR structures for Y4_TM1-TM2 showed a fair agreement with the location of the secondary structure as predicted from the homology model (Fig. 1). Nevertheless, we observed interruptions in the helical secondary structures in both TM1 and TM2.

The experimental evidence indicates the secondary structure of Y4_TM1-TM2 is less stable compared to that of the analogous Ste2p fragment, and that the packing of the two TM helices, if at all present, is only transient in the Y4 fragment. Specifically, much fewer medium-range NOEs occur in Y4_TM1-TM2 when compared to TM1-TM2 of Ste2p. Because of the absence of these medium range NOEs the helical structure of some regions of Y4_TM1-TM2 such as $D87^{2.50}$-$Q95^{2.58}$ in TM2 is mainly based on chemical-shift derived restraints. Moreover, significant amide proton exchange occurs in the putative TM region of Y4_TM1-TM2, e.g., between α2b and α3 as well as between α4 and α5 (see Fig. 5). This was not observed for similar regions of the Ste2p fragment (Fig. 3 *C* and Fig. S3). Spin-label data experiments probing the topology of micelle integration reveal fewer pronounced differences than expected for a unique topology. We attribute this unexpected result to extensive conformational exchange processes. Finally, backbone dynamics (Fig. 4) are indicative of positions of increased mobility within the TM portions of the Y4 fragment. We have observed a similar behavior in our studies of the isolated TM-7 of Ste2p (37), where positions of increased mobility were also correlated to the presence of polar residues. Although these polar residues have been proposed to mediate contacts between TM helices (38–41) and therefore contribute to the assembly of the 7-TM bundle, if they cannot efficiently locate their interaction partners in the smaller receptor fragment they will be exposed to the lipid environment. Consequently, both the secondary and the tertiary structure is perturbed, and the TM segments may not be stably anchored in the micelle and undergo additional rigid-body fluctuations that will partially transfer those residues into the interface. Johnsen et al. (42) have systematically probed the formation of interhelical contacts in a model helical hairpin membrane peptide, and observed a similar destabilization from polar residues. The fact that we failed to detect significant stabilization in the $D87^{2.50}N$ Y4_TM1-TM2 mutant, however, indicates that exchange of $Asp\text{-}87^{2.50}$ to $Asn\text{-}87^{2.50}$ may

reflect the protonation state of the wild-type Asp residue in the nonpolar lipid core, the unfavorable energetic consequences of a nonhydrogen-bonded Asn side chain, or the influence of other polar residues acting in concert with Asp-87$^{2.50}$.

Fig. 6 highlights the location of polar residues within the 7-TM bundle of Ste2p and the Y4 receptors, as derived from homology models based on the crystal structure of bovine rhodopsin. It is obvious that the Y4 receptor contains more polar side chains in the membrane interior than those of Ste2p. The top view of the Y4 receptor additionally reveals that most polar side chains are contained in helices TM1 and TM2, with TM3 making a number of additional polar contacts. Most contacts in the membrane interior are located on one continuous face of Y4_TM1-TM2, and together with residues from TM7 result in an amphiphatic interior of the helical bundle. Particularly remarkable is the high number of polar residues within TM1. In Ste2p polar residues are much more uniformly distributed across the TM helices and also seem to be located in a belt in the center of the TM bundle. The side-chain interactions of polar residues of TM1 and TM2 as derived from a model of the Y4 receptor are summarized in Table S4, and the location of those that form unsatisfied H-bonds to residues from TM3-7 based on the Y4 model are indicated in Fig. 3 *A*. Obviously, all these residues are in the vicinity of interrupts in the secondary structure.

The impact of uncompensated polar or charged sites can be estimated using the Wimley and White scale of free energies for amino acid transfer into the membrane-water interface or the membrane interior (43). A comparison of these values for the sequences of TM1-TM2 of the Y4 and the Ste2p receptors is shown in the Supporting Material (Fig. S5). The putative TM segments are recognized by the presence of a group of residues with near zero or negative ΔG_{trans} values compared to those from the loop regions. In the case of the Y4 receptor, polar and in particular charged residues, which result in high energetic penalties when placed in a membrane interior, are observed both in TM1 (residue E51$^{1.42}$) and in TM2 (residue D87$^{2.50}$). Close to these two residues we observe increased flexibility of the polypeptide backbone as demonstrated by lowered values of the $^{15}N\{^{1}H\}$-NOE, the absence of typical interhelical i,i+3 NOE contacts, or prediction of nonhelical torsion angles based on backbone chemical shifts (Figs. 3 and 4). Although residues with positive ΔG_{oct} values are also present in TM1 or TM2 of Ste2p, the energy penalties are much smaller, in particular on the ΔG_{wif} scale. Table 1 lists a comparison of these free energies summed over all residues of the TM helices in the Y4 and Ste2p receptors as computed using the MPEX program (44). The data indicate that the TM4-TM6 segment of NY4 may constitute a hydrophobic core that readily partitions into the membrane. In contrast the TM1 domain of the Y4 is quite polar and shows little tendency to partition into the membrane, whereas the TM1 helix of Ste2p displays a very favorable free energy for membrane insertion.

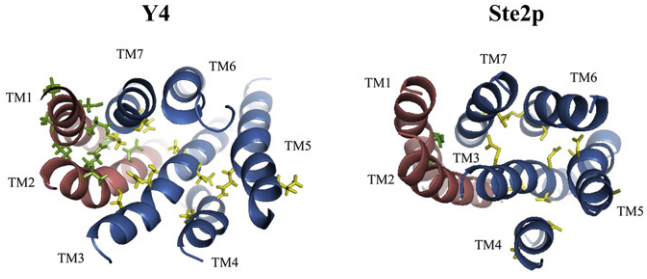

FIGURE 6 Structure comparison of models of the Y4 (*left*) (51) and Ste2p (*right*) (23) receptors. Only the backbone trace of the TM portions is shown. TM helices 1 and 2 are depicted in red. Charged or polar side-chains (R,K,E,D,N,Q,T,S) are displayed for residues from TM1 or TM2 in green and for those from the remainder of the receptor in yellow. Among these side chains, those not further away than a full helical turn from termini of the helices are removed for clarity.

A multiple sequence alignment for all Y4-TM1-TM2 sequences included in the GPCRDB database (Table S5) (45) shows that in general sequence conservation among the Y4 receptor is high. With the exception of Asn-59$^{1.50}$, Asn-82$^{2.45}$, and Asp-87$^{2.50}$ conservation of polar residues within the Y receptor family, however, is below 50%, e.g., for Glu-51$^{1.42}$ the consensus residue is actually Phe. Among class A (rhodopsin-like) GPCRs, only Asn-59$^{1.50}$ and Asp-87$^{2.50}$ are retained. A recent systematic study, using ligand binding as the selection assay, in which all codons in the neurotensin receptor were replaced with the remaining 63 possibilities, revealed that polar residues frequently represented so-called drift positions (46), implying that their replacement results in either more stable or better binding proteins. This surprising result indicates that these polar residues are primarily important for signaling rather than folding or ligand binding.

We believe that fragments of GPCRs containing transmembrane domains may be useful because they allow investigation of the partitioning and folding events that occur in the two-stage model of Popot and Engelman (20,21) (see also Jacobs and White (47) and the Introduction). Recent work has shown that TM segments do not diffuse directly from the ribosome exit tunnel into the membrane-water interface but rather are injected into the translocon. Depending on the hydrophobicity they are then allowed to pass across the membrane or are laterally gated to partition into the membrane (48). Considering the fact that proteins

TABLE 1 Free energies (in kcal/mol) for partitioning of individual helices into the membrane calculated using the MPEX program

	TM1	TM2	TM3	TM4	TM5	TM6	TM7
Y4	−1.99	−6.83	−8.58	−13.27	−13.75	−13.15	−8.36
Ste2p	−12.63	−4.88	−3.63	−8.6	−8.43	−12.42	−7.91

are synthesized starting at the N-terminus Y4_TM1 is expected to be expelled from the translocon first.

The work on Y4_TM1-TM2 and other GPCR fragments leads us to speculate that individual helices may not initially be fully formed and, in fact, they may be disrupted at polar residues, proline, or GXXXG motifs. In addition, some of the TM helices such as TM1 of the Y4 receptor possess little tendency to remain fully integrated in the membrane, but instead may prefer to become surface-associated or only partially integrated. For the Y4 receptor the energies for inserting individual TM helices into a membrane differ vastly, and only TM helices 4, 5, and 6 are expected to be stably integrated on their own. Therefore, when gated from the translocon into the membrane interior TM1 or TM2 likely will not remain inserted as full helices until all the required compensating interactions with residues of other TM segments are made. The Skach group has used photoaffinity cross-linking experiments to probe how individual TM helices of aquaporin-4 during biogenesis of the entire protein form contacts to Sec-61α, which is part of the mammalian translocon (49). Although this ion channel is very different from a GPCR the experiments demonstrated that TM1 becomes reassociated with Sec-61α after synthesis of TM3, indicating that some regions of membrane proteins may remain attached to the translocon until larger parts or the entire sequence have been synthesized. We recently have conducted an extensive study on the topogenesis of truncated fragments from the Y4 receptor in a native environment. These experiments have shown that the percentage of correctly inserted TM segments is strongly increased with increasing number of TM helices (50). It would be consistent to conclude that the translocon facilitates assembly of the TM bundle by stabilizing the N-terminal TM segments in a hydrophobic environment until further segments that provide the necessary compensating interactions have left the translocon. In other cases, such as TM1-TM2, the N-terminal hairpin is relatively stable and highly hydrophobic and can integrate into the bilayer to provide a starting anchor for receptor folding. Further experiments are certainly required to distinguish these cases and to follow the folding events. These experiments would be aided by the availability of longer GPCR fragments with three and more TMs that are designed to provide additional polar contacts in the bilayer. In any case, our data indicate that the view that individual TM helices insert into the membrane and diffuse as such to find their interacting counterparts may be too simplistic, at least in the case of receptors like Y4.

SUPPORTING MATERIAL

Five tables and five figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00786-2.

We thank the Swiss National Science Foundation (grant No. 31003A_124469) and the OPO Foundation for funding. Fred Naider acknowledges support from National Institute of General Medical Sciences grant 22087 and the Leonard and Esther Kurtz Term Professorship.

REFERENCES

1. Palczewski, K., T. Kumasaka, …, M. Miyano. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science.* 289:739–745.
2. Warne, T., M. J. Serrano-Vega, …, G. F. Schertler. 2008. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature.* 454:486–491.
3. Rasmussen, S. G., H. J. Choi, …, B. K. Kobilka. 2007. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature.* 450:383–387.
4. Jaakola, V. P., M. T. Griffith, …, R. C. Stevens. 2008. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science.* 322:1211–1217.
5. Chien, E. Y., W. Liu, …, R. C. Stevens. 2010. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science.* 330:1091–1095.
6. Wu, B., E. Y. Chien, …, R. C. Stevens. 2010. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science.* 330:1066–1071.
7. Shimamura, T., M. Shiroishi, …, S. Iwata. 2011. Structure of the human histamine H1 receptor complex with doxepin. *Nature.* 475:65–70.
8. Lebon, G., T. Warne, …, C. G. Tate. 2011. Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature.* 474:521–525.
9. Rosenbaum, D. M., C. Zhang, …, B. K. Kobilka. 2011. Structure and function of an irreversible agonist-β(2) adrenoceptor complex. *Nature.* 469:236–240.
10. Warne, T., R. Moukhametzianov, …, C. G. Tate. 2011. The structural basis for agonist and partial agonist action on a β(1)-adrenergic receptor. *Nature.* 469:241–244.
11. Xu, F., H. Wu, …, R. C. Stevens. 2011. Structure of an agonist-bound human A_{2A} adenosine receptor. *Science.* 332:322–327.
12. Standfuss, J., P. C. Edwards, …, G. F. Schertler. 2011. The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature.* 471:656–660.
13. Rasmussen, S. G., B. T. DeVree, …, B. K. Kobilka. 2011. Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature.* 477:549–555.
14. Gautier, A., J. P. Kirkpatrick, and D. Nietlispach. 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.
15. Gautier, A., H. R. Mott, …, D. Nietlispach. 2010. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat. Struct. Mol. Biol.* 17:768–774.
16. Reckel, S., D. Gottstein, …, V. Dötsch. 2011. Solution NMR structure of proteorhodopsin. *Angew. Chem. Int. Ed. Engl.* 50:11942–11946.
17. Zerbe, O. 2012. First solution structures of seven-transmembrane helical proteins. *Angew. Chem. Int. Ed. Engl.* 51:860–861.
18. Berglund, M. M., P. A. Hipskind, and D. R. Gehlert. 2003. Recent developments in our understanding of the physiological role of PP-fold peptide receptor subtypes. *Exp. Biol. Med. (Maywood).* 228: 217–244.
19. Naider, F., and J. M. Becker. 2004. The alpha-factor mating pheromone of *Saccharomyces cerevisiae*: a model for studying the interaction of peptide hormones and G protein-coupled receptors. *Peptides.* 25:1441–1463.
20. Popot, J. L., and D. M. Engelman. 1990. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry.* 29:4031–4037.
21. Popot, J. L., and D. M. Engelman. 2000. Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem.* 69:881–922.

22. Neumoin, A., L. S. Cohen, …, F. Naider. 2009. Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles. *Biophys. J.* 96:3187–3196.

23. Eilers, M., V. Hornak, …, J. B. Konopka. 2005. Comparison of class A and D G protein-coupled receptors: common features in structure and activation. *Biochemistry.* 44:8959–8975.

24. Zou, C., F. Naider, and O. Zerbe. 2008. Biosynthesis and NMR-studies of a double transmembrane domain from the Y4 receptor, a human GPCR. *J. Biomol. NMR.* 42:257–269.

25. Kay, L., G. Y. Xu, …, J. D. Forman-Kay. 1993. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson. B.* 101:333–337.

26. Olejniczak, E. T., R. X. Xu, and S. W. Fesik. 1992. A 4D HCCH-TOCSY experiment for assigning the side chain ^{1}H and ^{13}C resonances of proteins. *J. Biomol. NMR.* 2:655–659.

27. Sattler, M., J. Schleucher, and C. Griesinger. 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* 34:93–158.

28. Noggle, J. H., and R. E. Schirmer. 1971. The Nuclear Overhauser Effect - Chemical Applications. Academic Press, New York.

29. Hilty, C., G. Wider, …, K. Wüthrich. 2004. Membrane protein-lipid interactions in mixed micelles studied by NMR spectroscopy with the use of paramagnetic reagents. *ChemBioChem.* 5:467–473.

30. Shen, Y., F. Delaglio, …, A. Bax. 2009. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR.* 44:213–223.

31. Volk, J., T. Herrmann, and K. Wüthrich. 2008. Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *J. Biomol. NMR.* 41:127–138.

32. Güntert, P. 2004. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278:353–378.

33. Ballesteros, J. A., and H. Weinstein. 1995. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* 25:366–428.

34. Worth, C. L., G. Kleinau, and G. Krause. 2009. Comparative sequence and structural analyses of G-protein-coupled receptor crystal structures and implications for molecular models. *PLoS ONE.* 4:e7011.

35. Cohen, L. S., B. Arshava, …, F. Naider. 2011. Comparative NMR analysis of an 80-residue G protein-coupled receptor fragment in two membrane mimetic environments. *Biochim. Biophys. Acta.* 1808: 2674–2684.

36. Wehbi, H., A. Rath, …, C. M. Deber. 2007. Role of the extracellular loop in the folding of a CFTR transmembrane helical hairpin. *Biochemistry.* 46:7099–7106.

37. Neumoin, A., B. Arshava, …, F. Naider. 2007. NMR studies in dodecylphosphocholine of a fragment containing the seventh transmembrane helix of a G-protein-coupled receptor from *Saccharomyces cerevisiae*. *Biophys. J.* 93:467–482.

38. Adamian, L., and J. Liang. 2001. Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J. Mol. Biol.* 311:891–907.

39. Smith, S. O., M. Eilers, …, S. Aimoto. 2002. Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer. *Biophys. J.* 82:2476–2486.

40. Gratkowski, H., J. D. Lear, and W. F. DeGrado. 2001. Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. USA.* 98:880–885.

41. Zhou, F. X., H. J. Merianos, …, D. M. Engelman. 2001. Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. USA.* 98:2250–2255.

42. Johnson, R. M., C. L. Heslop, and C. M. Deber. 2004. Hydrophobic helical hairpins: design and packing interactions in membrane environments. *Biochemistry.* 43:14361–14369.

43. Wimley, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* 3:842–848.

44. Snider, C., S. Jayasinghe, …, S. H. White. 2009. MPEx: a tool for exploring membrane proteins. *Protein Sci.* 18:2624–2628.

45. Vroling, B., M. Sanders, …, G. Vriend. 2011. GPCRDB: information system for G protein-coupled receptors. *Nucleic Acids Res.* 39(Database issue):D309–D319.

46. Schlinkmann, K. M., A. Honegger, …, A. Plückthun. 2012. Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc. Natl. Acad. Sci. USA.* 109:9810–9815.

47. Jacobs, R. E., and S. H. White. 1989. The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices. *Biochemistry.* 28:3421–3437.

48. White, S. H., and G. von Heijne. 2008. How translocons select transmembrane helices. *Annu Rev Biophys.* 37:23–42.

49. Sadlish, H., D. Pitonzo, …, W. R. Skach. 2005. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat. Struct. Mol. Biol.* 12:870–878.

50. Marino, J., E. R. Geertsma, and O. Zerbe. 2012. Systematic studies of topology of fragments from the Y4 receptor, a human GPCR. *BBA Biomembranes*. In press.

51. Bissantz, C., A. Logean, and D. Rognan. 2004. High-throughput modeling of human G-protein coupled receptors: amino acid sequence alignment, three-dimensional model building, and receptor library screening. *J. Chem. Inf. Comput. Sci.* 44:1162–1176.