Biochimica et Biophysica Acta 1828 (2013) 1471–1483

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Structural features of the apelin receptor N-terminal tail and first transmembrane segment implicated in ligand binding and receptor trafficking

David N. Langelaan [a], Tyler Reddy [a,b], Aaron W. Banks [a], Graham Dellaire [a,e], Denis J. Dupré [c], Jan K. Rainey [a,d,*]

[a] Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2
[b] Structural Bioinformatics & Computational Biochemistry Unit, Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK
[c] Department of Pharmacology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2
[d] Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2
[e] Department of Pathology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2

ARTICLE INFO

Article history:
Received 17 October 2012
Received in revised form 17 January 2013
Accepted 13 February 2013
Available online 22 February 2013

Keywords:
Apelin receptor
Membrane protein structure
Divide and conquer
Biomolecular NMR spectroscopy
Homology model
Molecular dynamics simulations

ABSTRACT

G-protein coupled receptors (GPCRs) comprise a large family of membrane proteins with rich functional diversity. Signaling through the apelin receptor (AR or APJ) influences the cardiovascular system, central nervous system and glucose regulation. Pathophysiological involvement of apelin has been shown in atherosclerosis, cancer, human immunodeficiency virus-1 (HIV-1) infection and obesity. Here, we present the high-resolution nuclear magnetic resonance (NMR) spectroscopy-based structure of the N-terminus and first transmembrane (TM) segment of AR (residues 1–55, AR55) in dodecylphosphocholine micelles. AR55 consists of two disrupted helices, spanning residues D14-K25 and A29-R55$^{1.59}$. Molecular dynamics (MD) simulations of AR built from a hybrid of experimental NMR and homology model-based restraints allowed validation of the AR55 structure in the context of the full-length receptor in a hydrated bilayer. AR55 structural features were functionally probed using mutagenesis in full-length AR through monitoring of apelin-induced extracellular signal-regulated kinase (ERK) phosphorylation in transiently transfected human embryonic kidney (HEK) 293A cells. Residues E20 and D23 form an extracellular anionic face and interact with lipid headgroups during MD simulations in the absence of ligand, producing an ideal binding site for a cationic apelin ligand proximal to the membrane–water interface, lending credence to membrane-catalyzed apelin-AR binding. In the TM region of AR55, N46$^{1.50}$ is central to a disruption in helical character. G42$^{1.46}$, G45$^{1.49}$ and N46$^{1.50}$, which are all involved in the TM helical disruption, are essential for proper trafficking of AR. In summary, we introduce a new correlative NMR spectroscopy and computational biochemistry methodology and demonstrate its utility in providing some of the first high-resolution structural information for a peptide-activated GPCR TM domain.

1. Introduction

G-protein coupled receptors (GPCRs) are a large and diverse family of membrane embedded proteins that are responsible for converting the extracellular signal of a ligand into an intracellular signal. The importance of GPCRs is mirrored by their abundance, with >750 different receptors identified in humans [1]. Beyond their natural physiological importance, GPCRs are targeted by ~25% of current drugs [2] and have great potential for pharmaceutical development.

Due to the physiological and pharmacological importance of GPCRs, there has been considerable drive to characterize them structurally. In recent years, structures of 11 GPCRs have been determined at high resolution providing unprecedented insight into GPCR function [3–13]. However, it is important to note that in many cases the extracellular loops of these structures are poorly resolved or have been mutated in order to facilitate crystallization. Through these studies, it is clear that a

Abbreviations: AR, apelin receptor; AR55, residues 1–55 of AR; BSA, bovine serum albumin; CSI, chemical shift index; DMEM, Dulbecco's modified eagle medium; DPC, dodecylphosphocholine; DPPC, dipalmitoylphosphatidylcholine; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; DTT, dithitothreitol; ERK, extracellular signal-regulated kinase; FBS, fetal bovine serum; GPCR, G-protein coupled receptor; HA, hemagglutinin; HEK, human embryonic kidney; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, Luria broth; MD, molecular dynamics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; PCR, polymerase chain reaction; PEI, polyethylenimine; PDB, Protein Data Bank; RMSD, root mean square deviation; TFA, trifluoracetic acid; TM, transmembrane; TOCSY, total correlation spectroscopy

* Corresponding author at: Department of Biochemistry & Molecular Biology, Dalhousie University, 5850 College St., PO Box 15000, Halifax, NS Canada B3H 4R2. Tel.: +1 902 494 4632; fax: +1 902 494 1355.
E-mail address: jan.rainey@dal.ca (J.K. Rainey).

0005-2736
http://dx.doi.org/10.1016/j.bbamem.2013.02.005

network of polar residues residing in the membrane-spanning domain of GPCRs is responsible for signal transduction by allowing conformational change upon ligand binding. In particular, transmembrane (TM) segment VI is shown to undergo considerable movement during activation [14]. However, despite the similarities between structures there are also differences between them. For example, CXCR4 and rhodopsin contain a kink in the first TM helix while the other solved GPCRs do not.

The apelin receptor (AR or APJ) is a class A, rhodopsin-like GPCR identified in 1993 due to high sequence similarity to the angiotensin receptor AT_1 [15]. Angiotensin-II did not activate AR and it remained an orphan GPCR until 1999 when the endogenous peptide ligand was discovered and named apelin [16]. Various apelin isoforms from 13 to 36 residues in length are produced as C-terminal cleavage products in an as of yet unspecified manner from the preproapelin gene product [17].

Since the discovery of AR, it has been shown to be widely expressed in humans, with high levels in the central nervous system, cardiovascular system, adipocytes and many other tissues [2]. AR signaling has key roles in the cardiovascular system, with effects on vascular smooth muscle cells and a strong inotropic effect on cardiac muscle [17]. Plasma levels of apelin are correlated to heart failure progression [17] and body mass index [18]. Furthermore, AR is a known CD4 co-receptor for HIV type 1 [19].

Modulation of apelin signaling has therefore been identified as a therapeutic target, with a handful of non-natural ligands now reported. Recently, both agonist and antagonist peptide mimics [20–22] have been discovered based on our previous structural studies of apelin [23,24]. Blind screens have also been successful, with a non-peptidic agonist being discovered through a FRET based assay [25]. However, there are no published experimental structural data for AR itself, hampering the drug discovery process. Furthermore, there are no published structures of peptide binding GPCRs, making it unclear as to how many conclusions can be drawn from homologous structures.

At the molecular level, there have been two mutagenesis studies probing binding to and activation of AR by apelin. However, the coverage of the receptor is too sparse to make many conclusions about the ligand binding site or mechanism of activation. Zhou et al. [26] showed insensitivity to deletion of the first 10 residues of AR vs. a complete disruption of apelin binding competency upon deletion of the first 20 residues. Specifically, the E20A and D23A mutations prevented apelin binding despite normal receptor localization. Iturrioz et al. [27] demonstrated that mutation of $W152^{4.50}$ in rat AR TM IV affected folding while $F255^{6.44}$ and $W259^{6.48}$ mutants still allowed apelin binding but prevented AR internalization (Ballesteros nomenclature shown as superscript for TM residues [28]).

Full-length membrane proteins are challenging to characterize structurally [29]. The "divide and conquer" approach provides a tractable route to collect structural information about a membrane protein of interest. In this approach, dissected segments of membrane proteins are studied in a membrane mimetic environment. The underlying assumption, following from the multi-step cycle of membrane protein folding [30–33], is one of structural integrity with the same segment in the context of a full-length protein [34]. Although protein tertiary structure must be determined through other methods, the divide and conquer approach applied judiciously can provide reliable data about secondary structure and has been applied to several other GPCRs [35–43].

In this paper, we provide the NMR spectroscopy-derived high-resolution structure of the N-terminal tail and first TM segment of AR solubilized in dodecylphosphocholine (DPC) micelles. Residues involved in a TM domain helical disruption were subjected to mutagenesis, highlighting their involvement in functional AR folding. The NMR-based structure of AR55 was placed in the context of full-length AR using structural calculations and molecular dynamics simulations combining the experimental restraints and a CXCR4-based homology model of full-length AR. Considered in the perspective of mutagenesis data presented here and previously, key structural features for both AR folding and apelin binding are brought to light.

2. Materials and methods

2.1. Materials

pEXP5-CT vector, fetal bovine serum (FBS) and reagents for polymerase chain reaction (PCR) amplification of DNA were obtained from Invitrogen (Burlington, ON). A synthetic gene with bacterial codon optimized AR55 coding sequence as well as DNA primers for PCR amplification and mutagenesis were purchased from BioBasic Inc. (Markham, ON). Phusion and QuikChange mutagenesis kits were obtained from New England Biolabs (Pickering, ON) and Stratagene (Wilmington, DE), respectively. The cDNA clone for 3×-human influenza hemagglutinin (HA)-tagged human AR was obtained from the Missouri S&T cDNA Resource Center (www.cdna.org). cOmplete Protease Inhibitor Cocktail tablets were obtained from Roche (Laval, QC). HPLC grade acetonitrile, ampicillin, bovine serum albumin (BSA), dithiothreitol (DTT), isopropyl β-D-1-thiogalactopyranoside (IPTG), trifluoroacetic acid (TFA) and reagents for making Luria Broth (LB) medium were purchased from Fisher Scientific (Ottawa, ON). Uniformly ^{13}C enriched D-glucose and ^{15}N enriched NH_4Cl were purchased from Cambridge Isotope Laboratories (Andover, MA). Deuterated DPC (DPC-d_{38}), DTT-d_6, deuterium oxide (D_2O; 99.8 atom % D) and D_2O containing 1% (w/w) sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) were obtained from C/D/N isotopes (Pointe-Claire, QC). 0.2 μm filters were obtained from Millipore (Billerica, MA). Primary anti-HA antibodies were obtained from either Covance (Montréal, QC) or Santa Cruz Biotechnology Inc. (Santa Cruz, CA). Non-cross reactive anti-rabbit Dylight 549 conjugated and anti-mouse Dylight 649 conjugated antibodies were purchased from Rockland (Gilbertsville, PA). Donkey anti-mouse, Cy5 conjugated antibodies were obtained from Jackson ImmunoResearch (West Grove, PA). Paraformaldehyde was obtained from Electron Microscopy Sciences (Hatfield, PA) and fluorescence mounting medium was obtained from Dako (Burlington, ON). All other chemicals were obtained from Sigma-Aldrich (Oakville, ON).

2.2. Cloning and expression of AR55

PCR was used to amplify the purchased AR55 gene and clone it into pEXP5-CT vector. Sequencing (Biobasic Inc.) was used to check for correct vector insertion of the PCR product to code for the following amino acids: MEEGGDFDNY YGADNQSECE YTDWKSSGAL IPAIYMLVFL LGTTGNGLVL WTVFRKKGHH HHHH, corresponding to the first 55 residues of AR as well as a C-terminal KKG extension and His_6-tag. High yield expression of uniformly $^{13}C/^{15}N$-labeled AR55 (~5 mg/L) was achieved with BL-21(DE3) *Escherichia coli* grown in LB medium to an optical density at 600 nm (OD_{600}) of 0.6 and transferred to a 1/2-equivalent volume of minimal medium (100 mM NaH_2PO_4, 40 mM K_2HPO_4, 4 mM $MgSO_4$, 1.8 μM $FeSO_4$, 2 g/L $^{13}C_6$ D-glucose, 1 g/L of $^{15}NH_4Cl$ and 100 mg/L ampicillin, titrated with NaOH to pH 7.3), as in Markley et al. [44]. Expression was induced by adding IPTG to 0.5 mM followed by shaking (37 °C, 4 h) and centrifugation of the cells (3000 ×*g*, 15 min).

2.3. Purification of AR55

Cell pellets were resuspended (50 mM Tris–HCl buffer, pH 8.0, 25% w/v sucrose, 1 mM EDTA, 0.1% w/v NaN_3, 5 mM MgCl, 0.03 mg/mL DNase I and 1% Triton-X 100), sonicated and centrifuged (15,000 ×*g*, 30 min). The pellet was washed 4× (50 mM Tris–HCl buffer, pH 8.0, 0.5% Triton-X100, 100 mM NaCl, 1 mM EDTA). The resulting inclusion body-containing [45] pellet was resuspended in acetonitrile:H_2O (1:1 v:v) containing 0.1% TFA, filtered using a 0.2 μm filter and reduced with DTT at ~100 mM prior to reverse phase HPLC (Cosmosil C_{18} column, linear gradient from 45% to 65% acetonitrile in H_2O, 0.1% TFA over 15 min using a Beckman System Gold (Mississauga, ON)). Peaks were collected and immediately lyophilized, with AR55 peak identity

confirmed by matrix assisted laser desorption ionization MS using a Micromass M@LDI instrument (Waters; Mississauga, ON).

2.4. NMR spectroscopy of AR55

AR55 (1 mM) was prepared in 90% H_2O/10% D_2O containing 150 mM DPC-d_{38}, 20 mM $Na^+CD_3COO^-$, 1 mM DSS, 1 mM NaN_3 and 10 mM DTT-d_6 at pH 4.00 ± 0.05, without accounting for deuterium isotope effects. This low pH was employed due to precipitate formation in AR55/DPC micelle samples above pH 4.5. ^1H–^{15}N heteronuclear single quantum coherence (HSQC), HNCO, HNcaCO, HNCA, HNcoCA, HNCACB, HNcoCACB (500 MHz) as well as ^{15}N-edited nuclear Overhauser enhancement (NOE) spectroscopy (NOESY; 800 MHz, 125 ms mixing time following from a buildup curve for AR55 in non-deuterated lyso-palmitoyl phosphatidylglycerol micelles [46]) experiments were acquired using Varian (Palo Alto, CA) INOVA instruments at the QANUC Facility (Montreal, QC). HcCH total correlation spectroscopy (TOCSY) and ^{13}C-edited NOESY (125 ms mixing time) experiments were collected using a 700 MHz Bruker (Milton, ON) Avance III spectrometer at the NRC BMRF/NMR^3 Facility (Halifax, NS).

2.5. NMR structure calculation of AR55

NMR data were processed using NMRPipe [47] and analyzed manually using CCPNMR Analysis version 2.1.5 [48] following the main chain directed approach [49]. 3D-NOESY peak volumes were binned into distance restraint classes of 1.8–2.8 Å, 1.8–4 Å, 1.8–5 Å and 1.8–6 Å using CCPNMR Analysis. Following previous protocols [50], 100 member NOE-based ensembles were produced in XPLOR-NIH version 2.19 [51,52] and iteratively refined using a Tcl/Tk script to identify violating restraints requiring reassignment or lengthening due to poor peak integration and overlap. 11 rounds of iterative structure calculation led to a converged AR55 structural ensemble, as indicated by reported XPLOR-NIH total energies and minimal restraint violation. The 40 structures with lowest energy were retained in the final ensemble and validated with PROCHECK-NMR [53]. ϕ and ψ dihedral angle order parameters were analyzed to determine regions of converged structure, as detailed in [54]. Helical kink angles in the NMR ensemble were measured using the MC-HELAN algorithm [55].

2.6. General molecular dynamics (MD) simulation parameters

MD simulations were used both for building of the AR homology model (Section 2.7) and evaluating the validity of the "divide and conquer" derived NMR structure of AR55 in the context of the full-length AR in a lipid bilayer (Section 2.8). In all cases, MD simulations were carried out using the GROMACS 4.5.4 [56] software package, using the GROMOS53a6 forcefield [57]. Topology files were produced by standard GROMACS methodology. In all cases, structures were first subjected to a 200-step steepest descent energy minimization. Following this, canonical ensembles were subjected to 100 ps of simulation at 323 K, with a coupling constant of 0.1 ps. After stabilization of temperature, isothermal–isobaric ensembles were subjected to 1 bar pressure with a coupling constant of 5 ps for 1 ns. After equilibration, semiisotropic pressure coupling was introduced at 1 bar with a coupling constant of 1.0 ps for extended MD simulations suitable for analysis. In all cases, simulations were performed using the Nose–Hoover thermostat [58], and Parrinello–Rahman pressure coupling [59]. The Particle Mesh Ewald (PME) method [60] was employed for long-range electrostatics, with 1.2 nm cutoff for van der Waals and Coulomb interactions, updating every 10 fs. The LINCS algorithm [61] was used for covalent bond constraints, with periodic boundary conditions in all dimensions. In all cases, the system was stable for the length of the MD simulations indicated.

2.7. NMR restraint-informed homology modelling of AR

A homology model for AR was built based on the 2.5 Å X-ray CXCR4 structure (3ODU, [9]), with pairwise alignment using the conserved class A GPCR TM segment residues [62]. Insertions of Gly spacers or deletions were used to rectify length discrepancies in loops and termini. Each insertion or deletion was performed by producing a high-energy modified protein structure and subjecting it to a short (~2 ns) in vacuo MD simulation to allow polypeptide backbone relaxation. Conserved TM domain residue sidechains were fixed using the CXCR4 template for substitution of all required sidechains by SCWRL 4 [63]. The resulting model was subjected to a 10 ns MD simulation in vacuo.

From this initial model, distance restraints in the 5–12 Å range (± 1.4 Å from initial distance) were produced using ^1H–^1H distances within the TM domain. All restraints were retained for highly conserved residues vs. 30% elsewhere to reduce computational demand. Backbone dihedral angle restraints (± 40° from starting model) were calculated over the TM II–VII regions. These, along with the experimental AR55 NOE restraints and the disulfide bonds predicted for the extracellular domain [64], were used in and refined by XPLOR-NIH simulated annealing calculations. Potential scaling was carried out as in Section 2.5, with separate treatment of experimental NOEs and pseudo-NOEs. Unlike refinement of NMR-based structures, refinement of the model included only modification of model NOE-like and dihedral restraints, with fixed stringency. Those NOE-like restraints violating by > 0.3 Å in > 70% of structures were removed, while those violating by > 0.3 Å in > 50% of structures were lengthened by a 0.2 Å. In the case of dihedral restraints, those violating by > 10° in 80% of structures were widened by 5° iterations. Following 8 rounds of refinement by this method, model NOE-like restraints, and experimental NOEs contributed minimal energies to the structures. A final ensemble of 72 members was produced, from which the 18 with lowest energy were retained.

2.8. MD simulations and trajectory analysis of full-length AR

An arbitrarily selected member from the ensemble of NMR restraint-informed homology models (the selected member had no violations of experimental NOEs by > 0.5 Å and a root mean squared violation size of 0.022 Å for all experimental NOEs) was subjected to replicate MD simulations in explicit, hydrated dipalmitoylphosphatidylcholine (DPPC) bilayers. Protein and lipid were separately temperature-coupled to an external bath at 323 K with a 0.5 ps time constant (parameters based upon [65]). Membrane insertion of the AR model was performed at the center of a DPPC bilayer (482 molecules), with the AR long axis perpendicular to the membrane–solvent interface. This process was aided by the g_membed [66] computational tool. The membrane was then solvated with 41,148 water molecules using the simple point charge (SPC) model, and was neutralized via the replacement of arbitrary solvent with Cl^- counter-ions. Two replicate simulations were performed for 100 ns each. For analyses of non-time-dependent behavior over trajectories, the first 10 ns was disregarded in each case. Backbone dihedral angle order parameters (following Section 2.5), AR55 NOE violations (following Section 2.5), H-bonding patterns for N46 (based upon standard geometries, e.g. [67]) and the projected angle of the E20 and D23 sidechains over the course of the MD trajectories were determined using in-house tcl (dihedral angle and E20/D23 analysis) and Python code (NOE violations and H-bonding) to analyze Protein Data Bank (PDB) format files produced from simulation frames separated by 0.1 ns intervals (every 50 frames). The freely-available Reduce software package [68] was used to reintroduce protons to the PDB files for those analyses requiring explicit protons, as GROMOS53a6 is a united-atom forcefield without explicit aliphatic (non-polar) hydrogens. For RMSD, area per lipid, bilayer thickness and anionic face contact calculations described

below, trajectories were exposed by the MDAnalysis library [69] and analyzed using in-house Python code.

2.8.1. Tracking protein RMSD

Backbone atoms in the AR models were included in RMSD calculations and the reference structure was taken from the first frame of the first replicate simulation. For each replicate, the instantaneous protein coordinates in every hundredth frame (every 0.2 ns) were translated and rotated to produce an RMSD-minimized superposition with the reference structure [70,71]. An unweighted RMSD value was derived from the optimized superposition.

2.8.2. Area per lipid calculations

In every hundredth frame (every 0.2 ns) of the two replicate trajectories, the AR coordinates were used to define a central rectangular 'exclusion zone' in the xy-plane inside which no DPPC atoms were included in area per lipid measurements. As AR is an irregular geometric object, this vastly simplifies the calculation of the area occupied by DPPC at the cost of excluding approximately 15% of all DPPC atoms in each frame. The net area occupied by the remaining DPPC atoms was calculated as the difference between the areas of the complete simulation box and the exclusion zone (xy-plane). Some DPPC molecules straddled the exclusion boundary, but for higher accuracy those atoms that were outside the exclusion zone were nonetheless included to produce a fractional number of DPPC molecules to be considered in each analyzed frame. The average area per DPPC molecule was then calculated by dividing double the net exclusion area of a single leaflet by the fractional number of excluded DPPC molecules from both leaflets combined. The average area and standard deviation across the two replicates, excluding the first 10 ns of each replicate to allow for equilibration (900 frames considered), were also calculated.

2.8.3. Bilayer thickness calculation

DPPC phosphate atoms were assigned to a given leaflet only if they were within a given distance cutoff of another phosphate atom in that leaflet. In every 100 frames (every 0.2 ns), distance cutoffs for this networking procedure were optimized by varying between 3 and 17 Å in 2 Å steps, but in some cases the MDAnalysis-based algorithm could not identify two distinct leaflets. Phosphate particles within 12 Å of the phosphate center of geometry of the bilayer were excluded from analysis to prevent propagation of the selection network from one leaflet to the other. Only in rare cases was the algorithm not able to uniquely identify two leaflets, and in all frames the two largest networks of phosphate particles retained a consistent population suitable for estimating bilayer thickness. The last 90 ns of the two 100 ns replicate simulations was parsed for the phosphate center of geometry (Z coordinate) separation between leaflets either within an 18 Å shell of the protein or outside of it. The average and standard deviation of interphosphate bilayer thickness were both calculated.

2.8.4. Anionic face sidechain and contact analysis

Angles between the E20 and D23 sidechains were determined every 50 frames (every 0.1 ns) (i) through calculation of the pseudo-torsion angles between the centroids of the sidechain oxygens of each residue and the closest backbone heavy atom within one residue N- or C-terminal to each anionic residue and (ii) through the CB-CA-CA-CB torsion angle. Every 500 frames (every 1 ns) from the last 90 ns of the replicate trajectories was parsed for the full set of atoms within a 5 Å shell of the sidechain oxygens of either E20 or D23. Every residue in the topology (amino acid, DPPC molecule, or water molecule) was counted as a contact in a given frame if any of its atoms fell within one of these shells, and there was no duplicate counting of a given residue if many of its atoms fell within a given shell. The aggregate residue contact counts were normalized to the total number of frames parsed and the top ten most likely proximal residues were sorted for both E20 and D23 sidechain oxygens.

2.9. Mutagenesis, ERK phosphorylation assays and immunofluorescence microscopy

Single-point Ala and Ile mutants of 3HA-tagged AR were generated at G42$^{1.46}$, T43$^{1.47}$, T44$^{1.48}$, G45$^{1.49}$, N46$^{1.50}$ and G47$^{1.51}$ alongside N46D and N46Q using Phusion or QuikChange mutagenesis kits. Mutant constructs were sequenced for exactitude. HEK 293A cells were seeded into 6 well plates and grown in Dulbecco's modified eagle medium (DMEM) containing 10% FBS. At ~50% confluence, cells were transfected by adding 100 μL of a solution containing 2 μg of DNA, 20 μg of polyethylenimine (PEI) and 100 μM DMEM. The media was aspirated and replaced with DMEM 8 h later and stimulated 24 h post-transfection with 1 μM apelin-17 for 5 min. The cells were harvested with 1 mL of PBS, resuspended in 80 μL of radioimmune precipitation assay buffer (50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 150 mM NaCl, 0.5% sodium deoxycholate, 1% Nonidet P-40, 0.1% SDS, Complete Protease inhibitors, DNase I, 2 mM sodium orthovanadate and 1 mM sodium fluoride) and 20 μL of 5× SDS sample loading buffer, and cells were heated to 65 °C for 10 min. The lysate was resolved by SDS-PAGE on 12% polyacrylamide gels at 200 V for 45 min, with transfer onto nitrocellulose membrane at 115 V for 75 min. Membranes were incubated in blocking buffer (Tris buffered saline with 2% w/v BSA) for 1 h before overnight incubation at 4 °C in blocking buffer with primary antibodies for extracellular signal-regulated kinase (ERK; polyclonal rabbit anti-HA, 1:1000) and phospho-ERK (monoclonal mouse anti-HA, 1:500). Secondary antibodies (1:2000 donkey anti-rabbit Dylight 649 conjugated; 1:1000 donkey anti-mouse Dylight 549 conjugated) in blocking buffer were applied for 1 h. After final washing and drying the membrane was imaged on a VersaDoc 4000 (Biorad; Mississauga, ON).

For immunofluorescence microscopy, human embryonic kidney (HEK) 293A cells were grown on coverslips in DMEM containing 10% FBS until ~50% confluence. Transfection was performed as above, with the exception that only 12 μg of PEI was used and media was refreshed with DMEM with 10% FBS. Cells were washed in warm PBS and fixed in 2% paraformaldehyde in PBS for 10 min 24 h post transfection. Coverslips were blocked with PBS containing 5% normal donkey serum for 20 min then incubated with primary antibody for 1 h (1:1000 mouse anti-HA). After washing 3× in PBS for 5 min, coverslips were incubated with secondary antibody (1:500 donkey anti-mouse, Cy5 conjugated) for 30 min and mounted using fluorescence mounting medium. Fluorescence images were acquired on a Zeiss Axio Observer Z1 inverted microscope (Intelligent Imaging Innovations (3i); Boulder, CO) using a solid state spectra light source (Lumencor; Beaverton, OR). Cells were observed using a 40× objective (1.3 N.A) and images were recorded using an HQ2 charge-coupled device camera (Photometrics; Tucson, AZ) and Slidebook 5.0 software (3i; Boulder, CO). For each AR mutant, at least three representative fields of view containing 5–10 cells each were collected. Photoshop CS5 (Adobe) was used to adjust brightness and crop images.

2.10. Database accession details

The structural coordinates for the final ensemble of AR55 have been deposited in the PDB ([72]; accession # 2LOU). The chemical shift, spectral peak list and restraint data have been deposited to the Biological Magnetic Resonance Bank (BMRB [73], accession # 18225).

3. Results

3.1. Structure of AR55 in DPC micelles

Despite chemical shift overlap typical of a helical TM domain, spectral quality of AR55 in DPC micelles was sufficient to assign all 64 residues (Fig. 1) through the main chain directed backbone walk

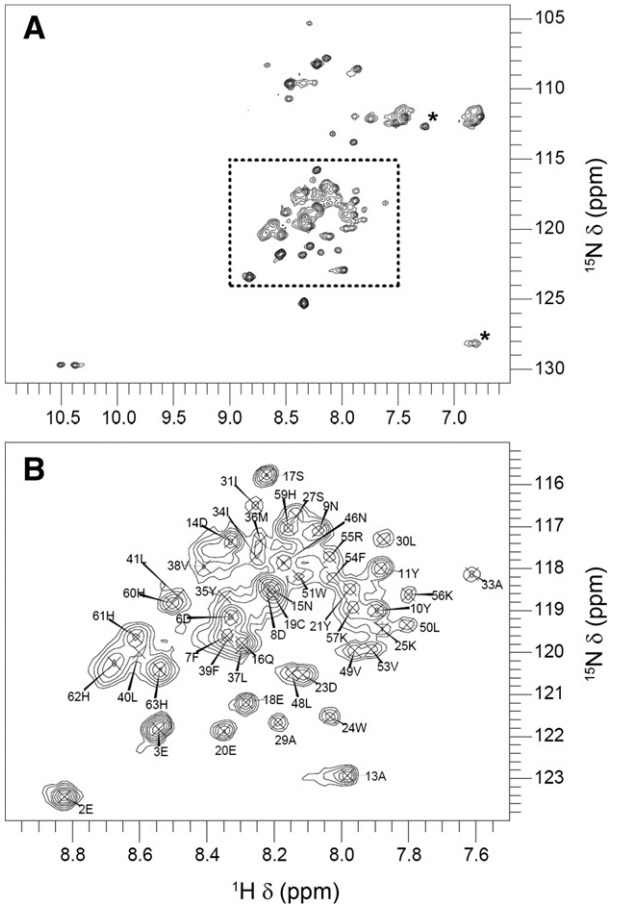

Fig. 1. Two-dimensional ^{1}H–^{15}N HSQC spectrum of AR55 solubilized in DPC micelles at 37 °C at 700 MHz ^{1}H frequency. (B) Zoom to the region boxed in A, with peak assignments annotated. Aliased peaks are indicated with *'s.

Table 1
Distance restraints, violations and Ramachandran plot statistics of the 40-member ensemble of AR55 solubilized in DPC micelles.

Unique distance restraints	
Total	841
Intraresidue	394
Sequential	212
Medium range (|i–j|≤4)	161
Long range (|i–j|>4)	0
Ambiguous	74
Ramachandran plot statistics	
Core	39.2%
Allowed	47.3%
Generously allowed	9.3%
Disallowed	4.2%
XPLOR-NIH energies (kcal/mol)	
Total	39.80 ± 1.75
NOE	1.07 ± 0.59
Violations	
NOE violations > 0.5 Å	0
NOE violations of 0.3–0.5 Å	1
NOE violations > 0.2–0.3 Å	0

approach [49]. This yielded chemical shift assignments for 98% of ^{1}H, 86% of ^{13}C and 80% of ^{15}N nuclei (BMRB Entry #18225). Some slow-exchange conformational sampling is also apparent in the form of minor, unassigned peaks in the ^{1}H–^{15}N HSQC spectrum (Fig. 1) and other triple-resonance experiments (not shown), most likely arising from conformational exchange in the N-terminal tail of AR55.

In total, 841 unique distance restraints were determined (Fig. 2; Table 1). The distribution of distance restraints is not uniform throughout the protein, with the N-terminal tail of AR55 (residues ~1–28) having fewer medium range restraints than the putative membrane-spanning region (residues ~29–51, Fig. 2A). Of the assigned distance restraints, canonical i,i + 3 and i,i + 4 helical contacts [74] are primarily observed from A29-L40$^{1.44}$ and from T44$^{1.48}$-R55$^{1.59}$, with some in the

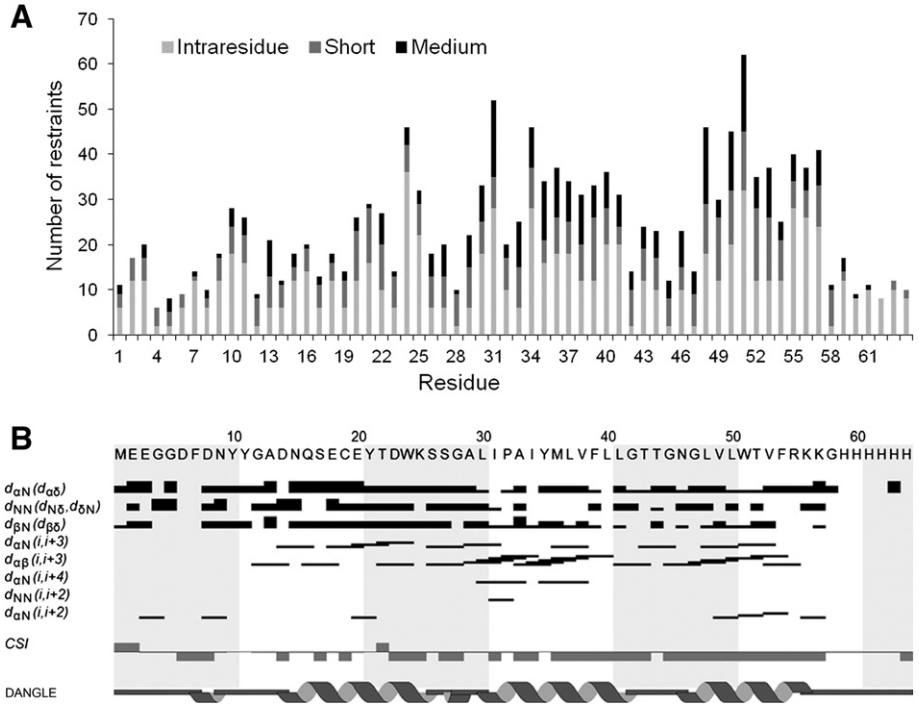

Fig. 2. Summary of NOE assignments for AR55 in DPC micelles. (A) Distance restraints on a per-residue basis (short is |i–j| = 1; medium is |i–j|≤4). (B) Graphical overview of distance restraints, chemical shift index [75] and DANGLE secondary structure prediction [76] generated using CCPNMR Analysis [48].

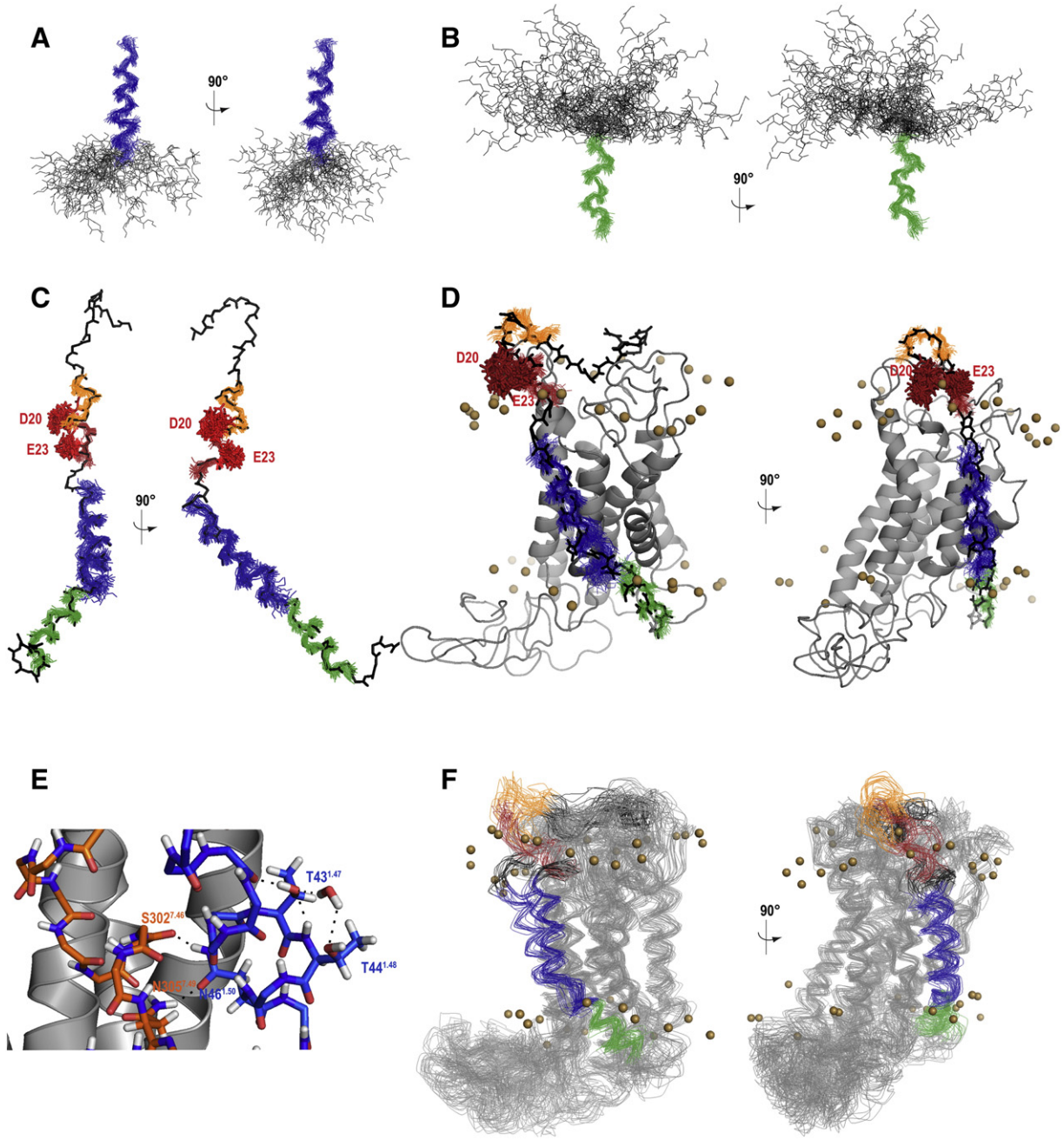

Fig. 3. Structure of AR55 in DPC micelle solution and in context of MD simulations of full-length AR in hydrated DPPC bilayer. (A–D) are superpositions of the 40-member AR55 NMR-based structural ensemble. In (A–B), the kinked helical TM domain (A29-K57$^{1.61}$) of AR55 is shown superposed over (A) A29-N46$^{1.50}$ (blue) or (B) G47$^{1.51}$-K57$^{1.61}$ (green). In (C–D), all four structurally converged regions of the NMR ensemble are superposed, with D14-C19 in orange, E20-K25 in dark red, A29-N46$^{1.50}$ in blue and G47$^{1.51}$-K57$^{1.61}$ in green; the E20 and D23 sidechains are shown for all 40 ensemble members with positions based on the backbone superposition of E20-K25. In (C), the superposition is onto the lowest-energy conformer of the NMR ensemble (backbone shown in black); in (D), onto a frame from the replicate MD simulation of full-length AR (backbone of residues 1–55 illustrated in black) of the 40 randomly selected frames in (F) with the lowest overall RMSD over the converged regions. (E) Detailed configuration of an arbitrary, representative simulation frame, including sidechain positions of T43$^{1.47}$, T44$^{1.48}$ and N46$^{1.50}$ (blue sticks), S302$^{7.46}$ and N305$^{7.49}$ (orange sticks) and a water molecule associated with TM I kink (red/grey). H-bonds in (E) are illustrated by dashed black lines. (F) Superposition over all 7 TM domains of 40 random frames from the AR simulation upon AR simulation frame used in D (colouring over AR55 region as in D). Spheres in D and F are phosphorus atom positions in proximal DPPC headgroups for the simulation frame used in D. All panels were produced in PyMol (Schrödinger).

N-terminal tail (Fig. 2B). The chemical shift index (CSI; [75]) suggests that AR55 is largely helical from D23-K57$^{1.61}$. Finally, the DANGLE algorithm [76] predicts the membrane spanning region of AR55 to be composed of two helices with a disruption in helical character from G42-N46 and the presence of a third helix in the N-terminal tail of AR from N15-K25 (Fig. 2B). Taken together, these data imply a helix from ~A29-K57$^{1.61}$ alongside some additional helical character in the N-terminal tail.

Using the NOE derived distance restraints, a structural ensemble (Fig. 3A–C; 40 lowest energy members of 100 calculated) was generated in excellent agreement with the experimental data (Table 1). Analysis of dihedral angle variability in the structural ensemble [54] indicates that the D14-C19, E20-K25, A29-N46$^{1.50}$ and G47$^{1.51}$-K57$^{1.61}$ regions of AR55 are well converged (Fig. 4A–B; Supplementary data Fig. S1). Superposition yields backbone atom root mean square deviation (RMSD) values of 0.70, 0.66, 1.33 and 0.99 Å for these 4 regions over

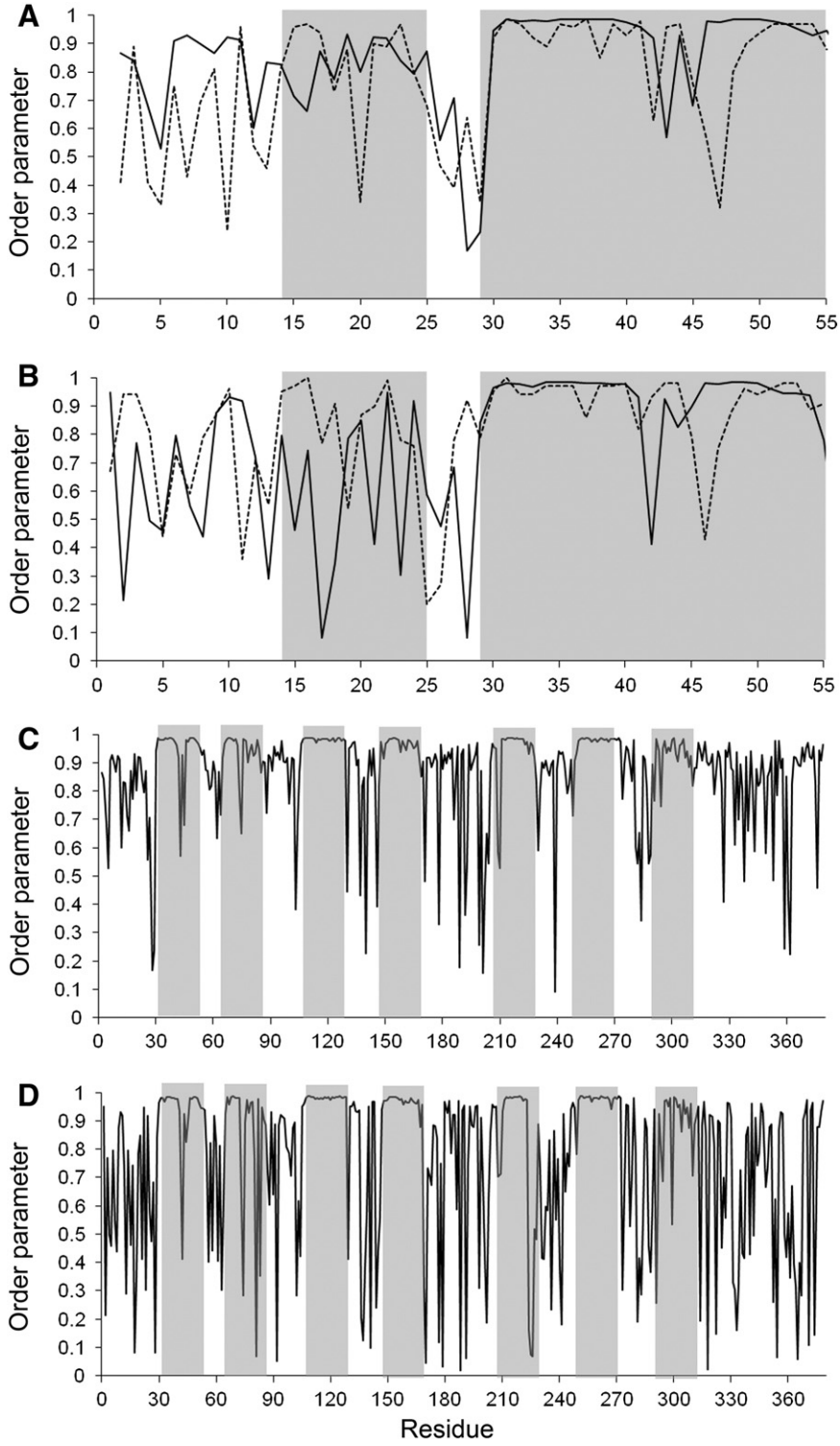

Fig. 4. Polypeptide backbone dihedral angle convergence in the AR55 NMR derived ensemble and replicate MD simulations. Dihedral angle order parameters [54] are shown for ϕ (A and C) and ψ (B and D) over the 40-member AR55 NMR ensemble (A and B; dashed lines) and for the ~60 ns replicate MD simulation trajectories (A–D, solid lines). All helical regions of AR55 are shaded in A and B, while the TM helices are shaded in C and D.

the NMR ensemble, respectively. The remainder of AR55, including parts of the N-terminal tail and the His_6-tag, is largely variable between ensemble members.

The N-terminal tail of AR55 has helical character from D14-K25, with a pivot point at C19-E20 allowing conformational sampling [54] of the helical segments to either side. The E20 and D23 sidechains fall on the same face of the E20-K25 helical segment, producing an anionic surface (Fig. 3C). The region of AR55 corresponding to TM I in AR has a helix-kink-helix structure, with the kink occurring in the relatively polar region (GTTGNG) around $N46^{1.50}$. A break in helical character at the kink is evident both from a lack of the characteristic helical NOE contacts in its vicinity and from the deviation away from chemical shift-based helix prediction at the kink according to both CSI and DANGLE. It should be noted that the relative positions and orientations of each section of the TM helix are not well defined (Fig. 3A–B). The observed variability in relative helix position and orientation is due to

a lack of medium- or long-range NOE contacts in this region and leads to a wide variability in the calculated kink angle [55], with a mean value of 69 ± 20°. While the relative orientations of the helical segments of the TM domain could likely be better defined using an approach such as measurement of residual dipolar couplings [49], this would be relatively challenging to perform in a cost-effective manner due to the extensive spectral overlap of 2D NMR experiments on AR55 (e.g., Fig. 1A) and relatively short transverse relaxation time constants for residues in the TM domain (average ^{15}N T_2 of ~45 ms over residues 30–54) with correspondingly broad line widths. We have therefore opted instead to characterize this in the context of full-length AR using MD simulations on an NMR restraint informed homology model (Section 3.3).

3.2. Function of 3xHA AR mutants

Following previous demonstration of AR stimulation leading to ERK phosphorylation in AR transfected HEK 293A cells [77], this process was monitored for transiently transfected 3xHA AR mutants upon stimulation with 1 μM apelin-17 for 5 min. Mutation of G42$^{1.46}$, G45$^{1.49}$ and N46$^{1.50}$ impaired AR activation compared to wild-type AR, with mutation to Ile having a more pronounced effect than mutation to Ala (Fig. 5). N46$^{1.50}$ was particularly sensitive to mutation, with N46A, N46D and N46I mutants all having severely impaired activation. The N46Q mutant, conversely, had approximately wild-type activity.

Through immunofluorescence microscopy, the general localization of 3xHA AR and each of the AR mutants was established. For the wild-type receptor and mutants with approximately wild-type activity, proper membrane trafficking and localization of 3xHA AR was apparent (Fig. 6; Supplementary data Fig. S2). Conversely, mutants with decreased activation (G42I, G45I, N46A, N46D and N46I) showed abnormal receptor localization on the membrane surface due to either improper trafficking or folding.

3.3. Simulation-based context for AR55 structure

A homology model of AR based on a 2.5 Å crystal structure of the related GPCR CXCR4 [9] was built. NMR-styled distance and dihedral angle restraints were produced using the model and supplanted by the NMR-derived restraints (Table 1; Fig. 2) in the AR55 region. Following simulated annealing structure calculations, an arbitrary member of the ensemble of NMR restraint-informed homology models was embedded into a hydrated DPPC bilayer and simulated twice using unified-atom MD methods for 100 ns. Simulations were performed in the complete absence of any distance or dihedral restraints and were carried out using well-parameterized DPPC bilayers with physiologically

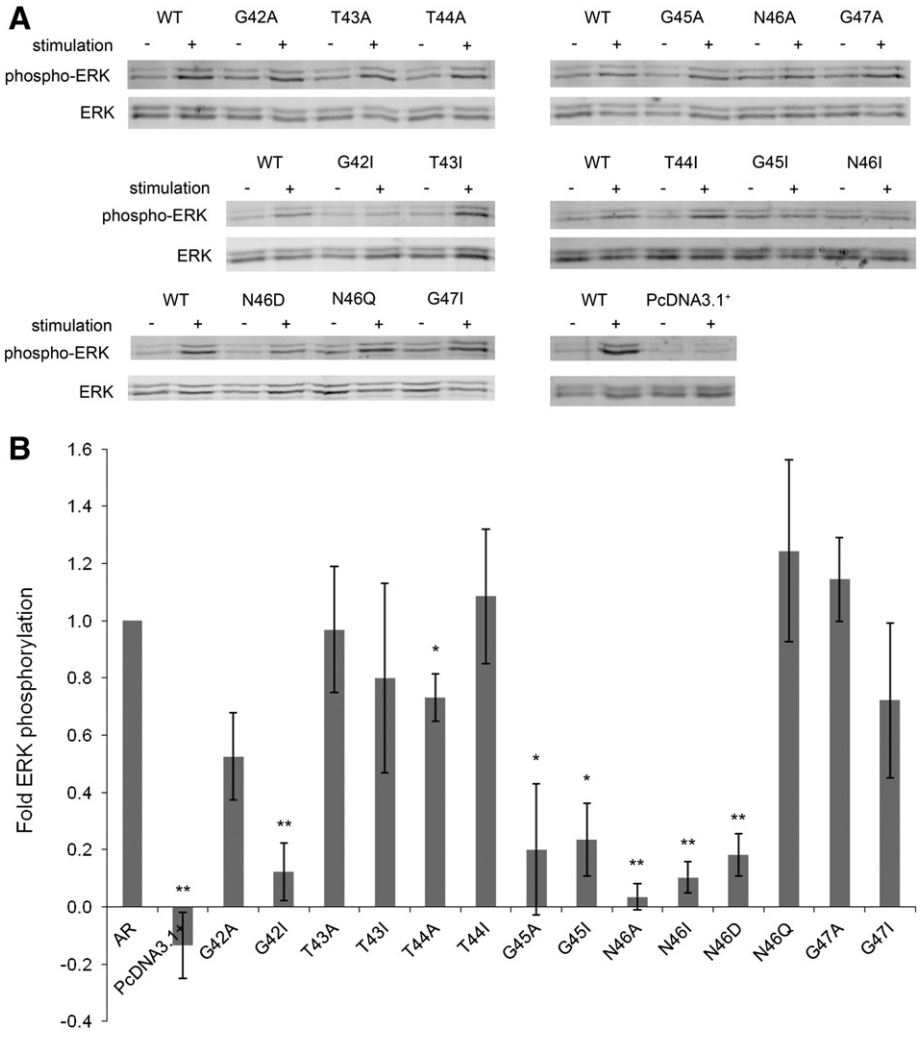

Fig. 5. Functional analysis of 3xHA AR mutants. The fold ERK phosphorylation compared to wild-type 3xHA receptor (AR) after stimulation by apelin-17 is shown for a representative experiment in (A) and plotted for all replicates in (B) for all mutants alongside the empty vector control (PcDNA 3.1$^+$). Standard error is shown, with each experiment replicated at least 3 times (* = $p<0.05$; ** = $p<0.01$).

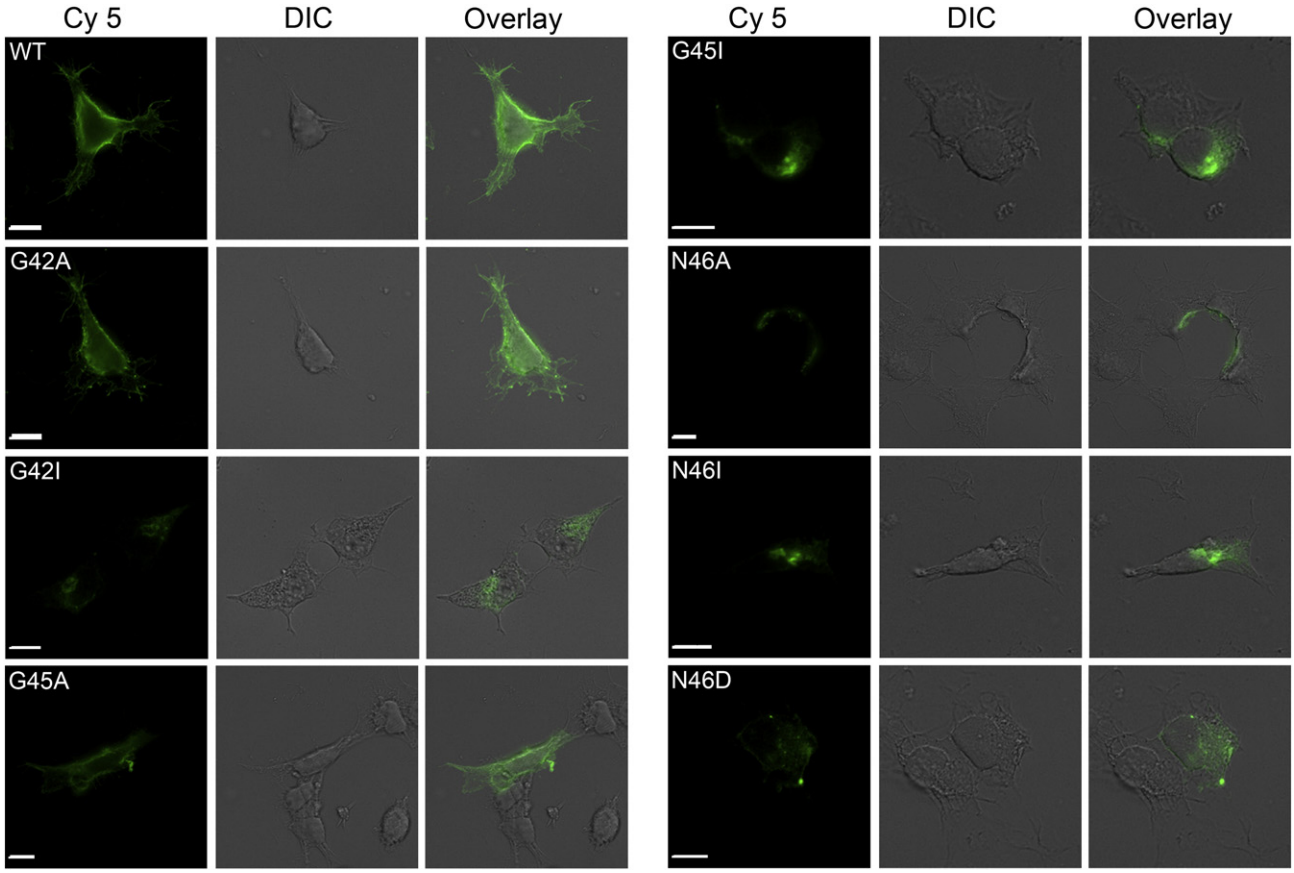

Fig. 6. Distribution of 3xHA AR in transiently transfected HEK 293A cells. Immunofluorescence microscopy showing a differential interference contrast (DIC) field of view overlaid by the signal from Cy5 bound to 3xHA AR is shown for WT AR and the functionally inhibited mutants. Scale bars are 10 μm.

relevant head and tailgroup chemistry. Equilibration of the MD simulations was validated both in terms of AR behavior and phospholipid behavior.

The RMSD of the AR reached a reasonably stable plateau in the two 100 ns replicate simulations (Fig. 7A), and the first 10 ns was discarded for the purpose of several analyses in order to allow for equilibration. Only backbone atoms of the AR are considered in this analysis of structural deviation relative to the first frame of the first replicate simulation. Examination of ϕ and ψ dihedral angle order parameters over the course of the merged 180 ns MD trajectory demonstrates well maintained structure in the TM domains (Fig. 4C–D) and more variability in the loops and termini. This is evident in the ensemble of 40 randomly chosen AR frames (20 from the last 90 ns of each replicate) superposed in Fig. 3F and provided as a PDB file in the Supplementary data. In terms of bilayer behavior, the average area per DPPC molecule was tracked as a function of time in each of our replicate AR MD simulations (Fig. 7B). The trends are stable and similar for the two replicates, consistent with an equilibrated system (i.e., a properly packed bilayer). Our overall average area per lipid at 323 K (61.0 ± 0.8 Å^2) is consistent with previous MD simulations of pure DPPC bilayers at 323 K reporting values of: ~62 Å^2 [78], 62 ± 1 Å^2 [79], ~64 Å^2 [80], 64.7 ± 0.2 Å^2 [81], 64.5 ± 1.0 Å^2 [82], 65 ± 1 Å^2 [83], and 66.8 ± 0.7 Å^2 at 350 K [84]. It is also noteworthy that the presence of a dynamic membrane protein in our simulations could be reasonably expected to modestly influence the average area per lipid, perhaps causing slightly tighter packing. Although occasional protein-local thinning of the DPPC bilayer was observed by visual inspection of the trajectories, a quantitative analysis of the average interphosphate bilayer thickness within an 18 Å shell of the AR (37.9 ± 0.7 Å) and outside of this shell (37.4 ± 0.5 Å), over 2×90 ns of simulation time, does not provide significant evidence of bilayer thinning within one standard deviation.

As a whole, the models derived from simulated annealing and the subsequent MD simulation trajectories are in good agreement with the experimental structure of AR55 (Fig. 3D) and mutagenesis data. While the initial models produced using NMR structure calculation protocols violated no experimental distance restraints, MD simulation led to AR conformations where ~10% of the unique NOEs would be violated in >50% of the frames sampled (Supplementary data Table S1) with varying degrees of mismatch between the NMR data and the simulation. NOE "violations" in the MD trajectories are not localized to a particular region of AR55, rather they are spread over the entire 55 residue segment. Of the NOEs violated in >5% of frames (Supplementary data Table S1), >80% involve one or more sidechain atoms, implying variant positioning of sidechains between the isolated AR55 segment and the full-length, bilayer embedded protein. Notably, the backbone atom-only NOE contacts that were violated almost exclusively fell in the N-terminal tail or intracellular loop 1 portions of AR55 (the same regions which display lower ϕ/ψ dihedral order parameter indicative of increased dynamics), with only 4 TM domain backbone NOEs producing an appreciable number of violations and all localized around the kink. The modest mismatch observed between the experimental AR55 distance constraints and MD trajectories is therefore likely almost exclusively caused by variation in sidechain positions and in backbone conformational dynamics outside of the TM domain. As such, the experimentally converged regions of AR55 agree well with stable structuring during the MD simulation trajectories, as is readily seen in superposition of the NMR ensemble over each of these regions upon an arbitrary simulation frame (Fig. 3D) and, more generally, in examination of the backbone dihedral angle order parameters for the set of MD trajectories over the AR55 region (Fig. 4A–B). Specifically, high dihedral angle order is seen in the helical transmembrane region but not near N46$^{1.50}$, which is where the kink is observed in AR55. Throughout the course of the MD

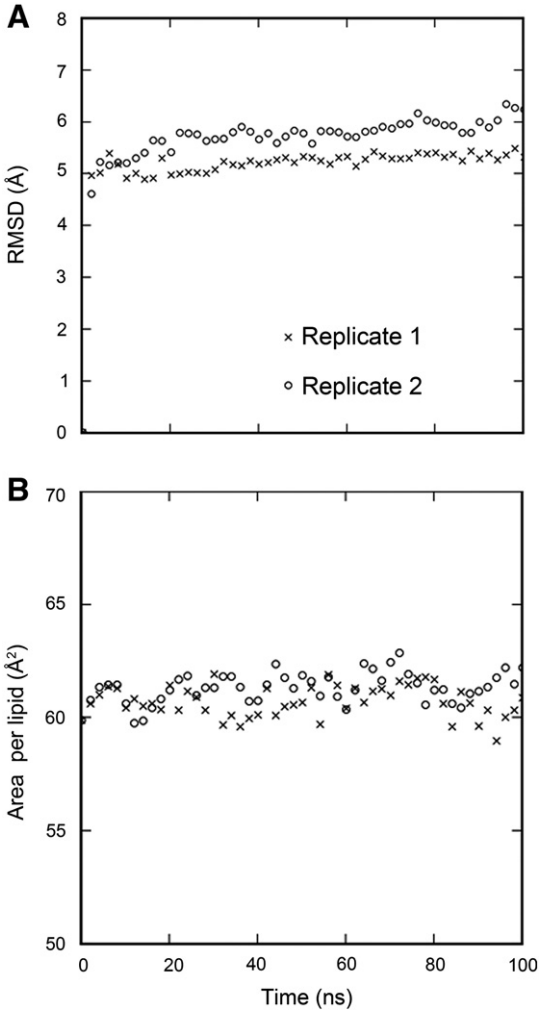

Fig. 7. Tracking protein and lipid behavior in the molecular dynamics simulations. (A) The RMSD of backbone heavy atoms in the apelin receptor was tracked over the length of each trajectory with reference to the first frame of the first replicate simulation. (B) The area per DPPC molecule was tracked in each replicate simulation after excluding all atoms and the corresponding area occupied by a rectangular zone bounding the apelin receptor. For both analyses, every tenth data point was plotted to improve clarity.

simulations, the hydrophilic and kinked GTTGNG segment of the TM I kink falls within the hydrophobic tailgroup region of the bilayer (e.g. the approximate DPPC headgroup phosphate moiety positions illustrated in Fig. 3D and F).

Although more variation in backbone structure is apparent in the MD trajectories than in the NMR ensemble over the N-terminal tail, in most of the simulation frames, the region surrounding E20 and D23 is ordered, with the sidechains of these residues forming an anionic surface. Over both MD trajectories, the torsion angle between the vectors formed between the closest backbone heavy atom and the centroid of the sidechain O^{γ} or O^{δ} atoms for these two residues had a mean of $9.0 \pm 35.3°$ (range $-123.1°$–$104.0°$; Fig. S3) while these centroids were separated by a mean distance of 8.9 ± 3.2 Å (range 4.3–15.6 Å). The E20 and D23 sidechain oxygens both localized prominently near DPPC molecules (Table 2; Fig. 3D and F), with most of the prominent amino acid interactions for E20 and D23 falling proximally within primary sequence rather than on neighbouring transmembrane segments or extracellular loop segments.

Table 2
Identities of the ten most frequently observed interacting residues for the E20 and D23 sidechain oxygens during the last 90 ns of two 100 ns MD simulation trajectories.

Residue in topology	% frames within 5 Å of E20 sidechain oxygens	Residue in topology	% frames within 5 Å of D23 sidechain oxygens
K25	9.9	L30	34.1
DPPC (658)	14.3	A29	36.3
C19	17.0	G28	44.5
Q16	26.4	DPPC (406)	55.5
W24	46.2	S27	56.0
D23	51.1	DPPC (496)	57.1
Y21	55.5	DPPC (404)	59.9
T22	66.5	T22	64.3
DPPC (406)	95.6	W24	78.6
E20	100.0	D23	100.0

4. Discussion

4.1. A kink is present in TM I of AR in a functionally sensitive region

In the AR55 NMR ensemble, a kink is present near residue $N46^{1.50}$. This kink is observed both in the isolated AR55 NMR structural ensemble and throughout the replicate MD trajectories on full-length AR in a hydrated bilayer (Fig. 3). During the MD trajectories, the exact kink location varies to some degree, generally falling close to $T44^{1.48}$ (Fig. 3F). Notably, both ground state and light activated bovine rhodopsin (PDB entries 1F88 and 4A4M) as well as inactive CXCR4 (3OE6) have kinks in TM I in proximity to the homologous position of the kink in AR55. In the case of rhodopsin, this clearly arises from a kink-inducing Pro residue [55] at position 1.48. It has been shown that TM helix kinks may arise from vestigial prolines [85]; however, none of the known AR sequences from related organisms include a proline residue near $T44^{1.48}$, making this explanation for the kink unlikely. CXCR4 is kinked at position 1.49 (as determined by the MC-HELAN algorithm [55]) but is hydrophobic at positions 1.47 and 1.48, with a sequence of $G^{1.46}$IVGN$G^{1.51}$ corresponding to $G^{1.46}$TTGN$G^{1.51}$ in AR. Notably, a number of other GPCR structures (e.g. the squid rhodopsin (2Z73), several β2-adrenegric (2R4R, 2R4S, 3D4S, 3KJ6, 3NY8, 3NY9, 3NYA, 3SN6), two adenosine A_2A (3EML, 3QAK), the dopamine D3 (3PBL) and the sphingosine 1-phosphate (3V2W) receptor structures) have an unkinked TM I.

In order to evaluate the importance of this kink, functional effects of mutations in its vicinity were tested. As is clear from both ERK phosphorylation (Fig. 5) and AR localization (Fig. 6), mutations at $G42^{1.46}$, $G45^{1.49}$ or $N46^{1.50}$ severely affect the folding and/or trafficking of AR. $N46^{1.50}$ is completely conserved in class A GPCRs [86], suggesting that it is essential for receptor folding or function. The N46D AR mutant was inactive, while the N46Q mutant showed the same response to apelin as wild-type AR. This indicates that the amide moiety of the sidechain is essential for proper AR folding and localization, but that there is ability to accommodate variation in size. The reason that an N to Q mutation at site 1.50 is not more prevalent in nature is thus an interesting question.

In the currently solved crystal structures of GPCRs, the sidechain of $N^{1.50}$ is H-bonded to residue 7.46 in TM VII through the backbone carbonyl oxygen of $S^{7.46}$ (β_1-adrenergic, β_2-adrenergic, A_2A adenosine, dopamine D3, histamine H1, M2 muscarinic, S1P(1), κ-opioid and μ-opioid receptors; [3–5,7,8,10–13]), $C^{7.46}$ (CXCR4; [9]) or $A^{7.46}$ (bovine rhodopsin; [6]). This connection is retained in our model and MD simulations of AR (last 90 ns of each replicate), with an H-bond between the $N46^{1.50}$ sidechain and the $S302^{7.46}$ backbone carbonyl oxygen being observed in 63.7% of simulation frames (e.g. Fig. 3E). An additional 52.1% of simulation frames have an H-bond between the sidechains of $N46^{1.50}$ and $D75^{2.50}$ of TM II. The overwhelming majority (93.7%) of simulation frames exhibit at least one of these H-bonding geometries, and 22.0% of frames are consistent with H-bonding to both partners simultaneously. A homologous contact between $N46^{1.50}$ and $D73^{2.50}$ is observed in the histamine H1 receptor structure [7]. This leads to the speculation that flickering between these H-bond configurations involving $N^{1.50}$ may be a more widespread and functionally significant feature of GPCR dynamics.

Residues G42$^{1.46}$ and G45$^{1.49}$ are important for receptor trafficking (Figs. 5 and 6) and are highly conserved in class A GPCRs (40–60% and 60–80% respectively, [86]) but do not show specific conserved inter-helix interactions in the manner of N46$^{1.50}$. The importance of Gly for inter-helical packing is well established [87] and these residues may play a role for helix packing in AR. In addition, glycine allows for helix flexibility [88], which would be essential given the dynamic nature of the nearby kink. Throughout our AR MD simulations, TM I is bundled against both TM II and TM VII. Although the specific interactions vary between frames, G42$^{1.46}$ lies against TM VII (predominantly) or TM II (infrequently) while G45$^{1.49}$ interfaces exclusively with TM VII. As a whole, the high sensitivity of residues near the TM I kink to mutation highlights its importance for proper membrane localization of AR.

The amino acid identities at positions 1.47 and 1.48 are conserved in 40–60% of class A GPCRs. Thr residues, as in AR, are quite rare, with only 4% and 3% of class A GPCRs having Thr at positions 1.47 and 1.48, respectively [86]. Even in the most homologous of the GPCRs, the AT_1 receptor (~40–50% sequence identity in the TM domain [15]), the corresponding positions are hydrophobic residues. However, neither T43$^{1.47}$ nor T44$^{1.48}$ appears to be very sensitive to mutation in the context of the activation studies employed herein.

In our AR model, these residues demonstrate considerable dynamics and have the γ-OH groups protruding into the hydrophobic region of the bilayer (see, e.g., headgroup phosphate positions illustrated in Fig. 3D and F), with many frames of the simulation showing H-bonding back to TM I of AR (Fig. 3E). During the replicate MD simulations of AR in a hydrated bilayer, water molecules consistently became associated with the γ-OH group of these sidechains within approximately 2 ns of simulation time. This does not preclude simultaneous H-bonding to the main chain carbonyl O of other residues in TM I (e.g. Fig. 3E). The variation in exact H-bonding to proximal water(s), the main chain of TM I residues, or to both during the MD simulations would lead to the modulation in kink properties and location described above. Further illustrating that this plasticity is reasonable, mutation of neither T43$^{1.47}$ nor T44$^{1.48}$ causes any appreciable loss of function.

4.2. The membrane-proximal E20/D23 anionic patch forms a likely binding site for apelin

In the otherwise disordered N-terminal tail of AR55, we observed a helix from D14 to K25 containing the residues E20 and D23 previously shown to be essential for apelin binding to AR [26]. Residues E20 and D23 are structured such that their anionic sidechains are in close proximity (~9 Å apart) on the same face of the tail, both in our experimental AR55 structure and in our MD simulations (Fig. 3C–D). The MD frame used in Fig. 3D gives rise to the E20 and D23 sidechains facing outwards upon superposition of the AR55 NMR ensemble, an orientation generally upheld in tracking of the atomic species in closest proximity over the entire pair of MD trajectories (Table 2). Specifically, >90% of frames show DPPC molecules within the 10 most proximal moieties to the anionic patch.

Although the N-terminal tail and TM I region of GPCRs are not generally implicated in ligand binding, they are, e.g., important for ligand specificity and efficacy in the dopamine D1 and D5 receptors [90]. A short helix preceding TM I has also been observed in at least 2 other GPCRs, in one case amphipathic [89] as with E20-K25 of AR, and in the other not [10]. In the former case (the Y4 receptor), the helical segment is postulated to undergo context specific restructuring in the presence or absence of its peptidic ligand; in the latter case ($S1P_1$), this helix forms part of a movable cap for the binding pocket for a lipid ligand. Given the diversity in class A GPCR N-terminal tail length and structure, variant structuring and roles of N-terminal tails seem reasonable. It should also be noted that in most of the GPCR crystal structures, this region is not resolved. Considering the N-terminal structuring of the Y4 and $S1P_1$ receptors in more detail, neither our MD simulation trajectories nor NMR structural data are indicative of capping of the TM domain, as in $S1P_1$, by D14-E25 in AR (Fig. 3C–D). Instead, a mechanism akin to that postulated for the Y4 receptor seems quite reasonable.

Although the binding mechanism of apelin to AR is not yet known, one popular model of peptide-GPCR interactions is a two-step mechanism where the ligand initially binds the GPCR extracellular domain and then works its way deeper into the GPCR TM domain to trigger activation [91–94], with a fly-casting mechanism frequently being invoked as part of this process [94,95]. With the N-terminal tail of AR55 being so flexible and with multiple conformations observed in the NMR data, it is possible that initial interactions of apelin with AR55 (the first step of the 2-step model) may be facilitated through fly-casting behavior. A general two-step mechanism fits well with the recent proposal of Iturrioz et al. that the C-terminus of apelin penetrates deeply into the TM bundle during activation [27]. In this previous study, mutagenesis results were rationalized in the context of a bacteriorhodopsin-based homology model of AR to state that residues W154$^{4.50}$, F257$^{6.44}$ and W261$^{6.48}$ are in close proximity to the C-terminal Phe of apelin in the bound state. Notably, some aspects of the model presented herein are inconsistent with this previous model. While our model shows a similar pocket of aromatic amino acids at the putative apelin binding site, TM IV is not involved. W154$^{4.50}$ is distant from the binding pocket and, instead, forms a polar contact with A69$^{2.44}$. This may help to rationalize the result of Iturrioz et al. that mutation to W154$^{4.50}$ severely affects folding and trafficking of AR without a requirement for proximity to the ligand-binding pocket.

The membrane-catalysis model, which we have proposed applies to apelin-AR binding [24], adds a necessary precursor step to ligand-receptor binding, where the ligand initially associates with the membrane and, potentially, undergoes a structural change prior to receptor recognition, binding and activation [96,97]. Apelin is highly cationic (pI of apelin-17 = 12.5) and the anionic surface created by E20 and D23 would therefore be a prime target for an electrostatic interaction. In the absence of ligand, our MD simulations indicate that both E20 and D23 have prominent interactions with a small number of DPPC molecules, but not with residues in other helices of AR (Table 2). The zwitterionic lipid headgroups may stabilize the anionic patch until they are displaced by the cationic ligand. This is reminiscent of the ligand-induced conformational change postulated previously for the Y4 receptor. Binding of apelin, very probably at the "RPRL" motif [20,23], to E20 and/or D23 and subsequent liberation of the D14-E25 helix from the membrane would bring the relatively hydrophobic C-terminus of apelin into sufficiently close proximity to find its way into its postulated binding site [27] within the AR helical bundle. Upon binding, the dynamic nature of the N-terminal tail (Fig. 3F), including the structurally converged anionic patch region, would permit subsequent conformational excursion (or, fly-casting) of the tail-ligand complex allowing the unbound and relatively unencumbered apelin C-terminus to access and bind in the core of the AR TM helical bundle, inducing receptor activation.

4.3. Conclusions

This study illustrates a strong correlation between functionally essential residues and structural features in the N-terminus and first TM segment of AR. In particular, we have identified an anionic patch in the N-terminal tail of AR55 comprised of residues essential for apelin binding. This patch is fully accessible on the protein surface and poised in proximity to the membrane surface, providing an ideal membrane-catalyzed binding site for a highly cationic membrane-associated apelin ligand. A noncanonical kink in the otherwise helical TM I segment is located near N46$^{1.50}$ and a pair of glycine residues allowing close inter-helical packing. AR TM I is also unusually polar in the vicinity of this kink, with two Thr sidechains oriented toward the nonpolar tailgroup region of the bilayer. The divide and conquer approach, expanded in scope using hybrid experimental data-

homology model derived MD simulations, has thus provided the first high-resolution structural detail about AR and strongly improved insight into apelin binding to and activation of AR.

Acknowledgements

This research was supported by Canadian Institutes of Health Research (CIHR) Operating Grants to JKR (MOP-111138 and ROP-91807) and GD (MOP-84260) and by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to DJD (RGPIN/355310-2008). ROP-91807 was partnered through the Nova Scotia Health Research Foundation and Dalhousie University. DNL was the recipient of a Canada Graduate Scholarship (CGS) from NSERC; TR was previously supported by an NSERC CGS and a Killam Predoctoral Scholarship and is currently supported by a CIHR Postdoctoral Fellowship; AWB was the recipient of a CGS from the CIHR and a Killam Predoctoral Scholarship; GD, DJD and JKR are supported by CIHR New Investigator Awards; and, GD is a Senior Scientist of the Beatrice Hunter Cancer Research Institute and the Cameron Research Scientist of the Dalhousie University Cancer Research Program.

NMR experiments were recorded (1) with support from Dr. Tara Sprules at the Québec/Eastern Canada High Field NMR Facility, supported by NSERC, the Canada Foundation for Innovation (CFI), the Québec ministère de la recherche en science et technologie, and McGill University; and, (2) at the NMR^3 Facility, supported by Dalhousie University, using the 700 MHz housed at the National Research Council of Canada's Institute for Marine Biosciences (NRC-IMB) and a He-cooled probe funded by an Atlantic Canada Opportunities Agency Grant to Dalhousie University. Simulations were performed using resources at Dalhousie University and, predominantly, at the Structural Bioinformatics & Computational Biochemistry Unit, Oxford, U.K. Grants from NSERC, the Dalhousie Medical Research Foundation and the CFI provided equipment essential for this work. We thank Dr. Philip Fowler (SBCB, Oxford, UK) for helpful advice regarding simulations. Finally, we are grateful to Dr. Devanand M. Pinto (NRC-IMB) for providing MALDI-MS access.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2013.02.005.

References

[1] D.K. Vassilatis, J.G. Hohmann, H. Zeng, F. Li, J.E. Ranchalis, M.T. Mortrud, A. Brown, S.S. Rodriguez, J.R. Weller, A.C. Wright, J.E. Bergmann, G.A. Gaitanaris, The G protein-coupled receptor repertoires of human and mouse, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 4903–4908.
[2] J.P. Overington, B. Al-Lazikani, A.L. Hopkins, How many drug targets are there? Nat. Rev. Drug Discov. 5 (2006) 993–996.
[3] E.Y. Chien, W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov, R.C. Stevens, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, Science 330 (2010) 1091–1095.
[4] V.P. Jaakola, M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y. Chien, J.R. Lane, A.P. Ijzerman, R.C. Stevens, The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist, Science 322 (2008) 1211–1217.
[5] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, G.F. Schertler, Structure of a beta1-adrenergic G-protein-coupled receptor, Nature 454 (2008) 486–491.
[6] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, Crystal structure of rhodopsin: a G protein-coupled receptor, Science 289 (2000) 739–745.
[7] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han, T. Kobayashi, R.C. Stevens, S. Iwata, Structure of the human histamine H1 receptor complex with doxepin, Nature 475 (2011) 65–70.
[8] S.G. Rasmussen, H.J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R. Ratnala, R. Sanishvili, R.F. Fischetti, G.F. Schertler, W.I. Weis, B.K. Kobilka, Crystal structure of the human beta2 adrenergic G-protein-coupled receptor, Nature 450 (2007) 383–387.
[9] B. Wu, E.Y. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi, D.J. Hamel, P. Kuhn, T.M. Handel, V. Cherezov, R.C. Stevens, Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, Science 330 (2010) 1066–1071.
[10] M.A. Hanson, C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer, M.G. Sanna, G.W. Han, P. Kuhn, H. Rosen, R.C. Stevens, Crystal structure of a lipid G protein-coupled receptor, Science 335 (2012) 851–855.
[11] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga, T. Kobayashi, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, Nature 482 (2012) 547–551.
[12] H. Wu, D. Wacker, M. Mileni, V. Katritch, G.W. Han, E. Vardy, W. Liu, A.A. Thompson, X.P. Huang, F.I. Carroll, S.W. Mascarella, R.B. Westkaemper, P.D. Mosier, B.L. Roth, V. Cherezov, R.C. Stevens, Structure of the human kappa-opioid receptor in complex with JDTic, Nature 485 (2012) 327–332.
[13] A. Manglik, A.C. Kruse, T.S. Kobilka, F.S. Thian, J.M. Mathiesen, R.K. Sunahara, L. Pardo, W.I. Weis, B.K. Kobilka, S. Granier, Crystal structure of the micro-opioid receptor bound to a morphinan antagonist, Nature 485 (2012) 321–326.
[14] S.G. Rasmussen, B.T. DeVree, Y. Zou, A.C. Kruse, K.Y. Chung, T.S. Kobilka, F.S. Thian, P.S. Chae, E. Pardon, D. Calinski, J.M. Mathiesen, S.T. Shah, J.A. Lyons, M. Caffrey, S.H. Gellman, J. Steyaert, G. Skiniotis, W.I. Weis, R.K. Sunahara, B.K. Kobilka, Crystal structure of the beta2 adrenergic receptor-Gs protein complex, Nature 477 (2011) 549–555.
[15] B.F. O'Dowd, M. Heiber, A. Chan, H.H. Heng, L.C. Tsui, J.L. Kennedy, X. Shi, A. Petronis, S.R. George, T. Nguyen, A human gene that shows identity with the gene encoding the angiotensin receptor is located on chromosome 11, Gene 136 (1993) 355–360.
[16] K. Tatemoto, M. Hosoya, Y. Habata, R. Fujii, T. Kakegawa, M.X. Zou, Y. Kawamata, S. Fukusumi, S. Hinuma, C. Kitada, T. Kurokawa, H. Onda, M. Fujino, Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor, Biochem. Biophys. Res. Commun. 251 (1998) 471–476.
[17] B. Chandrasekaran, O. Dar, T. McDonagh, The role of apelin in cardiovascular function and heart failure, Eur. J. Heart Fail. 10 (2008) 725–732.
[18] J. Boucher, B. Masri, D. Daviaud, S. Gesta, C. Guigne, A. Mazzucotelli, I. Castan-Laurell, I. Tack, B. Knibiehler, C. Carpene, Y. Audigier, J.S. Saulnier-Blache, P. Valet, Apelin, a newly identified adipokine up-regulated by insulin and obesity, Endocrinology 146 (2005) 1764–1771.
[19] H. Choe, M. Farzan, M. Konkel, K. Martin, Y. Sun, L. Marcon, M. Cayabyab, M. Berman, M.E. Dorf, N. Gerard, C. Gerard, J. Sodroski, The orphan seven-transmembrane receptor APJ supports the entry of primary T-cell-line-tropic and dualtropic human immunodeficiency virus type 1, J. Virol. 72 (1998) 6113–6118.
[20] N.J. Macaluso, R.C. Glen, Exploring the 'RPRL' motif of apelin-13 through molecular simulation and biological evaluation of cyclic peptide analogues, ChemMedChem 5 (2010) 1247–1253.
[21] N.J. Macaluso, S.L. Pitkin, J.J. Maguire, A.P. Davenport, R.C. Glen, Discovery of a competitive apelin receptor (APJ) antagonist, ChemMedChem 6 (2011) 1017–1023.
[22] A. Murza, A. Parent, E. Besserer-Offroy, H. Tremblay, F. Karadereye, N. Beaudet, R. Leduc, P. Sarret, E. Marsault, Elucidation of the structure-activity relationships of apelin: influence of unnatural amino acids on binding, signaling, and plasma stability, ChemMedChem 7 (2012) 318–325.
[23] D.N. Langelaan, E.M. Bebbington, T. Reddy, J.K. Rainey, Structural insight into G-protein coupled receptor binding by apelin, Biochemistry 48 (2009) 537–548.
[24] D.N. Langelaan, J.K. Rainey, Headgroup-dependent membrane catalysis of apelin-receptor interactions is likely, J. Phys. Chem. B 113 (2009) 10465–10471.
[25] X. Iturrioz, R. Alvear-Perez, N. De Mota, C. Franchet, F. Guillier, V. Leroux, H. Dabire, M. Le Jouan, H. Chabane, R. Gerbier, D. Bonnet, A. Berdeaux, B. Maigret, J.L. Galzi, M. Hibert, C. Llorens-Cortes, Identification and pharmacological properties of E339-3D6, the first nonpeptidic apelin receptor agonist, FASEB J. 24 (2010) 1506–1517.
[26] N. Zhou, X. Zhang, X. Fan, E. Argyris, J. Fang, E. Acheampong, G.C. DuBois, R.J. Pomerantz, The N-terminal domain of APJ, a CNS-based coreceptor for HIV-1, is essential for its receptor function and coreceptor activity, Virology 317 (2003) 84–94.
[27] X. Iturrioz, R. Gerbier, V. Leroux, R. Alvear-Perez, B. Maigret, C. Llorens-Cortes, By interacting with the C-terminal Phe of apelin, Phe255 and Trp259 in helix VI of the apelin receptor are critical for internalization, J. Biol. Chem. 285 (2010) 32627–32637.
[28] J.A. Ballesteros, H. Weinstein, [19] integrated methods for the construction of three-dimensional models and computational probing of structure–function relations in G protein-coupled receptors, Methods in Neurosciences, vol. 25, Academic Press, 1995, pp. 366–428.
[29] J. Torres, T.J. Stevens, M. Samso, Membrane proteins: the 'Wild West' of structural biology, Trends Biochem. Sci. 28 (2003) 137–144.
[30] D.M. Engelman, Y. Chen, C.N. Chin, A.R. Curran, A.M. Dixon, A.D. Dupuy, A.S. Lee, U. Lehnert, E.E. Matthews, Y.K. Reshetnyak, A. Senes, J.L. Popot, Membrane protein folding: beyond the two stage model, FEBS Lett. 555 (2003) 122–125.
[31] R.E. Jacobs, S.H. White, The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices, Biochemistry 28 (1989) 3421–3437.
[32] J.L. Popot, D.M. Engelman, Membrane protein folding and oligomerization: the two-stage model, Biochemistry 29 (1990) 4031–4037.
[33] J.L. Popot, S.E. Gerchman, D.M. Engelman, Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process, J. Mol. Biol. 198 (1987) 655–676.
[34] N. Bordag, S. Keller, Alpha-helical transmembrane peptides: a "divide and conquer" approach to membrane proteins, Chem. Phys. Lipids 163 (2010) 1–26.

[35] I.G. Tikhonova, S. Costanzi, Unraveling the structure and function of G protein-coupled receptors through NMR spectroscopy, Curr. Pharm. Des. 15 (2009) 4003–4016.
[36] C. Zou, F. Naider, O. Zerbe, Biosynthesis and NMR-studies of a double transmembrane domain from the Y4 receptor, a human GPCR, J. Biomol. NMR 42 (2008) 257–269.
[37] M. Katragadda, A. Chopra, M. Bennett, J.L. Alderfer, P.L. Yeagle, A.D. Albert, Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin, J. Pept. Res. 58 (2001) 79–89.
[38] L.S. Cohen, B. Arshava, R. Estephan, J. Englander, H. Kim, M. Hauser, O. Zerbe, M. Ceruso, J.M. Becker, F. Naider, Expression and biophysical analysis of two double-transmembrane domain-containing fragments from a yeast G protein-coupled receptor, Biopolymers 90 (2008) 117–130.
[39] H. Zheng, J. Zhao, W. Sheng, X.Q. Xie, A transmembrane helix-bundle from G-protein coupled receptor CB2: biosynthesis, purification, and NMR characterization, Biopolymers 83 (2006) 46–61.
[40] A. Neumoin, B. Arshava, J. Becker, O. Zerbe, F. Naider, NMR studies in dodecylphosphocholine of a fragment containing the seventh transmembrane helix of a G-protein-coupled receptor from *Saccharomyces cerevisiae*, Biophys. J. 93 (2007) 467–482.
[41] B. Arshava, I. Taran, H. Xie, J.M. Becker, F. Naider, High resolution NMR analysis of the seven transmembrane domains of a heptahelical receptor in organic-aqueous medium, Biopolymers 64 (2002) 161–176.
[42] R. Estephan, J. Englander, B. Arshava, K.L. Samples, J.M. Becker, F. Naider, Biosynthesis and NMR analysis of a 73-residue domain of a *Saccharomyces cerevisiae* G protein-coupled receptor, Biochemistry 44 (2005) 11795–11810.
[43] L.S. Cohen, B. Arshava, A. Neumoin, J.M. Becker, P. Guntert, O. Zerbe, F. Naider, Comparative NMR analysis of an 80-residue G protein-coupled receptor fragment in two membrane mimetic environments, Biochim. Biophys. Acta 1808 (2011) 2674–2684.
[44] J. Marley, M. Lu, C. Bracken, A method for efficient isotopic labeling of recombinant proteins, J. Biomol. NMR 20 (2001) 71–75.
[45] Q.R. Fan, D.N. Garboczi, C.C. Winter, N. Wagtmann, E.O. Long, D.C. Wiley, Direct binding of a soluble natural killer cell inhibitory receptor to a soluble human leukocyte antigen-Cw4 class I major histocompatibility complex molecule, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7178–7183.
[46] D.N. Langelaan, Structural studies of apelin and its receptor as well as the characteristics and causes of membrane protein helix kinks, PhD thesis, Dalhousie University, Halifax, (2012).
[47] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.
[48] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, Proteins 59 (2005) 687–696.
[49] J. Cavanagh, W.J. Fairbrother, A.G.I. Palmer, N.J. Skelton, M. Rance, Protein NMR Spectroscopy: Principles and Practice, Academic Press, Amsterdam, Boston, 2007.
[50] J. Ding, J.K. Rainey, C. Xu, B.D. Sykes, L. Fliegel, Structural and functional characterization of transmembrane segment VII of the Na+/H+exchanger isoform 1, J. Biol. Chem. 281 (2006) 29817–29829.
[51] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, The XPLOR-NIH NMR molecular structure determination package, J. Magn. Reson. 160 (2003) 65–73.
[52] C.D. Schwieters, J.J. Kuszewski, G. Marius Clore, Using XPLOR-NIH for NMR molecular structure determination, Prog. Nucl. Magn. Reson. Spectrosc. 48 (2006) 47–62.
[53] R.A. Laskowski, J.A. Rullmannn, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486.
[54] J.K. Rainey, L. Fliegel, B.D. Sykes, Strategies for dealing with conformational sampling in structural calculations of flexible or kinked transmembrane peptides, Biochem. Cell Biol. 84 (2006) 918–929.
[55] D.N. Langelaan, M. Wieczorek, C. Blouin, J.K. Rainey, Improved helix and kink characterization in membrane proteins allows evaluation of kink sequence predictors, J. Chem. Inf. Model. 50 (2010) 2213–2220.
[56] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[57] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1656–1676.
[58] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695–1697.
[59] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[60] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[61] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[62] J.M. Baldwin, G.F. Schertler, V.M. Unger, An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors, J. Mol. Biol. 272 (1997) 144–164.
[63] G.G. Krivov, M.V. Shapovalov, R.L. Dunbrack Jr., Improved prediction of protein side-chain conformations with SCWRL4, Proteins 77 (2009) 778–795.
[64] A.I. Alexandrov, M. Mileni, E.Y. Chien, M.A. Hanson, R.C. Stevens, Microscale fluorescent thermal stability assay for membrane proteins, Structure 16 (2008) 351–359.
[65] D. Poger, A.E. Mark, On the validation of molecular dynamics simulations of saturated and cis-monounsaturated phosphatidylcholine lipid bilayers: a comparison with experiment, J. Chem. Theory Comput. 6 (2010) 325–336.
[66] M.G. Wolf, M. Hoefling, C. Aponte-Santamaria, H. Grubmuller, G. Groenhof, g_membed: efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation, J. Comput. Chem. 31 (2010) 2169–2174.
[67] T.E. Creighton, Proteins: Structures and Molecular Properties, W.H. Freeman, New York, 1993.
[68] J.M. Word, S.C. Lovell, J.S. Richardson, D.C. Richardson, Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation, J. Mol. Biol. 285 (1999) 1735–1747.
[69] N. Michaud-Agrawal, E.J. Denning, T.B. Woolf, O. Beckstein, MDAnalysis: a toolkit for the analysis of molecular dynamics simulations, J. Comput. Chem. 32 (2011) 2319–2327.
[70] P. Liu, D.K. Agrafiotis, D.L. Theobald, Fast determination of the optimal rotational matrix for macromolecular superpositions, J. Comput. Chem. 31 (2010) 1561–1563.
[71] D.L. Theobald, Rapid calculation of RMSDs using a quaternion-based characteristic polynomial, Acta Crystallogr. A 61 (2005) 478–480.
[72] H.M. Berman, T. Battistuz, T.N. Bhat, W.F. Bluhm, P.E. Bourne, K. Burkhardt, Z. Feng, G.L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J.D. Westbrook, C. Zardecki, The Protein Data Bank, Acta Crystallogr. D: Biol. Crystallogr. 58 (2002) 899–907.
[73] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R. Kent Wenger, H. Yao, J.L. Markley, BioMagResBank, Nucleic Acids Res. 36 (2008) D402–D408.
[74] K. Wüthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.
[75] D.S. Wishart, B.D. Sykes, F.M. Richards, The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, Biochemistry 31 (1992) 1647–1651.
[76] M.S. Cheung, M.L. Maguire, T.J. Stevens, R.W. Broadhurst, DANGLE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure, J. Magn. Reson. 202 (2010) 223–233.
[77] B. Bai, J. Tang, H. Liu, J. Chen, Y. Li, W. Song, Apelin-13 induces ERK1/2 but not p38 MAPK activation through coupling of the human apelin receptor to the Gi2 pathway, Acta Biochim. Biophys. Sin. (Shanghai) 40 (2008) 311–318.
[78] S.A. Pandit, M.L. Berkowitz, Molecular dynamics simulation of dipalmitoylphosphatidylserine bilayer with Na^+ counterions, Biophys. J. 82 (2002) 1818–1827.
[79] O. Berger, O. Edholm, F. Jahnig, Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature, Biophys. J. 72 (1997) 2002–2013.
[80] M. Patra, M. Karttunen, M.T. Hyvonen, E. Falck, P. Lindqvist, I. Vattulainen, Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions, Biophys. J. 84 (2003) 3636–3645.
[81] C. Anézo, A.H. de Vries, H. Höltje, D.P. Tieleman, S. Marrink, Methodological issues in lipid bilayer simulations, J. Phys. Chem. B 107 (2003) 9424–9433.
[82] M. Patra, M. Karttunen, M.T. Hyvönen, E. Falck, I. Vattulainen, Lipid bilayers driven to a wrong lane in molecular dynamics simulations by subtle changes in long-range electrostatic interactions, J. Phys. Chem. B 108 (2004) 4485–4494.
[83] S. Leekumjorn, A.K. Sum, Molecular simulation study of structural and dynamic properties of mixed DPPC/DPPE bilayers, Biophys. J. 90 (2006) 3951–3965.
[84] A.K. Sum, J.J. de Pablo, Molecular simulation study on the influence of dimethylsulfoxide on the structure of phospholipid bilayers, Biophys. J. 85 (2003) 3636–3645.
[85] S. Yohannan, S. Faham, D. Yang, J.P. Whitelegge, J.U. Bowie, The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 959–963.
[86] T. Mirzadegan, G. Benko, S. Filipek, K. Palczewski, Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin, Biochemistry 42 (2003) 2759–2767.
[87] M.M. Javadpour, M. Eilers, M. Groesbeek, S.O. Smith, Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association, Biophys. J. 77 (1999) 1609–1618.
[88] J.N. Bright, M.S.P. Sansom, The flexing/twirling helix: exploring the flexibility about molecular hinges formed by proline and glycine motifs in transmembrane helices, J. Phys. Chem. B 107 (2003) 627–636.
[89] C. Zou, S. Kumaran, S. Markovic, R. Walser, O. Zerbe, Studies of the structure of the N-terminal domain from the Y4 receptor – a G protein-coupled receptor – and its interaction with hormones from the NPY family, ChemBioChem 9 (2008) 2276–2284.
[90] J.P. D'Aoust, M. Tiberi, Role of the extracellular amino terminus and first membrane-spanning helix of dopamine D1 and D5 receptors in shaping ligand selectivity and efficacy, Cell. Signal. 22 (2010) 106–116.
[91] C. Parthier, M. Kleinschmidt, P. Neumann, R. Rudolph, S. Manhart, D. Schlenzig, J. Fanghanel, J.U. Rahfeld, H.U. Demuth, M.T. Stubbs, Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 13942–13947.
[92] C.R. Grace, M.H. Perrin, M.R. DiGruccio, C.L. Miller, J.E. Rivier, W.W. Vale, R. Riek, NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 12836–12841.
[93] S.R. Hoare, Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors, Drug Discov. Today 10 (2005) 417–427.
[94] Y. Kofuku, C. Yoshiura, T. Ueda, H. Terasawa, T. Hirai, S. Tominaga, M. Hirose, Y. Maeda, H. Takahashi, Y. Terashima, K. Matsushima, I. Shimada, Structural basis of the interaction between chemokine stromal cell-derived factor-1/CXCL12 and its G-protein-coupled receptor CXCR4, J. Biol. Chem. 284 (2009) 35240–35250.
[95] B.A. Shoemaker, J.J. Portman, P.G. Wolynes, Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 8868–8873.
[96] D.N. Langelaan, J.K. Rainey, Membrane catalysis of peptide-receptor binding, Biochem. Cell Biol. 88 (2010) 203–210.
[97] H. Inooka, T. Ohtaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, M. Shirakawa, Conformation of a peptide ligand bound to its G-protein coupled receptor, Nat. Struct. Biol. 8 (2001) 161–165.