The G protein-coupled receptor N-terminus and receptor signalling: N-tering a new era

James L.J. Coleman, Tony Ngo, Nicola J. Smith

PII: S0898-6568(17)30045-1
DOI: doi: 10.1016/j.cellsig.2017.02.004
Reference: CLS 8845
To appear in: Cellular Signalling
Received date: 1 February 2017
Accepted date: 6 February 2017

Please cite this article as: James L.J. Coleman, Tony Ngo, Nicola J. Smith, The G protein-coupled receptor N-terminus and receptor signalling: N-tering a new era. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. CLs(2016), doi: 10.1016/j.cellsig.2017.02.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

COPYRIGHT STATEMENT. This is the author's version of a work that was accepted for publication. Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this work. For a definitive version of this work please refer to the published source.


Final publication is available from http://dx.doi.org/10.1016/j.cellsig.2017.02.004

This article may be used for non-commercial purposes in accordance with the Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 International licence http://creativecommons.org/licenses/by-nc-nd/4.0/
Title: The G Protein-Coupled Receptor N-terminus and Receptor Signalling: N-tering a new era

Authors: James L. J. Coleman*1,2, Tony Ngo1,2, Nicola J. Smith*1,2

*To whom correspondence should be addressed: Dr Nicola J. Smith (n.smith@victorchang.edu.au) or Mr. James L. J. Coleman (j.coleman@victorchang.edu.au).

Affiliations:

1. Molecular Pharmacology Group, Victor Chang Cardiac Research Institute, Darlinghurst, NSW, 2010, Australia

2. St Vincent’s Clinical School, University of New South Wales, Darlinghurst, NSW, 2010, Australia

Abstract: G protein-coupled receptors (GPCRs) are a vast family of membrane-traversing proteins, essential to the ability of eukaryotic life to detect, and mount an intracellular response to, a diverse range of extracellular stimuli. GPCRs have evolved with archetypal features including an extracellular N-terminus and intracellular C-terminus that flank a transmembrane structure of seven sequential helices joined by intracellular and extracellular loops. These structural domains contribute to the ability of a GPCR to be correctly synthesised and inserted into the cell membrane, to interact with its cognate ligand(s) and to couple with signal-transducing heterotrimeric G proteins, allowing the activated receptor to selectively modulate a number of signalling cascades. While well known for its importance in receptor translation and trafficking, the GPCR N-terminus is underexplored as a participant in receptor signalling. This review aims to discuss and integrate recent advances in knowledge of the vital roles of the GPCR N-terminus in receptor signalling.

Key words: GPCR, G protein-coupled receptor, N-terminus, proteolysis, signalling
Acknowledgements: This work was supported by research fellowships from the National Health and Medical Research Council of Australia and National Heart Foundation of Australia (NJS), a St Vincent’s Clinic Foundation grant-in-aid (NJS), the Mostyn Family Foundation (NJS) and Australian Postgraduate Scholarships to JLJC and TN. JLJC is additionally supported by a Simon and Michal Wilkenfeld Scholarship.
1. Introduction

G protein-coupled receptors (GPCRs) are paramount to the ability of eukaryotic organisms to detect and respond to stimuli in their extracellular environment; a property that perhaps made GPCRs integral to the evolution and diversification of multicellular eukaryotic life itself [1]. Each GPCR has evolved remarkable specificity towards biological detection of discrete stimuli including photons, ions, odorants, nucleotides, amino acids, peptides and proteins [2], association with which biases the membrane-spanning GPCR towards an active conformation, facilitating interaction with heterotrimeric G proteins which transduce the once extracellular signal along a number of potential intracellular signalling pathways [3]. Crucial to the ability of GPCRs to serve as a nexus between extracellular and intracellular signals is their membrane-traversing structure; a GPCR consists of an extracellular N-terminus, seven serial transmembrane helices joined by intracellular and extracellular loops, and an intracellular C-terminus [3]. Whilst we refer the reader to a number of comprehensive reviews of the relationship between the GPCR structural domains and receptor function [3, 4], the GPCR N-terminus is an often overlooked and ever-evolving area of research, with increasing evidence that it is truly a dynamic participant in the exquisitely regulated mechanisms of GPCR signalling. This review will briefly introduce the traditional paradigm of N-terminus function before exploring more recent knowledge of N-terminus-mediated GPCR signalling, including both protease-modulated and –independent processes, primarily in the context of the rhodopsin-like (class A) receptors.

2. Traditional roles of the G protein-coupled receptor N-terminus

The importance of the GPCR N-terminus is typically understood in the context of receptor trafficking, from the endoplasmic reticulum (ER) and golgi apparatus [5], or as being involved in ligand binding for secretin-like (class B) [6] and glutamate-like (class C) GPCRs [7]. Regarding trafficking, the GPCR N-terminus can be subject to post-translational modification in the form of N-linked glycosylation occurring at the motif NxS/T, proving important for correct integration of a number of GPCRs into the
cell membrane [5]. Notable examples include the β₂-adrenoceptors [8], vasoactive intestinal peptide 1 receptor [9], angiotensin II type 1 (AT₁) receptor [10], the melanin-concentrating hormone receptor 1 [11] and the melanocortin 2 receptor [12] - all displaying reduced surface expression in response to site-directed mutagenesis of N-terminal, N-linked glycosylation motifs.

Another means by which the GPCR N-terminus may influence surface expression is by possessing a hydrophobic signal peptide sequence; present on 5-10% of all GPCRs and typically essential for correct protein translation and export from the ER for these receptors [13]. GPCRs without an N-terminal signal peptide are thought to rely on uncleaved signal anchor sequences, commonly located within the first transmembrane domain, that are recognized by the signal recognition particle (SRP) following synthesis by cytosolic ribosomes, leading to a temporary arrest of translation. This ultimately mediates transfer of the ribosome and nascent protein to the ER translocon complex, where translation resumes, and the GPCR is co-translationally inserted into the ER membrane. For these receptors, the N-terminus is synthesised in the cytoplasm before the signal anchor sequence is synthesised and recognized, so must be post-translationally translocated through the ER membrane [13]. This represents a challenge for GPCRs with longer N-termini or those with a high abundance of positively charged amino acids (Lys and Arg) in their N-termini and, concordantly, GPCRs with these properties more commonly possess an N-terminal hydrophobic signal peptide, such that the N-terminus can be translocated into the ER as it is synthesised [14]. Similarly, the presence of stably folded domains in the N-terminus may prohibit post-translational translocation, hence the postulated role of the N-terminal signal peptide for the endothelin B (ET₉) receptor [15]. These signal peptides are typically cleaved off by the signal peptidases of the ER following translation, but not in all cases. The corticotropin-releasing factor receptor 2 is such an example, possessing an uncleaved signal peptide that reduces receptor expression at the plasma membrane and inhibits receptor dimerisation (as the signal peptide carries bulky N-linked glycans) [16].
In contrast, the predicted signal peptide of the α_{2C}-adrenoceptor has been reported to remain uncleaved, resulting in an increase in binding-competent receptor in the cell membrane [17].

In addition to these roles in receptor surface expression and trafficking, the N-terminus also regulates ligand-binding for many GPCRs, though primarily outside of the rhodopsin-like class. For peptide ligands of the secretin-like GPCRs, the ligand first binds to the N-terminus before interacting with the receptor’s transmembrane domain, exemplified by parathyroid hormone and its cognate receptor [18]. The glutamate-like receptors show yet another distinct mode of ligand binding, as seen in the metabotropic glutamate receptors, for which ligand binding takes place solely within the N-terminus, with these receptors having so-called ‘venus fly trap’ domains on their extensive N-termini [19].

Despite these well characterised roles for N-termini, ongoing research is elucidating new ways in which the N-terminus can directly contribute to the signalling capacity of GPCRs. For the purposes of this review, GPCRs with signalling activity that can be attributed to the presence of the receptor’s N-terminus (or defined regions within) have been categorised as being either protease-modulated or protease-independent and these classifications will now be explored further.

3. The N-terminus and protease-modulated GPCR signalling

Perhaps no better example of protease-modulated GPCR signalling exists than the protease activated receptors (PARs) (Table 1). PARs contain their own ligand within their N-terminus but this ‘tethered ligand’ is only accessible and able to activate the receptor upon N-terminal proteolysis immediately upstream of the ligand sequence [20] (Figure 1 A). The first discovered PAR was PAR1, reported in 1991 to be the receptor for the serine protease thrombin [21, 22], which until that point was considered primarily as a catalyst for the formation of fibrin in blood clotting. Vu, et al. [21] found that microinjection of mRNA from thrombin-responsive Dami cells conferred thrombin-induced Ca^{2+} release to Xenopus oocytes. Using size fractionation of mRNA and size-selected cDNA library synthesis, the
thrombin receptor was subsequently cloned, with the receptor cRNA then shown to induced a 100-fold greater thrombin-induced Ca\(^{2+}\) response in Xenopus oocytes than the unenriched Dami cell cRNA [21]. Further characterization of this thrombin-receptor cDNA identified the encoded protein to be 425 amino acids long, with an N-terminus of approximately 75 amino acids that intriguingly contained the motif LDPRS – noted as strongly resembling the thrombin-cleavage site within zymogen protein C (LDPRI) and suggesting thrombin can cleave this newly identified receptor between Arg41 and Ser42. Indeed, mutation of Arg41 to alanine abolished receptor-induced Ca\(^{2+}\) release in Xenopus oocytes, yet elegantly, signalling was restored to wild-type levels following addition of an exogenous peptide mimicking the region of the N-terminus distal to the hypothesised cleavage site (SFLLRNPNKDYEPF) [21]. It was shown soon after that only the first seven of these residues are necessary for the agonist activity of the peptide [23].

Following on from the discovery of a novel N-terminal-dependent mechanism of GPCR activation, several more protease-activated receptors were discovered. Trypsin was shown to cause activation of a receptor identified in a mouse genomic library as having 28% amino acid similarity to mouse PAR1, subsequently named PAR2, with a trypsin-cleavage site identified on the N-terminus between Arg34 and Ser35 [24]. A similar Ca\(^{2+}\) efflux assay was used in Xenopus oocytes to show that cRNA encoding a PAR2 receptor with a mutation to interrupt the putative trypsin cleavage site was functionally silent, yet addition of a peptide corresponding to the six amino acids distal to the site (SLIGRL) restored signalling [24]. Discovery of PAR3 [25] and PAR4 [26] followed suit, both being substrates for N-terminal proteolysis by thrombin, similar to PAR1. Whilst discussed in detail elsewhere [20, 27], thrombin and trypsin are not the only enzymes that may activate the PARs, with a plethora of other proteases shown to cleave PARs at their canonical sites for activation. Of these other proteases, metalloproteases (including the matrix metalloproteases [MMPs] and a disintegrin and metalloproteases [ADAMs]) warrant specific mention, and their role in the regulation of a number of other GPCRs will be discussed
later in this review. Importantly, the discovery of the PARs challenged traditional paradigms of GPCR signalling and demonstrated new roles for proteases in physiology.

Whilst PARs become activated following protease-mediated unveiling of a tethered agonist on the GPCR N-terminus, the thyroid-stimulating hormone (TSH) receptor exists as an example of a GPCR that may be activated by protease-mediated removal of an inhibitory N-terminal domain, without implicating a tethered agonist. The TSH receptor consists of two subunits [28], shown to result from a metalloprotease-mediated cleavage between the membrane-spanning domain and the large N-terminal extracellular domain of the receptor, which occurs at the cell surface [29]. The propensity of an HA-tagged version of the TSH receptor to constitutively stimulate cAMP production was increased up to sevenfold by deletion of the extracellular N-terminal domain, as measured via radioimmunoassay in COS cells [30]. Whilst working with untagged receptors is preferred where possible, the use of the HA tag rescued low expression typically seen when studying truncated glycoprotein hormone receptors, and allowed accurate normalization of receptor activity levels to absolute differences in expression levels (by use of detection with HA-specific antibodies). In this manner it was observed that the truncated TSH receptor had greater constitutive activity relative to the wild-type TSH receptor, even before adjusting for the higher expression of the wild-type receptor. Hypothesising that the remaining 10 amino acid region (409-418) of the N-terminus may be acting as a tethered ligand following removal of the extracellular domain (ECD), Zhang, et al. [30] used a peptide corresponding to that region in an attempt to stimulate the receptor in CHO-JP09 cells, but reported no effect even at concentrations as high as 0.79 mM. The authors concluded that the N-terminal domain of TSH receptor may interact with the transmembrane domain to cause intra-molecular inhibition of constitutive signalling however, whether the peptide was screened against the full-length or truncated receptor was not reported. Similarly, Vlaeminck-Guillem, et al. [31] found a rhodopsin-tagged (RT) TSH receptor with its ectodomain removed to be sixfold more effective at constitutive cAMP generation than a RT-wild-type TSH receptor, using a
comparable radioimmunoassay in COS cells. Moreover, this enhanced constitutive activity was not of the magnitude achieved with a TSH-stimulated wild-type TSH receptor (indicating that a signal maximum had not been reached), yet TSH could not further stimulate the truncated receptor. This suggests the N-terminal domain structurally inhibits TSH receptor constitutive signalling until liberated by metalloprotease activity or altered by interaction with TSH, allowing the ectodomain to adopt a conformation capable of maximally stimulating the transmembrane domain of the receptor [31]. A Disintegrin and Metalloprotease 10 (ADAM10) was subsequently implicated as a possible mediator of this process [32], representing a novel example of interaction between the ADAM enzymes and GPCR signalling. Prior to this, crosstalk between ADAMs and GPCRs was known only in the context of G protein-dependent activation of ADAMs and subsequent surface shedding of extracellular growth factor receptor (EGFR) ligands [33-35]

In contrast to the TSH receptor and PARs, GPCRs that both exhibit enhanced signalling following N-terminal proteolysis, our laboratory has recently reported a GPCR to be inactivated by this process (Figure 1 B) (Table 1) [36]. GPR37L1, an orphan GPCR without a validated pairing with a cognate ligand, is abundant in the central nervous system [37] where it has been shown to regulate cerebellar development and motor learning [38]. Following observation of fragments of the GPR37L1 N-terminus in human cerebrospinal fluid, we stably expressed GPR37L1 (with a C-terminal eYFP tag) in HEK293 cells, and using Western blotting, discovered two isoforms of the receptor: full-length GPR37L1 and a variant missing the entire N-terminus. Using broad spectrum inhibition of ADAM/MMP enzymes in culture, we prevented the appearance of the truncated receptor and increased the abundance of full-length GPR37L1-eYFP. The failure of the four human tissue inhibitors of metalloprotease (TIMPs) to prevent receptor cleavage in vitro suggested the protease responsible to be an ADAM; yet another manner in which this class of protease may interact with GPCRs. In investigating the consequences of N-terminal proteolysis for GPR37L1, we observed that untagged GPR37L1 constitutively couples to Gα, G proteins in
both yeast and mammalian reporter assays, but interestingly, genetic deletion of the N-terminus created a functionally silent receptor (despite confirming translation and expression of the receptor)[36]. Crucially, inhibition of MMPs/ADAMs was sufficient to cause a concordant increase in constitutive Gαs signalling, as full-length receptor was preserved. We were also able to detect these two GPR37L1 isoforms in mouse and rat cerebellum where, most intriguingly, cleaved GPR37L1 was found to be the predominant form, suggesting metalloprotease enzymes tightly regulate the silencing of a constitutively active receptor [36].

A similar protease-mediated inactivation method has been reported for the ETb receptor. When isolated from human placenta by affinity chromatography, biotinylated ETb receptor was found to exist as two polypeptides of different sizes [39]. Amino acid sequence analysis revealed a larger species consisting of the ETb receptor minus 26 residues from the start of the N-terminus (consistent with removal of the receptor’s signal peptide) and a smaller species missing 64 residues from the start of the N-terminus [39]. A similar phenomenon was observed when the human ETb receptor was purified from baculovirus-infected Sf9 insect cells but in this expression system, the smaller species was reported as missing only 63 residues [40]. Whilst Akiyama, et al. [39] found exogenous endopeptidase Arg-C could cleave the larger ETb receptor species at Arg64-Ser65, the work of Grantcharova, et al. [41] implicated a metalloprotease as the true mediator of ETb N-terminal proteolysis. EDTA and other broad-spectrum metalloprotease inhibitors were used to inhibit the N-terminal proteolysis of GFP-tagged ETb receptor (ETb receptor-GFP) in HEK293 cells. The cleavage was found to be dependent on binding of the endogenous agonist endothelin-1 (ET-1) [41], which perturbed biphasic extracellular signal-regulated kinase 1 and 2 (ERK1/2) signalling following receptor activation. Levels of total and phosphorylated ERK1/2 from lysates of HEK293 cells stably transfected to express wild-type ETb receptor-GFP or Δ2-64-
ET\textsubscript{B} receptor-GFP (a mutant analogous to the cleaved receptor) were assessed by immunoblot following a time course of treatment with ET-1 [42]. ET\textsubscript{B} receptor-GFP activation caused a peak of ERK1/2 phosphorylation 3-5 min after addition of ET-1, declining within 30-60 min, before a second response phase occurring 80-180 min after ET-1 addition. Whilst the first phase of ERK signalling was preserved following ET-1 stimulation of the Δ2-64-ET\textsubscript{B} receptor-GFP (although dampened), the second phase of phosphorylation was absent [42]. The authors also showed that this biphasic to monophasic switch in ERK1/2 signalling occurred in vascular smooth muscle cells (VSMCs) transiently transfected with the same receptor constructs and stimulated with the selective ET\textsubscript{B} receptor agonist IRL1620 [43]. Whilst one would expect metalloprotease inhibition to enhance the second phase of ERK phosphorylation, metalloprotease inhibition was found to abolish the second ERK1/2 phosphorylation stage.

Grantcharova, et al. [43] then discovered that IRL1620-stimulation of the wild-type ET\textsubscript{B} receptor-transfected VSMCs caused an increase in VSMC secretion of MMP2 – a metalloprotease previously implicated in shedding of heparin-binding EGF and subsequent transactivation of EGFR [44]. Indeed, MMP-2 siRNA, EGFR tyrosine kinase inhibition (AG 1478) and α-EGFR antibody treatment all abolished the second phase of ERK phosphorylation in response to IRL1620 [44]. Immunoblotting also showed that 120 min stimulation of wild-type ET\textsubscript{B} receptor-GFP with IRL1620 caused phosphorylation of EGFR, coinciding with the second phase of ERK phosphorylation, but this response was absent for the Δ2-64-ET\textsubscript{B} receptor-GFP. Although it should be acknowledged that truncation of ET\textsubscript{B}-GFP causes drastic decreases in cell surface expression and potentially confounds studies of the truncated mutant [41], these studies raise new possibilities about diverse modes of GPCR signalling. Notwithstanding the well-established role for N-terminal cleavage in activating PARs, it is fascinating to reflect on N-terminal proteolysis triggering GPCRs to change signalling modalities.

Whilst yet to be described in terms of functional consequences for the receptor, N-terminal proteolysis of the AT\textsubscript{1} receptor has been reported by one group as part of a greater series of proteolytic events that
follow receptor stimulation. Using \( \text{AT}_1 \) receptors tagged with an N-terminal ECFP and a C-terminal EYFP and expressed in COS-7 cells, Cook, et al. [45] observed separation of the fluorescent tags following 30 min treatment with angiotensin II, resulting in nuclear accumulation of the C-terminal \( \text{AT}_1 \) receptor fragment. Immunoblotting revealed this C-terminal fragment to be the size of the EYFP (\(~30\) kDa) plus the \( \text{AT}_1 \) receptor C-terminus (\(~7\) kDa). Metalloproteases were implicated in this process, as metal-chelating agents EDTA and 1,10-ortho-phenanthroline monohydrate inhibited the nuclear accumulation of the C-terminus. Intriguingly, it was observed during immunoblotting experiments that no cleavage product consistent with the size of ECFP and the \( \text{AT}_1 \) receptor (minus its carboxy terminus) could be detected, suggesting a second proteolytic event was cleaving the N-terminal ECFP moiety from the receptor [45]. Accordingly, the authors performed immunoblot experiments on conditioned culture media from \( \text{AT}_1 \) receptor-transfected CHO-K1 cells and detected a \(~33\) kDa protein – the expected size of the ECFP epitope (\(~30\) kDa) joined to the \( \text{AT}_1 \) receptor N-terminus (\(~3\) kDa). Treatment of the cells with angiotensin II increased the amount of \( \text{AT}_1 \) receptor N-terminus recovered from the media. Whilst a precise role for N-terminal proteolysis of the \( \text{AT}_1 \) receptor is yet to be elucidated, the agonist-enhanced nature of the phenomenon suggests it may be involved in receptor degradation following C-terminal proteolysis, or more speculatively, could permit conformational changes necessary for metalloprotease-mediated C-terminal proteolysis.

Aligning with a hypothesised role of \( \text{AT}_1 \) receptor N-terminal proteolysis in receptor degradation, a number of GPCRs have been shown to be subject to N-terminal proteolysis in a manner that regulates receptor turnover and cell surface expression and, thus, receptor signalling (Table 1). Hakalahti, et al. [46] described metalloprotease-mediated N-terminal proteolysis occurring at the human \( \beta_1 \)-adrenoceptor. This phenomenon was found to be ligand-enhanced and hypothesised by the authors to attenuate signalling and to initiate receptor turnover, although this was not directly demonstrated. The aforementioned ET\(_b\) receptor was found to undergo a similar phenomenon. Grantcharova, et al. [41]
report that a mutant receptor analogous to the N-terminally truncated ET₉, whilst equally responsive to ET-1, had 15-fold lower cell surface expression than the full length ET₉, yet this may have been an artefact of the mutant lacking a cleavable signal peptide and should be kept in mind when considering the reported role of ET₉ N-terminal cleavage and ERK phosphorylation, as previously discussed.

These examples of the β₁-adrenoceptor and ET₉ receptor contrast that of the α₁D-adrenoceptor, for which proteolytic removal of the first 91 residues of the receptor was found to be necessary for proper trafficking to the cell surface (Table 1). Plasma membrane expression in HEK293 cells was enhanced in a SNAP-tagged ∆1-91 mutant α₁D-adrenoceptor, as compared to a SNAP-tagged wild-type receptor, causing decreased receptor ER retention, facilitating correct surface expression and explaining increased agonist efficacy that the authors observed with an untagged ∆1-91 α₁D-adrenoceptor [47]. A similar result was observed for the orphan receptor GPR37, whereby genetic removal of the N-terminal-most 210 residues from a FLAG-tagged GPR37 greatly enhanced surface expression in HEK293 cells, compared to a FLAG-tagged wild-type GPR37 [48]. However, like the phylogenetic-related GPR37L1 [36], GPR37 has recently also been shown to undergo efficient metalloprotease-mediated N-terminal cleavage at the cell surface in HEK293 cells [49]. Mattila, et al. [49] note that metalloprotease-mediated N-terminal proteolysis would also remove an N-terminal tag if it were placed prior to the cleavage site (residues 167-168), potentially confounding studies of N-terminally tagged GPR37. The effect of N-terminal proteolysis on GPR37 signalling was not investigated.

4. The N-terminus and protease-independent signalling
Distinct from the preceding examples are those GPCRs for which the N-terminus modulates signalling in a protease-independent manner (Table 1). This is well illustrated by the case of the melanocortin-4 (MC₄) receptor where a disproportionate number of disease-causing single-point mutations have been located within the receptor’s N-terminus. Srinivasan, et al. [50] investigated a series of obesity-
associated N-terminal MC₄ receptor mutations using a cAMP signalling assay in HEK293 cells. Unusually, the mutants all responded to α-melanocyte stimulating hormone (αMSH) similarly to wild-type MC₄ receptor, yet all had lower constitutive activity. To further understand the importance of the MC₄ receptor N-terminus, Srinivasan, et al. [50] removed the N-terminus, and added a signal peptide to enhance membrane localisation, generating a construct termed ATG24-MC4R. This construct mirrored the behaviour of the point mutants, displaying unperturbed responsiveness to αMSH but abolished basal activity, as compared to an equivalently tagged wild-type MC₄ receptor. The necessity of the N-terminus in conferring the MC₄ receptor with constitutive activity was confirmed by the observed restoration of wild-type cAMP signalling to the ATG24-MC4R by co-expression of the wild-type MC₄ receptor N-terminus on a CD8 transmembrane domain [50]. Subsequently, the same group reported that a six amino acid peptide, corresponding to a region of the MC₄ N-terminus with high conservation between species, was sufficient to restore constitutive activity to the N-terminus-lacking mutant MC₄ receptor [51]. This elegant work demonstrates a different mechanism from that of the PARs, in that no prior proteolysis is required for these regions in the wild-type receptor to confer the MC₄ receptor with constitutive activity. Furthermore, this mechanism explains the fascinating observation that agouti-related peptide functions as an endogenous MC₄ receptor inverse agonist [52, 53], as it antagonizes the receptor’s own N-terminus [51] (Figure 1C). Remarkably, N-terminal-mediated agonism and αMSH ligand driven agonism are mechanistically distinct, with the former perhaps being the first to evolve, due to its dependency on a well conserved class A GPCR aromatic cluster (Phe201⁵.47-Trp258⁶.48-Phe262⁶.52) found within transmembrane helices 5 and 6 [3].

A similar tethered agonist model has been reported for central nervous system orphan receptor GPR61. In order to explore the role of the N-terminus of GPR61, Toyooka, et al. [54] expressed mutants of GPR61 in Sf9 cells. All were fused to a Gα₅ G protein (the reported cognate G protein for GPR61) and with the most N-terminal 10, 15, 20 or 25 amino acids deleted. Whilst mutants missing up to 20 amino
acids had unchanged expression and constitutive activity relative to full-length GPR61-Gs, removing 25 amino acids from the N-terminus (GPR61del25-Gs) resulted in decreased receptor presence in crude membrane preparations (indicative of lower surface expression) and also decreased constitutive Gαs activity as assessed by [35S]GTPγS radioassay. Note that in this study, loss of constitutive activity was defined as a loss of responsiveness to reported GPR61 inverse agonist 5-(nonyloxy)tryptamine (5-NT). Additionally, co-expression of a CD8 helix fused with the first 48 amino acids of GPR61, alongside GPR61del25-Gs, restored constitutive activity. Toyooka, et al. [54] then scanned residues 18-25 of the N-terminus with single-point alanine mutations to find that a Val19Ala GPR61 mutation was sufficient to recapitulate the decreased expression and loss of constitutive activity seen for the GPR61del25-Gs mutant. All other mutants were comparable to wild-type. This is inconsistent with the authors’ observation that GPR61del20 did not differ from wild-type, because that mutant would also lack the supposedly critical Val19 residue. Nevertheless, evidence supports the presence of tethered agonist domains within the N-termini of GPR61 and MC4 receptor contributing to receptor constitutive activity, independent of proteolysis, demonstrating the diverse ways by which the N-terminus can mediate signalling. This example, and that of the MC4 receptor tethered-agonist making possible the evolution of agouti-related peptide as an endogenous inverse agonist, serve to challenge and further our understanding of the GPCR N-terminus.

Conversely, regions of the N-terminus may be involved in negative regulation of GPCR signalling (Table 1). The N-terminus of GPR83 harbours a domain that may constrain receptor activity, as demonstrated by Muller, et al. [55]. A mutant murine GPR83 construct missing its entire N-terminus (yet containing its signal peptide and an HA tag downstream), termed del18-65, was found to be expressed at the cell surface of COS-7 cells at 19% the abundance of the full-length GPR83, yet had a fourfold increase in Gαq/11 activity of the full-length GPR83, as measured by reporter gene assays. Similarly, a shorter N-terminal deletion (del 36-65) reduced surface expression to 72% of the full-length GPR83 but possessed
1.4 times the $\alpha_{q/11}$ activity of the full-length GPR83 [55]. More recently, the same group has shown that similarly-tagged human GPR83 had approximately 150% the surface expression of the mouse orthologue and that deletion of the N-terminus from the human GPR83 (del17-60) was deleterious to surface expression, just as previously demonstrated for the mouse GPR83 [56]. The human GPR83 del17-60 mutant also mirrored the enhanced $\alpha_{q/11}$ reporter gene activity seen in the mouse counterpart. Hypothesising that this phenomena was due to distinct well-conserved or previously described residues in the GPR83 N-terminus, discrete residues were mutated to alanine and assessed in the same assays. Whilst the alanine mutations did not alter surface expression, Tyr61Ala, Glu64Ala, Ser65Ala and Gln66Ala showed enhanced $\alpha_{q/11}$ activity (amongst other enhanced signalling effects), implicating the side chains of these specific residues in negative regulation of signalling [56].

In another example of the suppressive effects of the N-terminus on GPCR activity, deletion of a region comprising residues 2-10 of the yeast pheromone $\alpha$-factor receptor Ste2p, has been shown to relieve an inhibitory influence on downstream pheromone-responsive promoter FUS1, resulting in increased FUS1-lacZ induction in response to increasing concentrations of $\alpha$-factor [57]. Yeast expressing the Ste2pΔ2-10 mutant also had more than sixfold increase in mating efficiency, compared to yeast with the wild-type Ste2p. Crucially, this phenomenon was not the result of altered receptor expression or $\alpha$-factor binding affinity, with Ste2pΔ2-10 not significantly differing from wild-type Ste2p in these parameters [57]. The authors hypothesise that the N-terminus may influence the conformation of other receptor domains that are important for activation.

Similarly, the human 5-HT$_{2B}$ receptor N-terminus has a role in constraining agonist-induced signalling (Figure 1 D). Belmer, et al. [58] expressed a 5-HT$_{2B}$ receptor containing two nonsynonymous N-terminal polymorphisms (Arg6Gly;Glu42Gly, previously identified within a drug-abusing population), in COS-7 cells and found an increase in affinity for radioligand [$^3$H]-5-HT binding and an increase in intracellular
inositol phosphate accumulation in response to partial agonist 2,5-dimethoxy-4-iodoamphetamine hydrochloride (DOI), compared to the wild-type receptor. The Arg6Gly;Glu42Gly, variant also displayed increased constitutive inositol phosphate accumulation. Moreover, removal of the 5-HT<sub>2B</sub> receptor N-terminus (ΔNter) resulted in further increases in constitutive activity, agonist affinity and agonist-stimulated activity. Importantly, co-expression of the receptor N-terminus alongside the ΔNter 5-HT<sub>2B</sub> restored the inositol phosphate response induced by DOI to that of the wild-type receptor, elegantly demonstrating the role of the N-terminus in ‘tuning’ agonist-mediated activation for this receptor [58].

5. Structural perspectives on the GPCR N-terminus
While examples of GPCRs with N-terminal-mediated signalling are being continually discovered, this paradigm of research is still in its infancy, thus representing an opportunity for the discovery of novel GPCR structural elements. Yet, crystallography of GPCRs is not trivial and is subject to many technical hurdles including the intrinsic flexibility within various GPCR domains, particularly non-transmembrane regions, that hinders their stabilisation, and the large quantities of functional receptor necessary for crystallisation [59]. Testament to the challenges of protein stability is the total or partial absence of the N-terminus from the majority of published GPCR crystal structures. However, a number of published rhodopsin-like GPCR structures are exceptions to this.

The structure of the human cannabinoid 1 (CB<sub>1</sub>) receptor in complex with the synthetic antagonist AM6538 was recently solved by Hua, et al. [60] and revealed an exciting potential role for the receptor’s N-terminus in ligand recognition - a phenomena unusual in rhodopsin-like GPCRs (Figure 2 A). Although truncated by 98 amino acids, the remaining CB<sub>1</sub> receptor N-terminus formed a V-shaped loop that capped the AM6538-bound pocket, with AM6538 forming hydrophobic interactions with not only all transmembrane helices (except IV), but also with extracellular loop (ECL) 2 and, excitingly, the N-terminus. Whilst Hua, et al. [60] note that crystal packing interactions may influence the N-terminus
conformation, the ordered form of the CB₁ N terminus is not unprecedented. The lysophosphatidic acid receptor 1 (LPA₁) (modified with a thermostabilised apocytochrome b₅₆₂RIL inserted into ICL3, a 38 amino acid C-terminal truncation and engineered disulfide bonds) was crystallised in complex with synthetic ligand ONO-9780307 and found to possess an ordered conformation of the N-terminus, whereby the six-turn alpha helix of the N-terminus packed against ECL1 and ECL2, formed a disulfide bond with ECL2 and facilitated interactions within the ligand-binding pocket via charged and polar amino acid side chains, potentially forming polar interactions with the co-crystallised ligand [61] (Figure 2 B). Previously, a similar conformation was reported for the sphingosine 1-phosphate receptor in complex with a sphingolipid mimetic, showing the binding pocket capped by the N-terminus, which packed on top of ECL1 and ECL2 (yet without the disulfide bond seen in the LPA₁ structure) [62] (Figure 2 C). Other noteworthy rhodopsin-like GPCRs with structural resolution of the N-terminus include the CXCR4 [63] (for which interaction between receptor N-terminus and viral chemokine vMIP-II was demonstrated) (Figure 2 D) and human orexin 1 receptor [64] (hypothesised to contribute to ligand binding or recruitment of ligand orexin-A) (Figure 2 E). Also of note is the reported crystal structure of PAR1 [65], of significance to this review as the archetype of protease-dependent GPCR signalling. Yet, this crystal structure was obtained in complex with synthetic antagonist vorapaxar, and not the endogenous tethered ligand of PAR1, presumably owing to the aforementioned flexibility of the N-terminal domain, leaving the structural basis of PAR activation unresolved.

Clearly, structural biology offers unique insight into the role of the GPCR N-terminus for receptor activation, although this current knowledge is primarily in relation to the conformation of the N-terminus in receptor states selected by (often synthetic) ligands. Yet these examples are all recent, demonstrating the rapid advancement in the field of GPCR crystallography, and optimistically, demonstrates the potential of structural biology for understanding many of the remarkable examples of GPCR signalling previously discussed in this review.
6. Adhesion GPCRs

The previous discussion has examined the role of the N-terminus in GPCR signalling, in ways that are both independent of and dependent on proteolysis. Whilst separate from the rhodopsin-like GPCRs, the adhesion GPCRs (so named for their roles in cell adhesion and protein-protein interactions) warrant mention due to their existence as a fascinating intermediate between receptors with protease-mediated and protease-independent N-terminus-regulated GPCR signalling. The adhesion receptors are comparable to PARs, in that a tethered agonist sequence [designated the stachel sequence by Liebscher, et al. [66]] is revealed following removal of the N-terminal ECD, yet this proceeds without involvement of a separate protease. The N-terminal tethered agonist is exposed following autoproteolysis at a GPCR proteolysis site (GPS), catalysed by the broader GPCR-autoproteolysis inducing domain (GAIN) [67]. The presence of a tethered agonist within the N-terminus of two exemplary adhesion receptors, GPR126 and GPR133, was demonstrated by Liebscher, et al. [66]. Despite poor surface expression, HA-tagged mutants of GPR126 and GPR133, both engineered to lack the ECDs that would be removed following autoproteolysis (termed CTF mutants), had increased constitutive cAMP accumulation, as assessed in COS-7 cells and compared to HA-tagged wild-type counterparts. Surface expression was enhanced following fusion of the P2Y12 receptor N-terminus to all mutants, and this system was used to show that further mutation removing the entire native GPR126 and GPR133 N-termini, including the GPS, (termed ΔGPS-CTF mutants), not only abolished the increase in constitutive signalling, but reduced activity to below that of their wild-type equivalents. This suggested that residues important for activation lay C-terminally to the GPS, and indeed, exogenous application of peptides mimicking these key regions of GPR126 or GPR133 to GPR126-ΔGPS-CTF and GPR133-ΔGPS-CTF respectively, restored signalling. Indeed, these features are exemplified and largely conserved within the adhesion GPCR family although it should be noted that not all adhesion GPCRs are cleaved or activated in the manner described above [68].
7. Conclusion

Recent advances in the field of GPCRs show the humble N-terminus to be more than a ligand-binding site, a substrate for glycosylation or an anchor for a signal peptide, as captured in the diverse examples summarised in Table 1. The protease-activated receptors are exemplars of this, revealing that GPCRs can harbour their own agonists within their N-termini, uncovered following a regulated proteolytic event. Similarly, the MC₄ receptor N-terminus contains an agonist for its own receptor, allowing evolution of agouti-related peptide as an endogenous inverse agonist, through its ability to antagonize the N-terminus. In contrast, we report orphan receptor GPR37L1 to have constitutive activity maintained by its N-terminus and abolished by metalloprotease-mediated removal of the region. This differs again from the N-terminus of yeast pheromone α-factor receptor Ste2p, which reportedly serves to constrain the receptor’s response to its agonist. The receptors that have been discussed in this review represent emerging examples of novel modalities of GPCR signalling, whereby the N-terminus is shown to be a crucial and dynamic contributor to signal transduction.
Figure 1. The N-terminus can be a dynamic participant in GPCR signalling

Diagram shows diverse examples of how the GPCR N-terminus can directly influence receptor signalling.

A. Protease-activated receptor 1 (PAR1) contains a tethered ligand sequence within its N-terminus that is revealed following N-terminal cleavage by thrombin, causing receptor activation. B. Orphan receptor GPR37L1 has constitutive activity until proteolysis by an MMP/ADAM removes the entire N-terminus and silences receptor signalling. C. The melanocortin-4 (MC4) receptor N-terminus contains a tethered ligand sequence that requires no proteolysis. This facilitates the actions of endogenous inverse agonist agouti-related peptide (AgRP), which antagonizes the receptor’s tethered ligand. D. Single residue mutations in the 5-HT2B N-terminus, identified within a drug-abusing population, cause increased constitutive and ligand (5-HT)-mediated activity. In this way, the wild-type N-terminus serves to constrain receptor activity.
Figure 2: The structural diversity of GPCR N-termini.

Crystal structures of GPCRs with their N-termini partially resolved are shown. A. The cannabinoid receptor 1 (CB₁) N-terminus forms hydrophobic interactions with AM6538 (PDB: 5TGZ). B-C. The N-termini of the lysophosphatidic acid receptor (LPA1) (PDB: 4Z34) and sphingosine 1-phosphate receptor (S1P1) (PDB: 3V2Y) form a cap over the ligand binding pocket, allowing for polar and charged interactions with the ligand. D. The CXCR4 receptor in complex with a viral chemokine vMIP-II displays extensive contacts with the receptor’s N-terminus (PDB: 4RWS). E. The ordered structure of orexin 1 (OX1) receptor N-terminus contributes to ligand binding of the endogenous peptide orexin-A, but not with small molecule ligands (4ZJ8). The N-terminus of each receptor is coloured in purple. Except for D., ligands are displayed in a space-filling model, with CPK colouring where the carbons are yellow. N-terminus residues involved in ligand interactions are space-filled in the side orientation, coloured by
their chemical nature: polar (pink), positive (blue), negative (red) and hydrophobic (yellow). In the top view, residues are displayed in ball-and-stick.
Table 1 Summary table of the diverse modalities by which the GPCR N-terminus can directly and indirectly contribute to receptor signalling.

1. Signalling modified by N-terminal proteolysis

<table>
<thead>
<tr>
<th>Protease-activated</th>
<th>Protease-inactivated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptor</td>
<td>Mechanism</td>
</tr>
<tr>
<td>Protease-activated receptors 1-4</td>
<td>Regulated cleavage of N-terminus exposes new N-terminal sequence – a tethered ligand that auto-activates the receptor [20].</td>
</tr>
<tr>
<td>Thyroid-stimulating hormone receptor</td>
<td>Proteolytic removal of N-terminal domain may liberate the receptor from constraints on constitutive signalling [30, 31].</td>
</tr>
</tbody>
</table>

2. Expression modified by N-terminal proteolysis

<table>
<thead>
<tr>
<th>Protease-augmented</th>
<th>Protease-diminished</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptor</td>
<td>Mechanism</td>
</tr>
<tr>
<td>α1D-adrenoceptor</td>
<td>Partial removal of N-terminus necessary for proper trafficking to cell surface [47]</td>
</tr>
<tr>
<td>β1-adrenoceptor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Signalling modified by N-terminus (protease-independent)

<table>
<thead>
<tr>
<th>N-terminus-enhanced</th>
<th>N-terminus-constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptor</td>
<td>Mechanism</td>
</tr>
<tr>
<td>GPR61</td>
<td>Potential tethered agonist domains within N-terminus [54].</td>
</tr>
<tr>
<td>Melanocortin-4 Receptor</td>
<td>N-terminus contains a tethered agonist [50, 51].</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References:


