The third extracellular loop of G-protein-coupled receptors: more than just a linker between two important transmembrane helices

Z. Lawson and M. Wheatley
School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

Abstract

GPCRs (G-protein-coupled receptors) are a large family of structurally related proteins, which mediate their effects by coupling with G-proteins. Despite responding to a range of very diverse stimuli, these receptors exhibit a conserved tertiary structure comprising a bundle of seven TM (transmembrane) helices linked by alternating ECLs (extracellular loops) and ICLs (intracellular loops). The hydrophobic environment formed by the cluster of TM helices is functionally important. For example, the 11-cis retinal chromophore of rhodopsin forms a protonated Schiff base linkage to a lysine in TM7, deep within the helical bundle, and small ligands, such as amine neurotransmitters and non-peptide analogues of peptide hormones, also bind within the corresponding region of their cognate receptors. In addition, activation of GPCRs involves relative movement of TM helices to present G-protein interaction sites across the intracellular face of the receptor. Consequently, it might be assumed that the ECLs of the GPCR are inert peptide linkers that merely connect important TM helices. Focusing on ECL3 (third ECL), it is becoming increasingly apparent that this extracellular domain can fulfill a range of important roles with respect to GPCR signalling, including agonist binding, ligand selectivity and receptor activation.

Introduction

GPCRs (G-protein-coupled receptors) can be divided into three major groups on the basis of sequence homology, namely the rhodopsin/β-AR (β-adrenergic receptor) family (Family A), the secretin receptor family (Family B) and metabotropic glutamate receptors (Family C). The largest and most extensively characterized of these is the rhodopsin/β-AR family. Despite being activated by a wide variety of stimuli from photons to glycoproteins, all of these different classes of GPCR exhibit a common protein fold comprising seven TM (transmembrane) domains connected by alternating ECLs (extracellular loops) and ICLs (intracellular loops). The loop regions of GPCRs are the most variable structural elements of the receptor with respect to both length and sequence. Even subtypes in the same receptor family can exhibit very low sequence homology in their loop regions. Consequently, it would not be unreasonable perhaps to conclude that this lack of sequence conservation may be indicative of a lack of functional significance. Furthermore, the crystal structure of bRho (bovine rhodopsin) revealed the binding contacts of retinal with opsin to be deep within the helical bundle [1]. Similarly, the ligand-binding site for small ligands such as biogenic amines and non-peptide analogues of peptide hormones has been mapped to a pocket defined by the TM cluster of the relevant GPCR. Although ICLs are required for G-protein interaction and receptor regulation through kinases, arrestins, scaffolding proteins etc., a role for the ECLs is not so apparent perhaps. An exception to this is ECL2, which in the bRho crystal structure forms contacts with 11-cis retinal by plunging deep into the TM bundle [1]. All of this may give the impression that ECL1 and ECL3 are peptide linkers incorporated into the GPCR structure merely to join functionally important TM helices together. However, there is now increasing evidence that this is not the case and that ECLs fulfill important functional roles within the GPCR architecture. In the present study, we examine the functional significance of ECL3, focusing primarily on Family A receptors.

Basic features of ECL3

Sequence analysis of 270 members of Family A GPCRs (including orphan receptors) indicates that ECL3 varies in length between 4 and 27 residues, with an average length of 10–11 residues. Overall homology between the ECL3 of different receptors is low, with no single residue absolutely conserved, in contrast with the cysteine conserved in ECL2. The defining feature of this loop is that it connects TM6 and TM7. At the cytoplasmic terminus of these two helices are ICL3 and helix 8 respectively, both of which have been implicated in G-protein activation and subsequent second-messenger generation (reviewed in [2]). ECL3 is usually a short loop and as such may constrain TM6 and TM7, thereby limiting their motion and preserving the ground state of the
The functional importance of ECL3 is independent of the size of the ligand

The role of ECL3 has been investigated using a chimaeric receptor strategy in which the ECL3 sequence of one GPCR was replaced by the corresponding sequence of a different GPCR. A β2AR, which incorporated the ECL3 of the α1aAR, exhibited higher binding affinity for agonists and a higher agonist-independent basal adenylate cyclase activity than wild-type β2AR, whereas antagonist affinity was not affected [6]. These effects are commensurate with the generation of a constitutively active mutant. A probable explanation for the constitutively active mutant is that replacing the ECL3 of the β2AR with the corresponding α1aAR formed an appropriately folded receptor but relaxed the conformational constraint on TM6 and TM7 provided by the wild-type β2AR ECL3 [6]. In marked contrast, substituting the ECL3 of the TSHR (thrytropin receptor) with the ECL3 of the β2AR did not affect basal signalling but decreased the potency and magnitude of the TSH (thyroid-stimulating hormone or thyrotropin)-stimulated cAMP response [7]. In this case, it appears that the β2AR ECL3 in the TSHR/β2AR chimaera imposed a greater constraint on TM6 and TM7 than the TSHR sequence. Taken together, these two observations are consistent with the β2AR ECL3, imposing a greater constraint than the corresponding segment of either the TSHR or the α1aAR. The role of ECL3 is not restricted to regulating GPCR transition to the active conformation but also encompasses ligand recognition. This was shown using AT1R/AT2R chimaeric constructs, which revealed that the ECL3 domain provides the main molecular discriminator underlying the AT1R-R-selective binding of the agonist CGP 42112A [N-α-nicotinoyl-Tyr-(N-α-CBZ-Arg)-Lys-His-Pro-Ile-OH] [8]. It is noteworthy that all three of the receptor examples cited above are activated by agonists of different chemical classes ranging in mass from 170 Da (noradrenaline; biogenic amine) to 1046 Da (AngII; octapeptide) to approx. 25 000 Da (TSH; glycohormone with α and β subunits), indicating that the functional importance of ECL3 is independent of the size or chemical nature of the ligand.

There is no conserved ‘hot spot’ of important residues within the ECL3 of different GPCRs

Individual residues within the ECL3 of a wide range of GPCRs have been identified as important for ligand binding and/or intracellular signalling. The approximate location and functional role of these essential residues from a range of GPCRs is presented in Figure 1, mapped on to a schematic representation of a typical ECL3. Table 1 lists the details of the individual residues in each case, which have been implicated using site-directed mutagenesis studies. From the information presented, it is apparent that there is no single locus or side-chain characteristic common to these key ECL3 residues. This contrasts with many functionally important residues in the TM helices, which are highly conserved throughout many GPCRs and which give rise to the characteristic ‘signature residues’ of Family A GPCRs (e.g. Asp252 [18] and the NPXXY motif in TM7). Although there is no functional ‘hot spot’ in ECL3, there does appear to be a tendency for such residues to occupy the proximal and distal juxtamembrane segments of the loop (Figure 1).

In conclusion, the ECL3 domain of GPCRs is not merely a peptide feature linking TM6 to TM7. The loop occupies recipient. Substitutions in ECL3 which relieve this constraint may well impact on R → R* transition. Indeed, this is exactly what was observed during random mutagenesis of the δ-opioid receptor, where activating mutations of the ECL3 led to the suggestion that ECL3 was the origin of the activation pathway for GPCRs [3].

ECL3 can also form contacts with the N-terminal domain. For example, the ECL3 of bRho forms a network of hydrogen bonds with the oligosaccharide chain on Asn65 [1]. However, such interaction will be receptor-specific as the N-linked glycosylation pattern varies considerably, even between closely related GPCRs [4]. For some Family A GPCRs, such as the AT1R (angiotensin II receptor), the N-terminus and ECL3 are covalently linked through a disulphide bond [5], which will impose additional conformational constraints.
Table 1 | Key residues in the ECL3 of GPCRs

For each receptor, the identity and location in the sequence of the key functional residue(s) in ECL3 are presented together with the appropriate reference. The approximate location of these residues within the ECL of each GPCR is shown in Figure 1. CCK1R, cholecystokinin 1 receptor; FSHR, follicle-stimulating hormone receptor; GnRHR, gonadotropin-releasing hormone receptor; LHR, luteinizing hormone/choriogonadotropin receptor; mACHR, muscarinic acetylcholine receptor; NTS1R, neurotensin receptor 1; P2Y1R, purinoceptor 2Y1.

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Key residue(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1R</td>
<td>Asp281</td>
<td>[9]</td>
</tr>
<tr>
<td>CCK1R</td>
<td>Arg136</td>
<td>[10]</td>
</tr>
<tr>
<td>FSHR</td>
<td>Leu583, Ile584 and Lys590</td>
<td>[11]</td>
</tr>
<tr>
<td>GnRHR</td>
<td>Val599 and Glu601</td>
<td>[12,13]</td>
</tr>
<tr>
<td>LHR</td>
<td>Lys583</td>
<td>[14]</td>
</tr>
<tr>
<td>M1 mACHR</td>
<td>Ser208 and Thr209</td>
<td>[15]</td>
</tr>
<tr>
<td>NTS1R</td>
<td>Trp139, Phe144 and Tyr147</td>
<td>[16]</td>
</tr>
<tr>
<td>P2Y1R</td>
<td>Arg287</td>
<td>[17]</td>
</tr>
</tbody>
</table>

a significant position in the GPCR architecture and it is becoming increasingly apparent that this enables ECL3 to provide important epitopes for normal receptor functioning.

Financial support from the Wellcome Trust and BBSRC is gratefully acknowledged.

References


Received 23 July 2004

©2004 Biochemical Society