Published in final edited form as:

Nat Struct Mol Biol. 2018 June ; 25(6): 538–545. doi:10.1038/s41594-018-0071-3.

Molecular mechanism of modulating arrestin conformation by GPCR phosphorylation

Andrija Sente^{1,*}, Raphael Peer¹, Ashish Srivastava², Mithu Baidya², Arthur M. Lesk^{1,3}, Santhanam Balaji¹, Arun K. Shukla², M. Madan Babu^{1,*}, and Tilman Flock^{1,4,5,*}

¹MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK

²Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India

³Department of Biochemistry and Molecular Biology and Huck Institutes of Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA

⁴Fitzwilliam College, Cambridge, CB3 0DG, UK

Abstract

Arrestins regulate signalling of ligand-activated, phosphorylated G protein-coupled receptors (GPCRs). Different patterns of receptor phosphorylations (phosphorylation barcode) can modulate arrestin conformations, resulting in distinct functional outcomes (e.g. desensitization, internalisation, downstream signalling). However, the mechanism of arrestin activation, and how distinct receptor phosphorylation patterns could induce different conformational changes on arrestin, are not fully understood. Here, we analyze how each arrestin amino acid contributes to its different conformational states. We identify a conserved structural motif that restricts the mobility of the arrestin finger-loop in the inactive state and appears to be regulated by receptor phosphorylation. Distal and proximal receptor phosphorylation sites appear to selectively engage with distinct arrestin structural motifs (i.e. micro-locks) to induce different arrestin conformations. These observations suggest a model where different phosphorylation patterns of the GPCR C-terminus can combinatorially modulate the conformation of the finger-loop and other

*Correspondence: as2475@cam.ac.uk, madanm@mrc-lmb.cam.ac.uk OR tf284@cam.ac.uk.

⁵Current address: Laboratory of Biomolecular Research, Division of Biology and Chemistry, Paul Scherrer Institute, Villigen PSI, Switzerland

Code availability

All packages and publically available computer code is described in the methods section of this manuscript. For availability of codes that were developed in-house, please contact the corresponding authors.

Data availability

All data generated during this study are included in this article and its supplementary files, or available from external sources: all protein structure data analyzed in this study are available in the RSCB Protein Data Bank (list of identifiers and references provided in Supplementary Tables 1). Mapping of Common Arrestin Numbering to available PDB structures is given in Supplementary Table 3. Source data of the bimane fluorescence are available with the paper online.

Contributions

AS and TF collected data. AS, RP and TF wrote scripts and processed the data. SB generated the arrestin alignment. AL helped with analysing conformational changes. ASr, MB, and AKS designed and performed bimane fluorescence experiments. All authors analysed and interpreted the results. AS, MMB, and TF wrote the manuscript with inputs from all authors. MMB and TF designed and supervised the project.

Competing financial interests

The authors declare no competing financial interests.

phosphorylation-sensitive structural elements to drive distinct arrestin conformation and functional outcomes.

Introduction

Arrestins are a family of proteins crucial for cellular signaling and regulation of G protein-coupled receptors (GPCRs), the largest class of human membrane protein receptors¹. GPCRs transduce diverse extracellular signals – such as hormones and neurotransmitters – into physiological responses through activation of intracellular heterotrimeric G proteins, which in turn regulate the production of second messengers to orchestrate different downstream signalling cascades^{2,3}. Subsequent receptor phosphorylation by G protein-coupled receptor kinases (GRKs) recruits arrestins, resulting in GPCR desensitization and (or) internalization^{4,5} (Fig. 1a). In addition to this regulatory role, arrestins are increasingly appreciated to link GPCRs to G protein-independent signalling cascades, such as through mitogen-activated protein kinases (MAPKs)⁶.

Although there are over 800 different human GPCRs, the human genome encodes only four arrestin genes to regulate the receptors. All four arrestin types share a common protein fold comprising an N-terminal and a C-terminal domain, each of which is made up of a seven-stranded β -sandwich. Despite this structural similarity, there are important functional distinctions: arrestin-1 and arrestin-4 (the ‘visual’ arrestins) are specific for rhodopsin and cone opsins, respectively, while arrestin-2 and arrestin-3 (β -arrestin 1 and 2) are ubiquitously expressed and interact less specifically with the majority of GPCRs².

The interaction with activated and phosphorylated receptor converts arrestin to its active state. Among the many aspects of arrestin activation that have been well studied⁷, it has been recently discovered that different patterns of phosphorylation on the receptor (termed the ‘phosphorylation barcode’) mediated by GRKs can induce distinct arrestin conformations, which are capable of mediating specific functions such as internalization, desensitization and downstream signaling^{8,9}. In addition, it has recently been shown that G proteins and arrestins can simultaneously bind to a GPCR¹⁰. In these supercomplexes, arrestin interacts with the receptor only through the receptor’s phosphorylated carboxyl terminus (C-tail engaged complex). Importantly, the ‘tail’ conformation of the GPCR-arrestin complex is fully capable of mediating receptor endocytosis, ERK binding and activation, but not desensitization of G protein signaling^{11,12}. Instead, a fully-engaged GPCR-arrestin complex is required to quench G protein signaling^{12,13}.

Structural studies of arrestin have revealed three conserved motifs that maintain the inactive state of arrestin (Fig. 1b): (1) the three-element interaction (TE) consisting of α -helix 1, β -strand 1 and the C-terminal tail, (2) the polar core and (3) a number of hydrophobic inter-domain interactions⁷. Current models of arrestin activation and binding to GPCRs suggest a multistep process in which arrestin first engages with the phosphorylated C-terminus of the receptor^{7,14,15}. This disrupts the three-element interaction, resulting in the release of the C-terminal tail of arrestin¹⁵. This in turn destabilizes the polar core, leading to a major conformational change of arrestin, which is required to form the fully engaged GPCR-arrestin complex¹⁵. One of the most prominent changes is the movement of the finger-loop

(connecting strands S5 and S6), which in the fully engaged GPCR-arrestin complex, inserts into the crevice of the activated receptor and undergoes a disorder-to-order transition to adopt a helical conformation^{16,17}.

Despite these observations, the molecular mechanism by which different phosphorylation patterns on the receptor lead to different arrestin conformations, that mediate distinct functional outcomes, has remained elusive. Furthermore, the transition from the inactive state to the tail-engaged or the fully engaged GPCR-arrestin complex seems to have an important regulatory role in determining arrestin functions¹⁸. However, the exact mechanism by which this transition occurs remains unclear. Recent structure determinations have made it possible to address these questions. Over 18 Arrestin structures are available in different conformations, including the inactive, pre-activated (p44 splice variant), phosphopeptide-bound, receptor-bound and clathrin-bound conformations (Fig. 1; Methods and Supplementary Table 1). Here, we perform a systematic analysis of all available arrestin structures and introduce a common arrestin numbering (CAN) system to simplify the comparison of different arrestins and their functional regions. Integration of this information with sequence and structural data enabled us to trace structural rearrangements that govern transitions between arrestin functional states at the level of single residues.

Residue contact network analysis of arrestin protein structures

The analysis of non-covalent contacts between amino acids in protein structures (residue contact networks; RCNs) is a systematic way to compare every single residue contact in a protein, and allows rapid and unbiased comparison of multiple protein structures to detect small structural differences^{19–25}. In these networks, nodes represent amino acid residues, and edges between nodes denote non-covalent contacts between the corresponding residues. Hence, conformational states of proteins can be represented as networks of residue contacts and conformational changes correspond to rearrangements of contacts in these networks.

We compared contacts between topologically equivalent residues of different arrestin structures using a common arrestin numbering (CAN) system, which was developed using an approach that we used for G proteins (Fig. 2a, Supplementary Data Set 1)²³: A structural alignment of all arrestin structures was linked to a sequence alignment of additional arrestin sequences (Methods). With this, each arrestin position is referred to by a common identifier, which consists of three levels (D.S.P): the first part corresponds to the domain (N or C), the second part corresponds to the secondary structure element (SSE) and the third to the position within the corresponding SSE (Fig. 2a). The CAN defines nineteen β strands, denoted S1-S19 and one alpha helix (H1). Each loop was named by lowercase letters indicating their flanking SSEs; for instance, the CAN name for finger loop is N.s5s6, as it is within the N domain, between strands S5 and S6. The border between the N and C domains was defined as the end of strand 10 (N.S10). The CAN is added in superscript to a residue: for example, phenylalanine 9 in bovine arrestin-2 (*i.e.*, β -arrestin-1; *e.g.*, PDB ID: 1g4r) corresponds to Phe^{9N.S1.2} (second residue within the first strand of the N domain). The equivalent residue in bovine arrestin-1 (*e.g.*, PDB ID: 1ayr) is Phe^{13N.S1.2}.

We determined the residue contact networks (RCNs, Methods) of 18 arrestin structures. Together, all residue contacts of a particular protein structure represent a ‘contact fingerprint’ typical for that protein and its conformation^{24,26}. Clustering protein structures by their ‘contact fingerprints’ allows for an unbiased classification of distinct conformational states. The arrestin structures cluster into two major groups: inactive state (13 structures, Methods) and active state (5 structures). Within the cluster of active-state structures, the V2Rpp (vasopressin-2 receptor phosphopeptide)¹⁴ bound arrestin forms a sub-cluster with the inositol hexakisphosphate-bound arrestin²⁷. The fact that these two structures appear to have similar contact fingerprints suggests that the phosphorylation-dependent mechanism of arrestin activation is conserved irrespective of the source of phosphates. Another sub-cluster within the active state is formed by GPCR-bound arrestins and the p44 splice variant. Our method also clearly separates visual and non-visual arrestins into sub-clusters within the inactive state cluster, and rod and cone arrestins branch separately within the visual arrestin sub-cluster (Supplementary Fig. 1). Thus, the approach of investigating contact fingerprints is suitable for studying the structural basis of differences in function of visual and non-visual arrestins (Supplementary Fig. 2, Supplementary Fig. 3, and Supplementary Data Set 2).

By identifying those contacts that are present in all structures of a particular conformational state (Methods), we generated the ‘consensus contact network’ for each state (Fig. 2b). Comparison of consensus residue contacts across the different states of arrestin allows systematic inference of contacts that are (1) shared between all conformational states (i.e. active and inactive) and hence represent the conserved structural scaffold of the arrestin fold, (2) unique to a particular state and thus important for conformation-specific arrestin function, and (3) variable contacts that are unique to each protein structure (Fig. 2c). This dataset (Supplementary Data Set 3) offers a comprehensive, quantitative characterization of all available arrestin structures and sequences, and can be used to address multiple questions, such as uncovering the structural differences between visual and non-visual arrestins at the molecular level. Here, we focus on the identification of conserved structural motifs that are re-organized upon arrestin activation. In line with previous studies, we find major structural rearrangements in consensus contact network involving the polar core, the three element and the inter-domain interaction motifs^{7,28}. Through this objective approach, we detected a previously uncharacterized motif that appears to control the movement of the finger loop and play an important role in arrestin activation by interacting with phosphorylated receptor residues.

An ionic lock controls release of the finger-loop

The ‘finger-loop’, which links the 5th and 6th sheet of the N-terminal domain of arrestin, is known to play an important role in interacting with the activated receptor^{16,29}. The consensus contacts that are unique to inactive state arrestin reveal a conserved structural motif in the finger-loop that is formed by a set of residue contacts between highly evolutionary conserved negative charges (EDxD; N.s5s6.3, N.s5s6.4, N.s5s6.6; 76, 93 and 91% sequence identity in >50 arrestin sequences covering a diverse range species; Methods) and two conserved arginines in the β sheets S5 and S6 of the N-domain (N.S5.11, N.S6.2; 96 and 91% sequence identity) (Fig. 3a-c). In the inactive form of arrestin, these ionic contacts appear to lock the finger loop in a conformation where it folds back onto the surface

of the N-domain. In this conformation, the finger loop residues are presumably prevented from interacting with the activated receptor.

This ‘ionic lock’ of the finger loop is present in all 11 inactive state structures (Supplementary Fig. 4a). In some structures with multiple chains in the asymmetric unit, however, we observe that the lock is broken in some of the chains. By analysing the inter-chain residue contacts between different monomers in such asymmetric units, we find that this is due to structural distortions arising from crystal packing contacts of the finger loop (which has been ascribed to be an effect of dimer formation of the inactive state of arrestin)³⁰. To ensure that no other crystal lattice contacts interfere with the finger loop, we additionally compared residue contacts between protein chains of neighbouring crystal units for all investigated structures (Supplementary Table 2 for full details). This shows that the lock is indeed engaged in all arrestin chains of the inactive conformation, and only chains in which this region is involved in crystal packing tend to have slightly deviating conformations (Supplementary Fig. 4b-c). In contrast, in all active state structures (both, in all active structures not bound to a GPCR, and active structures in complex with GPCR, where the finger loop adopts a helical conformation) the finger loop is present in an open (extended) conformation. Direct experimental analysis of the finger-loop position in the tail-engaged and the fully-engaged receptor- β -arrestin-1/2 (arrestin-2/3) complexes using bimane fluorescence spectroscopy further supports the existence of multiple finger-loop conformations and a conformational transition during this biphasic interaction (Supplementary Fig. 5; formation of the tail-engaged complex followed by the fully engaged complex with the receptor). We analysed the interactions between the V₂R phosphopeptide (V₂Rpp) and β -arrestins-1 and 2, mimicking the tail-only engaged receptor-arrestin conformation. Interaction of V₂Rpp with β -arrestins labelled with mBBr (monobromobimane) in the finger loop results in an increase in fluorescence intensity. This suggests the disruption of the ionic lock and the release of the finger-loop. We then analysed the interactions between β -arrestins-1 and 2 with the active (agonist-bound) and phosphorylated receptor (β_2 V₂R – a chimeric β_2 -adrenergic receptor that harbours the V₂R tail). Interaction of the receptor with β -arrestins labelled with mBBr in the finger loop leads to a significant decrease in bimane fluorescence arising from the quenching effect due to the core interaction between the receptor and the finger loop (Supplementary Fig. 5).

Earlier studies using EPR spectroscopy reported immobilization of spin labels in the finger-loop when arrestin-1 interacts with inactive phosphorylated rhodopsin, which is further enhanced if rhodopsin is both activated and phosphorylated³¹. In arrestin-2, however, the spin label appears already strongly immobilized even when arrestin interacts with inactive phosphorylated receptor, with little increase if the receptor is both active and phosphorylated³². This suggest that the finger-loop –once released by arrestin interaction with receptor-attached phosphates– is able to engage with the receptor regardless of its activation state. Furthermore, DEER spectroscopy revealed a dominant population of arrestin with a distance between L68C (in the finger-loop) and V167C (on the surface of N-domain) of ~20 Å in the free state. When activated phosphorylated receptor was added, another population emerges with L68C-V167C distance of ~45 Å³³.

Taken together, while we observed finger-loop to be in a locked state in the crystal structures, *in vivo*, it has to be in equilibrium between the locked and the open state (i.e. 'statistically' dynamic) to allow opening of the finger loop upon interaction with the phosphates. In other words, the loop can sample different conformations, but interaction with the phosphorylated tail of the receptor will increase the probability of breaking the lock and releasing the finger loop, and consequently form a helix in the fully-engaged receptor bound state (and hence is less flexible). Thus, finger loop conformations (bent conformation in the inactive state and released to interact with the receptor) are likely linked to the engagement of the receptor C-tail with β -arrestin and further stabilization upon the interaction with the activated receptor core.

Regulation of the ionic-lock by receptor phosphates

Which elements of the receptor C-tail trigger the conformational transition of the finger loop? Given that receptor-attached phosphates are known to be important for arrestin activation, we sought to investigate a potential link between receptor phosphorylation and finger-loop conformation. To this end, we identified all arrestin residues that interact with receptor-attached phosphates based on the existing structures. Interestingly, we find that the conserved arginine (N.S5.11; see above) of the ionic lock interacts with a phosphate in the phosphopeptide-bound structure. Hence, phosphorylation of the receptor appears to break the ionic finger loop lock by competing with the EDxD-motif of the finger-loop for one or both arginines in the N-domain of arrestin. Therefore, these receptor-attached phosphates could release the finger loop into an 'open conformation' that can interact with the receptor (Fig. 3c).

If this were true, one should expect that mutating one or both arginines of the finger-loop lock would mimic the effect of receptor-attached phosphates by pre-releasing the finger loop, thus increasing the affinity of arrestin for the receptor. To investigate this, we obtained data from two independent alanine mutagenesis studies on arrestins^{29,34}, in which every arrestin residue was mutated into alanine and binding to the receptor was measured. We then mapped the mutation effects on the residues from the finger-loop lock. As expected, we find that mutating R^{N.S5.11} into alanine dramatically increased binding affinity of the mutated arrestin for the receptor, similar to mutations in the polar core and the TE interaction. However, mutating R^{N.S6.2} only showed increased affinity in one of the studies, and no effect in the other (Supplementary Fig. 6, see Methods for a discussion of potential reasons of this discrepancy). Although mutations of the EDxD motif would be expected to pre-release the finger loop, they might interfere with finger-loop interaction with the receptor as it forms part of the receptor-arrestin interface. Therefore, it is not surprising that mutations in this region do not appear to strongly increase affinity for the receptor.

Different phosphorylated GPCR regions regulate distinct arrestin motifs

Next, we sought to investigate if similar behaviour could be observed in other arrestin residues that interact with receptor-attached phosphates. To address this question, we mapped the mutation effect on every arrestin residue that was identified to interact with phosphates (Fig. 4). We find that mutating several positions decreases arrestin binding to the

receptor, suggesting that they might contribute to affinity by providing an electrostatic patch that interacts with the phosphates. Strikingly, we find three clusters of residues on arrestin which when mutated to alanine increase the ability of arrestin to bind the receptor. Moreover, a majority of these positions do not directly form a part of the fully engaged receptor-arrestin interface. Together this suggests that mutating these positions likely induces conformational changes at the interface position on arrestin and thereby facilitate binding to the receptor. These clusters of residues that increase binding contain residues that form part of the finger-loop lock, residues from regions surrounding the polar core and the three-element interaction. This suggests that alanine mutations at these positions likely result in altering the conformations of key structural motifs, that act as inactive state locks, to mediate better interaction with the receptor, possibly mimicking the binding of distinct receptor attached phosphates.

Analysis of the phosphopeptide-bound arrestin structure shows a pattern of distribution of phosphate groups, whereby the proximal phosphates on the receptor appear to break the ionic finger loop lock, while the more distal phosphates interact with residues surrounding the polar core (e.g. $K^{N.s8s9.9}$ interacts with a phosphate which disrupts its contact with $Y^{N.S5.12}$), and the TE residues, disrupting conserved contacts that stabilise it, such as those between strand 1 and the C-tail ($V^{N.S1.1}-K^{C.s19c.42}$ and $K^{N.S1.1}-E^{C.s19c.43}$). Thus, different phosphorylated regions of the receptor C-tail (proximal or distal) appear to preferentially engage with distinct structural motifs on arrestin. Together, these observations suggest that different combinations of phosphorylation of receptor residues can lead to selective engagement and disruption of each of these structural motifs and thereby introduce distinct conformations of arrestin (Fig. 4).

Discussion

Different phosphorylation patterns of GPCRs have been suggested to induce arrestin conformations capable of performing distinct functions. Most prominently, a tail engaged and a fully engaged GPCR-arrestin complex have been observed. It has recently been shown that in the 'tail conformation', G proteins and arrestins can simultaneously bind to a GPCR¹⁰. Which factors modulate the transition between the tail- and fully-engaged complex is less well understood. In this analysis, we identified a new arrestin regulatory structural motif that suggests a molecular mechanism by which phosphorylation at different receptor positions by kinases could mediate distinct arrestin conformations. Furthermore, distinct regions on the receptor tail appear to preferentially engage with distinct structural motifs that are characteristic of arrestin conformational states. Together, they could influence arrestin activity and determine whether a tail- or fully-engaged GPCR-arrestin complex is formed, resulting in different functional outcomes.

In particular, the inactive state of arrestin is maintained by a set of conserved structural motifs: the three-element interaction, the polar core, and the here-identified finger-loop lock (Fig. 5). Given that these motifs are distributed across the arrestin surface and can be disrupted by phosphorylated residues on the receptor, different GPCR phosphorylation patterns could disrupt them in a combinatorial manner, thereby inducing distinct conformational changes in arrestins. If the receptor carboxyl terminus is phosphorylated

only on the positions that break the three-element interaction and the polar core (the ‘distal’ phosphates), arrestins are capable of binding to the receptor tail, internalization and G protein-independent downstream signaling. However, because the finger loop lock is still ‘locked’, arrestins are incapable of forming the fully engaged complex and hence cannot desensitize G protein signaling. Therefore, distal phosphorylation of GPCR tails is also compatible with the formation of GPCR-G protein-arrestin supercomplexes and signaling from intracellular compartments.

On the other hand, in addition to the distal phosphates, GPCRs are also phosphorylated at positions that can break the finger loop lock (‘proximal’ phosphates, closer to the receptor helix 8). These proximal phosphates might have a modulatory role by interacting with the new structural motif identified here, hence possibly playing a role in switching arrestin to form the fully-engaged GPCR-arrestin complex. In this case, G protein signaling would be quenched due to steric hindrance and the formation of GPCR-G protein complexes and signaling from intracellular compartments may not be possible^{37,38}. In addition to receptor-attached phosphates, recent studies suggest a mechanism for receptor-independent arrestin activation by binding of two molecules of inositol hexakisphosphate (IP₆)²⁷.

Interestingly, in the refined version of the fully-engaged, receptor-bound complex (PDB ID 5w0p), the proximal phosphates are missing and the receptor tail seems more distant from the finger-loop, which is already in its ‘open’ conformation. This may be because for the initial engagement with arrestin, the phosphorylated receptor tail could function as ‘interaction hot spot’ that is mainly driven by electrostatic interaction to increase the probability of receptor-arrestin co-localisation. If the receptor C-terminus is proximally phosphorylated, the finger-loop would be ‘pre-released’ to allow subsequent formation of the fully-engaged complex, eventually displacing the proximal phosphates during arrestin engagement with the receptor core. Hence, the proximal phosphates are ‘place holders’ to the finger loop until the receptor binds and the long receptor tail increases the probability to pre-recruit arrestin. This scenario would suggest that the two active state structures (tail-engaged and fully-engaged) represent two steps in a ‘sequential model’ of arrestin activation, where tail engagement precedes full engagement and can explain pre-complex formation that have been observed before¹⁵. Furthermore, while the distal part of the GPCR C-terminus is precisely positioned to break the polar core and the three-element interaction, thus facilitating arrestin activation, there could be multiple conformations of the proximal part of the C-terminus (supported by superposition of tail-engaged and fully-engaged structures) that weakly interacts with distributed charges of the surface of the arrestin N-domain, hence forming an ensemble of conformations. Such a phenomenon is often observed for disordered protein regions (often referred to as ‘fuzzy complexes’)^{35,36}, which tends to be dynamic in different time scales and hence may not be captured in simulation studies.

It should be noted that in the fully engaged complex, the proximal sites of the GPCR C-tail are not phosphorylated, which may also explain the different conformation while still forming the fully engaged complex as arrestin has been mutated to mimic the fully-activated state. The differences could also be due to the fact that the rhodopsin-arrestin structure is a covalent fusion of the two proteins, which might impose some constraints on arrestin

conformation. Moreover, both the receptor and arrestin are of different subtypes in the two structures (rhodopsin-arrestin-1 and V2Rpp- β -arrestin 1). Importantly, there is a difference in the number of phosphate groups present on the receptor C-tail in the two structures, which could also affect its conformation.

Alanine scanning mutagenesis suggest that the finger loop motif is important for arrestin binding to the receptor, while mutation of the region where the proximal phosphates bind in the fully-engaged structure do not affect complex formation²⁹. Although rhodopsin lacks the phosphorylation sites that break the finger loop lock in non-visual arrestins, a conserved set of negative charges (DDE) is found at the equivalent position in the C-terminal tail. These negative charges might play a role in opening the finger loop in visual arrestin and might be a factor for the differences observed in activation of visual and non-visual arrestins^{37,38}. For instance, by making arrestin more prone to induce fully engaged complex, thereby increasing rate of desensitisation.

In addition, some variants of arrestin, such as the p44 splice variant and squid arrestin, have the capacity to bind receptors independent of receptor phosphorylation status³⁹. The p44 splice variant lacks part of the C-tail that stabilizes the three-element interaction and polar core, and is likely already in a pre-activated state capable of binding the receptor (which may or may not include destabilization of the finger-loop lock). This is also supported by clustering results from the residue contacts of the p44 structure (Supplementary Fig. 1), which places p44 closer to the fully engaged complex. Although the finger-loop motif is present in squid arrestin sequence, it is unclear how the phosphorylation-independent binding is achieved. It has previously been speculated that this could arise due to a difference in the receptor as the squid rhodopsin contains a cluster of nine negatively charged residues in the C-tail that could mimic the role of phosphates^{39,40}.

The mechanism proposed here implies distinct roles of the proximal and distal phosphates on the GPCR C-tail. Such a mechanism provides a framework that can explain why and how different GRK mediated phosphorylation of specific receptor residues regulates arrestin functions by distinct phosphorylation patterns. Thus, phosphorylation-sensitive structural motifs on arrestin seem to function as 'micro-locks' that are distributed across the arrestin surface and can be 'unlocked' by different phosphorylation patterns (Fig. 5), suggesting a structural basis and a mechanistic link between the phosphorylation barcode and distinct downstream functional outcomes.

Online Methods

Common Arrestin Numbering (CAN)

The Common Arrestin Numbering system was developed for the systematic comparison of topologically equivalent residues across different protein structures and sequences of different arrestin subtypes, following our previously published protocol²³. For each position in the alignment, the CAN assigns a common name which consists of three parts: the domain (D), the secondary structure element (S) within which the residue is located, and the position within the particular secondary structure element (P). The boundary between domains is at the end of strand X (N.S10). This assignment of each residue position and its

secondary structure element is based on a structural alignment of all chains of all arrestin structures constructed using Mustang41. For this alignment, the consensus secondary structure was computed for each alignment position. This was done by using the STRIDE algorithm to consistently assign secondary structure for each arrestin structure. The most prominent secondary structure type at each topologically equivalent position of all arrestin structures (mean and standard deviation of secondary structure type at each CAN position) was defined as consensus secondary structure (SSEs). The consensus SSEs was validated using a 3D-structure alignment. There are 19 strands, 1 helix and 21 loops. To map arrestin sequences without structural information to the CAN, sequences alignments of different arrestin orthologs was mapped to the structural arrestin alignment. Orthologous sequences of arrestin were identified through JACKHMMER42 searches against the Uniprot database (e-value cut off 0.001 and sequence length match of at least 70% for both forward and reverse searches). Representative orthologues from major animal lineages were considered and a multiple sequence alignment was constructed using MSAProbs43 and further manually refined based on a structural arrestin alignment to arrive at the final alignment.

Residue Contact Network (RCN) analysis

Non-covalent residue contact networks—The topology, conformation and stability of a protein structure can be captured by its non-covalent contacts between amino acids²⁰. For each of the 18 arrestin structures^{14,16,17,27,30,44–55} (11 inactive structures, two structures in a pre-activated state (phosphopeptide-bound arrestin, inositol hexakisphosphate-bound and the p44 splice variant, two structure of activated arrestin in a fully engaged complex with the receptor core, and two clathrin-bound arrestin structures; see Supplementary Tables), van der Waals interactions between residues, as well as non-canonical interaction such as π – π stacking were determined as previously described²³ using a local version of the RINerator 0.556, PCA57 and NCI58. All calculations, analysis and processing were performed in R.

RCN Fingerprinting and Clustering—RCNs of each PDB structure were clustered using hierarchical clustering with complete linkage and Pearson correlation as distance measure between the fingerprints. Based on the clustering results, each arrestin structure was assigned to one of two signaling states: (1) inactive and (2) active state. The two clathrin bound structures clustered with the inactive structures, and we excluded them from the ‘consensus RCN’ of the inactive state (see below). Altogether, we analysed 11 structures in the inactive state and 5 structures in the active state: p44 splice variant (4j2q), phosphopeptide-bound arrestin (4jqj), rhodopsin-bound arrestins (4zwj, 5w0p) and inositol hexakisphosphate-bound arrestin (5tv1). One structure (4zrg) contains a mutation in the polar core region (R175E), and was thus excluded from the analysis. Additionally, structure of the finger loop peptide bound to rhodopsin (4pxf) was excluded from the analysis as it did not contain a full length arrestin.

All motifs described in the manuscript are conserved even in the evolutionary distant tiger salamander structure (PDB ID: 1suj), where the finger-loop lock is also found in a bent conformation. In addition, we identified corresponding residues in human and tiger salamander cone arrestins using the structure of phosphorylated vasopressin peptide-bound

β -arrestin 1, the arrestin-rhodopsin complex structures and the Common Arrestin Numbering system. We find that 17 of 19 arrestin positions that interact with receptor attached phosphates and 22 of 27 interface residues that interact with the receptor helix bundle are identical between human and tiger salamander cone arrestin. Thus, the identified motifs is conserved even in more distantly related receptor-arrestin interactions.

Consensus RCNs—To generate the consensus RCNs, only contacts that are present in all chains of a particular structure were considered. Subsequently, for each signaling state, contacts that were present in all structures in the particular signaling state were defined as consensus contacts. These two filtering steps help to increase stringency of the contact definition and reduce noise from contact networks.

Mapping contacts to structure—To study the re-organization of residue contacts in the different conformational states of arrestin, the RCytoscape interface was used to export consensus RCNs from R to Cytoscape. To visualize consensus contacts in 3D, structure coordinate files were created to show C-alpha atoms as spheres and residue contacts as lines between them using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.)

Alanine mutagenesis data

Results from two independent full alanine mutagenesis screens of arrestin-129,34 validate our observations. The data were mapped to CAN in order to assign them to residues in the consensus networks. The mutation effects are shown as percentages relative to wild type (wild type is 100%). In Ostermeier et al.²⁹, both R66 (N.S5.11) and R80 (N.S6.2) show enhanced binding to receptor upon mutation into alanine. In Peterhans et al.³⁴, the authors measure binding of mutant arrestins to both Meta-II-P rhodopsin (phosphorylated and ligand-bound rhodopsin) and Ops-P (phosphorylated receptor with no ligand bound). In our Analysis, we only used data for binding to Meta-II-P, which shows that mutation of R66 increases binding to the receptor relative to WT, while mutation of R80 has no effect. In Peterhans et al. buffer conditions for pull-down experiments were changed, and receptors are phosphorylated more as compared to Ostermeier et al. This is the reason why WT arrestin has higher ‘baseline’ binding affinity in Peterhans et al. Thus, different buffer conditions and higher phosphorylation levels might minimise the effect of single-residue mutations.

Arrestin subtype analysis

To study differences between visual and β -arrestins, we generated contact fingerprints of 11 inactive state arrestin structure by considering contacts present in 75% of chains of a particular structure. For the analysis, we did not consider clathrin-bound arrestin structures. The analysis included 5 structures of visual arrestins: 1ayr, 1suj, 1cf1, 3ugu, 3ugx, and 6 structures of β -arrestins: 2wtr, 1jsy, 1zsh, 3p2d, 1g4m and 1g4r. Altogether, there were 185 contacts present only in visual arrestin structures, 721 contacts present in both visual and β -arrestins arrestins, and 142 contacts present only in β -arrestins. To identify conserved contacts that are unique to visual arrestins, we considered contacts present in more than 80% (i.e. 4/5) structures of visual arrestins, but absent from β -arrestin structures. Likewise, to identify conserved contacts unique to β -arrestins, we considered contacts present in more

than 80% (i.e. 5/6) structures of β -arrestins, but absent from visual arrestin structures. This resulted in 35 contacts unique to visual arrestins and 16 contacts unique to β -arrestins.

In addition, we performed a sequence-based analysis to identify subtype-specific residues in arrestins. For each human arrestin subtype, an alignment of orthologous sequences, obtained from the Orthologous Matrix Database⁵⁹, was constructed using the Multiple Sequence Comparison by Log-Expectation (Muscle) software. In order to identify arrestin subtype-specific sequence motifs, the consensus sequence of each alignment was determined using R (default cut-off 90% sequence identity; varying cut-offs to ensure robustness of identified residues). An alignment of the consensus sequences was used to identify residues (1) conserved across all arrestins (conserved in orthologous alignment and conserved in consensus alignment), (2) residues specifically conserved in an arrestin subtype (conserved in orthologous alignment, different in consensus alignment), and (3) residues that are neutrally evolving (not conserved in ortholog alignment). This allows to identification of residues that are specifically conserved in an arrestin subtype and allows, for instance, studying differences in visual and non-visual arrestins (Supplementary Data Set 2).

Link to interactive analysis of Arrestin orthologs and paralogs to identify subtype specific sequence patterns: <https://tilman.shinyapps.io/ArrestinSubtypeViewer/> The server shows an alignment of the consensus sequence of each arrestin subtype (obtained from an alignment of close orthologs), the human arrestin sequences and the overall arrestin consensus sequence. Colors denote different conservation types (Flock et al., Nature 2017). Threshold for calculating consensus sequence of orthologs and paralogs can be adjusted.

Bimane fluorescence assay

Bimane labeling of purified β arr1^{L68C} and subsequent fluorescence spectroscopy experiments have been described previously^{11,12}. The coding sequence of analogous β arr2 mutant for bimane assay, referred to as β arr2^{L69C}, was synthesized (Genscript) and sub-cloned in pGEX4T-1 plasmid. β arr2^{L69C} was purified essentially following the same protocol as for β arr1^{L68C} except that it was collected in the flow-through of the Q-sepharose anion-exchange chromatography, and it was bimane labeled using the same protocol as for β arr1^{L68C}. V₂Rpp was synthesized by the Tufts University Core Facility (USA) as described previously¹⁴. Activated and phosphorylated β_2 V₂R was purified using baculovirus mediated expression in Sf9 cells as described previously¹¹. For bimane fluorescence assay, labeled β arr1^{L68C} or β arr2^{L69C} proteins (in 20mM Hepes, pH 7.4, 100mM NaCl, 0.01% MNG) were mixed with either V₂Rpp or β_2 V₂R at 1:3 molar ratio, and incubated for 1h at room-temperature (25°C). Subsequently, bimane fluorescence was measured using a Fluorometer (Perkin Elmer, USA model LS-55) in photon counting mode as described previously¹¹. Data were normalized and plotted as indicated in the figure legend.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J. Standfuss for discussing results of previous mutagenesis experiments, and X. Deupi, D. Veprintsev, D. Meyer, G.F.X Schertler, S. Chavali, P. Lakshminarasimhan and H. Harbrecht for their comments on the manuscript. This work was supported by the Medical Research Council (MC_U105185859; M.M.B., T.F., A.S.) and the Boehringer Ingelheim Fond (T.F.). A.S. was funded by Wolfson College Research Grant, MRC Summer Studentship and The Lister Institute Summer Studentship. M.M.B. is a Lister Institute Research Prize Fellow and is supported by an ERC Consolidator Grant. Research program in AKS's laboratory is supported by an Intermediate Fellowship from the Wellcome Trust DBT India Alliance (IA/I/14/1/501285). T.F. is a Research Fellow of Fitzwilliam College, University of Cambridge, UK.

References

- Shukla AK, Xiao K, Lefkowitz RJ. Emerging paradigms of β -arrestin-dependent seven transmembrane receptor signaling. *Trends Biochem Sci.* 2011; 36:457–469. DOI: 10.1016/j.tibs.2011.06.003 [PubMed: 21764321]
- Lefkowitz RJ, Shenoy SK. Transduction of receptor signals by β -arrestins. *Science.* 2005; 308:512–517. DOI: 10.1126/science.1109237 [PubMed: 15845844]
- Kobilka BK. G protein coupled receptor structure and activation. *Biochim Biophys Acta.* 2007; 1768:794–807. DOI: 10.1016/j.bbame.2006.10.021 [PubMed: 17188232]
- Luttrell LM, Lefkowitz RJ. The role of β -arrestins in the termination and transduction of G-protein-coupled receptor signals. *J Cell Sci.* 2002; 115:455–465. [PubMed: 11861753]
- Komolov KE, et al. Structural and Functional Analysis of a β 2-Adrenergic Receptor Complex with GRK5. *Cell.* 2017; 169:407–421 e416. DOI: 10.1016/j.cell.2017.03.047 [PubMed: 28431242]
- DeWire SM, Ahn S, Lefkowitz RJ, Shenoy SK. β -arrestins and cell signaling. *Annu Rev Physiol.* 2007; 69:483–510. DOI: 10.1146/annurev.ph.69.013107.100021 [PubMed: 17305471]
- Gurevich VV, Gurevich EV. The molecular acrobatics of arrestin activation. *Trends Pharmacol Sci.* 2004; 25:105–111. DOI: 10.1016/j.tips.2003.12.008 [PubMed: 15102497]
- Nobles KN, et al. Distinct phosphorylation sites on the β (2)-adrenergic receptor establish a barcode that encodes differential functions of beta-arrestin. *Sci Signal.* 2011; 4:ra51.doi: 10.1126/scisignal.2001707 [PubMed: 21868357]
- Yang F, et al. Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and (19)F-NMR. *Nat Commun.* 2015; 6:8202.doi: 10.1038/ncomms9202 [PubMed: 26347956]
- Thomsen AR, et al. GPCR-G Protein- β -Arrestin Super-Complex Mediates Sustained G Protein Signaling. *Cell.* 2016; 166:907–919. DOI: 10.1016/j.cell.2016.07.004 [PubMed: 27499021]
- Kumari P, et al. Functional competence of a partially engaged GPCR- β -arrestin complex. *Nat Commun.* 2016; 7 13416. doi: 10.1038/ncomms13416
- Kumari P, et al. Core engagement with β -arrestin is dispensable for agonist-induced vasopressin receptor endocytosis and ERK activation. *Mol Biol Cell.* 2017; 28:1003–1010. DOI: 10.1091/mbc.E16-12-0818 [PubMed: 28228552]
- Cahill TJ 3rd, et al. Distinct conformations of GPCR- β -arrestin complexes mediate desensitization, signaling, and endocytosis. *Proc Natl Acad Sci U S A.* 2017; 114:2562–2567. DOI: 10.1073/pnas.1701529114 [PubMed: 28223524]
- Shukla AK, et al. Structure of active β -arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. *Nature.* 2013; 497:137–141. DOI: 10.1038/nature12120 [PubMed: 23604254]
- Shukla AK, et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature.* 2014; 512:218–222. DOI: 10.1038/nature13430 [PubMed: 25043026]
- Kang Y, et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature.* 2015; 523:561–567. DOI: 10.1038/nature14656 [PubMed: 26200343]
- Zhou XE, et al. Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. *Cell.* 2017; 170:457–469 e413. DOI: 10.1016/j.cell.2017.07.002 [PubMed: 28753425]

18. Ranjan R, Dwivedi H, Baidya M, Kumar M, Shukla AK. Novel Structural Insights into GPCR-beta-Arrestin Interaction and Signaling. *Trends Cell Biol.* 2017; 27:851–862. DOI: 10.1016/j.tcb.2017.05.008 [PubMed: 28651823]
19. Sukhwal A, Bhattacharyya M, Vishveshwara S. Network approach for capturing ligand-induced subtle global changes in protein structures. *Acta Crystallogr D Biol Crystallogr.* 2011; 67:429–439. DOI: 10.1107/S0907444911007062 [PubMed: 21543845]
20. Greene LH. Protein structure networks. *Brief Funct Genomics.* 2012; 11:469–478. DOI: 10.1093/bfgp/els039 [PubMed: 23042823]
21. Kornev AP, Haste NM, Taylor SS, Eyck LF. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A.* 2006; 103:17783–17788. DOI: 10.1073/pnas.0607656103 [PubMed: 17095602]
22. del Sol A, Fujihashi H, Amoros D, Nussinov R. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol Syst Biol.* 2006; 2 2006 0019. doi: 10.1038/msb4100063
23. Flock T, et al. Universal allosteric mechanism for Galpha activation by GPCRs. *Nature.* 2015; 524:173–179. DOI: 10.1038/nature14663 [PubMed: 26147082]
24. Venkatakrishnan AJ, et al. Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature.* 2016; 536:484–487. DOI: 10.1038/nature19107 [PubMed: 27525504]
25. Flock T, et al. Selectivity determinants of GPCR-G-protein binding. *Nature.* 2017; 545:317–322. DOI: 10.1038/nature22070 [PubMed: 28489817]
26. Venkatakrishnan AJ, et al. Molecular signatures of G-protein-coupled receptors. *Nature.* 2013; 494:185–194. DOI: 10.1038/nature11896 [PubMed: 23407534]
27. Chen Q, et al. Structural basis of arrestin-3 activation and signaling. *Nat Commun.* 2017; 8 1427. doi: 10.1038/s41467-017-01218-8
28. Gurevich VV, Gurevich EV. The structural basis of arrestin-mediated regulation of G-protein-coupled receptors. *Pharmacol Ther.* 2006; 110:465–502. DOI: 10.1016/j.pharmthera.2005.09.008 [PubMed: 16460808]
29. Ostermaier MK, Peterhans C, Jaussi R, Deupi X, Standfuss J. Functional map of arrestin-1 at single amino acid resolution. *Proc Natl Acad Sci U S A.* 2014; 111:1825–1830. DOI: 10.1073/pnas.1319402111 [PubMed: 24449856]
30. Hirsch JA, Schubert C, Gurevich VV, Sigler PB. The 2.8 Å crystal structure of visual arrestin: a model for arrestin's regulation. *Cell.* 1999; 97:257–269. [PubMed: 10219246]
31. Hanson SM, et al. Differential interaction of spin-labeled arrestin with inactive and active phosphorhodopsin. *Proc Natl Acad Sci U S A.* 2006; 103:4900–4905. DOI: 10.1073/pnas.0600733103 [PubMed: 16547131]
32. Vishnivetskiy SA, et al. Few residues within an extensive binding interface drive receptor interaction and determine the specificity of arrestin proteins. *J Biol Chem.* 2011; 286:24288–24299. DOI: 10.1074/jbc.M110.213835 [PubMed: 21471193]
33. Zhuo Y, Vishnivetskiy SA, Zhan X, Gurevich VV, Klug CS. Identification of receptor binding-induced conformational changes in non-visual arrestins. *J Biol Chem.* 2014; 289:20991–21002. DOI: 10.1074/jbc.M114.560680 [PubMed: 24867953]
34. Peterhans C, Lally CC, Ostermaier MK, Sommer ME, Standfuss J. Functional map of arrestin binding to phosphorylated opsin, with and without agonist. *Sci Rep.* 2016; 6 28686. doi: 10.1038/srep28686
35. Miskei M, et al. Fuzziness enables context dependence of protein interactions. *FEBS Lett.* 2017; 591:2682–2695. DOI: 10.1002/1873-3468.12762 [PubMed: 28762260]
36. van der Lee R, et al. Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep.* 2014; 8:1832–1844. DOI: 10.1016/j.celrep.2014.07.055 [PubMed: 25220455]
37. Gurevich VV, Benovic JL. Visual arrestin interaction with rhodopsin. Sequential multisite binding ensures strict selectivity toward light-activated phosphorylated rhodopsin. *J Biol Chem.* 1993; 268:11628–11638. [PubMed: 8505295]

38. Gurevich VV, Benovic JL. Visual arrestin binding to rhodopsin. Diverse functional roles of positively charged residues within the phosphorylation-recognition region of arrestin. *J Biol Chem.* 1995; 270:6010–6016. [PubMed: 7890732]
39. Robinson KA, et al. The effect of phosphorylation on arrestin-rhodopsin interaction in the squid visual system. *J Neurochem.* 2015; 135:1129–1139. DOI: 10.1111/jnc.13366 [PubMed: 26375013]
40. Galliera E, et al. beta-Arrestin-dependent constitutive internalization of the human chemokine decoy receptor D6. *J Biol Chem.* 2004; 279:25590–25597. DOI: 10.1074/jbc.M400363200 [PubMed: 15084596]
41. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM. MUSTANG: a multiple structural alignment algorithm. *Proteins.* 2006; 64:559–574. DOI: 10.1002/prot.20921 [PubMed: 16736488]
42. Finn RD, et al. HMMER web server: 2015 update. *Nucleic Acids Res.* 2015; 43:W30–38. DOI: 10.1093/nar/gkv397 [PubMed: 25943547]
43. Liu Y, Schmidt B, Maskell DL. MSAProbs: multiple sequence alignment based on pair hidden Markov models and partition function posterior probabilities. *Bioinformatics.* 2010; 26:1958–1964. DOI: 10.1093/bioinformatics/btq338 [PubMed: 20576627]
44. Granzin J, et al. Crystal structure of p44, a constitutively active splice variant of visual arrestin. *J Mol Biol.* 2012; 416:611–618. DOI: 10.1016/j.jmb.2012.01.028 [PubMed: 22306737]
45. Granzin J, et al. X-ray crystal structure of arrestin from bovine rod outer segments. *Nature.* 1998; 391:918–921. DOI: 10.1038/36147 [PubMed: 9495348]
46. Han M, Gurevich VV, Vishnivetskiy SA, Sigler PB, Schubert C. Crystal structure of beta-arrestin at 1.9 Å: possible mechanism of receptor binding and membrane Translocation. *Structure.* 2001; 9:869–880. [PubMed: 11566136]
47. Hwang J, et al. The structural basis for the negative regulation of thioredoxin by thioredoxin-interacting protein. *Nat Commun.* 2014; 5 2958. doi: 10.1038/ncomms3958
48. Kang DS, et al. Structure of an arrestin2-clathrin complex reveals a novel clathrin binding domain that modulates receptor trafficking. *J Biol Chem.* 2009; 284:29860–29872. DOI: 10.1074/jbc.M109.023366 [PubMed: 19710023]
49. Kim YJ, et al. Crystal structure of pre-activated arrestin p44. *Nature.* 2013; 497:142–146. DOI: 10.1038/nature12133 [PubMed: 23604253]
50. Milano SK, Kim YM, Stefano FP, Benovic JL, Brenner C. Nonvisual arrestin oligomerization and cellular localization are regulated by inositol hexakisphosphate binding. *J Biol Chem.* 2006; 281:9812–9823. DOI: 10.1074/jbc.M512703200 [PubMed: 16439357]
51. Milano SK, Pace HC, Kim YM, Brenner C, Benovic JL. Scaffolding functions of arrestin-2 revealed by crystal structure and mutagenesis. *Biochemistry.* 2002; 41:3321–3328. [PubMed: 11876640]
52. Polekhina G, et al. Structure of the N-terminal domain of human thioredoxin-interacting protein. *Acta Crystallogr D Biol Crystallogr.* 2013; 69:333–344. DOI: 10.1107/S0907444912047099 [PubMed: 23519408]
53. Qi S, O'Hayre M, Gutkind JS, Hurley JH. Insights into beta2-adrenergic receptor binding from structures of the N-terminal lobe of ARRDC3. *Protein Sci.* 2014; 23:1708–1716. DOI: 10.1002/pro.2549 [PubMed: 25220262]
54. Sutton RB, et al. Crystal structure of cone arrestin at 2.3 Å: evolution of receptor specificity. *J Mol Biol.* 2005; 354:1069–1080. DOI: 10.1016/j.jmb.2005.10.023 [PubMed: 16289201]
55. Zhan X, Gimenez LE, Gurevich VV, Spiller BW. Crystal structure of arrestin-3 reveals the basis of the difference in receptor binding between two non-visual subtypes. *J Mol Biol.* 2011; 406:467–478. DOI: 10.1016/j.jmb.2010.12.034 [PubMed: 21215759]
56. Doncheva NT, Klein K, Domingues FS, Albrecht M. Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci.* 2011; 36:179–182. DOI: 10.1016/j.tibs.2011.01.002 [PubMed: 21345680]
57. Kayikci M, et al. Visualization and analysis of non-covalent contacts using the Protein Contacts Atlas. *Nat Struct Mol Biol.* 2018; 25:185–194. DOI: 10.1038/s41594-017-0019-z [PubMed: 29335563]
58. Babu MM. NCI: A server to identify non-canonical interactions in protein structures. *Nucleic Acids Res.* 2003; 31:3345–3348. [PubMed: 12824323]

59. Altenhoff AM, et al. The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces. *Nucleic Acids Res.* 2018; 46:D477–D485. DOI: 10.1093/nar/gkx1019 [PubMed: 29106550]

Figure 1. The arrestin signalling cycle.

a, (left) Inactive arrestin is present in the cytosol in the absence of external signals. (centre) Phosphorylation of agonist-activated receptors by GRKs stimulates the formation of the C-tail engaged arrestin-GPCR complex. (right) High affinity binding of arrestin requires formation of the fully-engaged GPCR-arrestin complex. Arrestin binding to a receptor acts to preclude G protein signalling (desensitization), but also mediates arrestin-dependent signaling pathways. Activated arrestin can expose clathrin-binding sites, which can lead to receptor internalization. **b**, Structure of inactive-state arrestin highlighting the three-element interaction, polar core and hydrophobic inter-domain interactions (PDB ID: 1cf1, chain D; side view: left; top view right).

Figure 2. Systematic comparison of arrestin structures.

a, Common Arrestin Numbering (CAN) system. (*left*) In order to compare different arrestins, each arrestin residue is described in three levels: Domain (D), secondary structure element (S) and position within the secondary structure element (P). (*right*) The CAN was derived through a structural alignment of all arrestin structures, from which the ‘consensus secondary structure’ was determined. **b**, Consensus residue contact network analysis of arrestins. Non-covalent interactions between all residues in each arrestin structure have been calculated (‘contact fingerprints’) to group arrestin structures according to their signalling state (fingerprint clustering, see Supplementary Fig. 1). For each group/signalling state, the fingerprints can be transformed into consensus contact networks, where nodes represent amino acid positions according to the CAN, and edges represent non-covalent contacts present in that state. Sequence data are used to identify nodes (residues) that are conserved. **c**, Number of contacts that are unique or shared between the different signalling states of arrestin, which are shown in Fig 1.

Figure 3. An ionic lock keeps the finger-loop in the inactive state.

a, Sequence analysis of the finger-loop lock motif shows that all residues are highly conserved. **b**, Residue contact network of the finger-loop lock. Dashed lines indicate that either both or at least one salt bridge is present in the inactive state. **c**, Structures of finger loop in the inactive and active state. In the inactive state, the finger loop is folded ('locked') whereas in the active state it adopts an 'open' (extended) conformation.

Figure 4. Analysis of arrestin residues that interact with receptor-attached phosphates. Three structural motifs (bottom – three-element interaction; middle – polar core; top – the here identified finger-loop lock) can induce conformational change on arrestin upon interaction with the phosphorylated receptor tail. For each structural motif, the inactive and active state consensus network, as well as a structure representation of the active state, is shown (circles – arrestin residues; red or blue – negatively or positively charged, respectively; orange squares – receptor phosphates). The effect of alanine mutations of motif residues on formation of the receptor/arrestin complex are shown as bar chart (right panel). Arrestin residues are sorted based on the position of the phosphate they interact with (top – proximal, closer to receptor helix 8; bottom – distal from helix 8; see left panel).

Figure 5. Summary of phosphorylation-sensitive motifs for arrestin activation.

a, Different phosphates on the GPCR C-tail can preferentially interact with different structural motifs on arrestin: the ‘finger-loop lock’, and the previously reported polar core and three element (TE) interaction. Panel 1 (top left) shows a schematic representation of the finger loop lock: negative charges of the finger loop interact with positive charges in strand 5 and 6; Panel 2 (top right) summarizes changes in the polar core and the three-element interaction induced by receptor-attached phosphates. **b-c**, The distribution of these motifs in arrestin suggests that different phosphorylation patterns on the receptor tail can interact with different structural motifs on arrestin and activate arrestin with different efficiencies or induce different conformations.