Neurotransmitter 2016; 3: e1273. doi: 10.14800/nt.1273; © 2016 by Marcelo O. Ortells

http://www.smartscitech.com/index.php/nt

Special Issue Edited by Dr. Hugo R. Arias | Functional and therapeutic aspects of the nicotinic acetylcholine receptor family.

REVIEW

Structure and function development during evolution of pentameric ligand gated ion channels

Marcelo O. Ortells

Faculty of Medicine, University of Morón and CONICET, Machado 914, 4to piso, 1708 Morón, Argentina

Correspondence: Marcelo O. Ortells
E-mail: mortells@retina.ar
Received: March 27, 2016
Published online: April 22, 2016

Ligand Gated Ion Channels (LGICs) are receptors widely involved in vertebrate neuronal and non-neuronal signalling as well as muscle activation. Until very recently these receptors, were only known to exist in animal cells and were named the Cys-loop receptor superfamily because they all share this sub-substructure in their amino acid sequences. Nowadays, these receptors are known as "pentameric" LGICs (pLGICs), to differentiate from other LGICs such us tetrameric glutamate or trimeric ATP LGICs. The first phylogenetic analyses of pLGICs have been performed almost 20 years ago, where only eukaryotic members were known. Currently, we have many more sequences from invertebrate species, but most importantly, we know that these receptors are present in a variety of prokaryotic organisms, a fact predicted in those early studies. Equally important, we have information about the tertiary and quaternary structures from prokaryotic to eukaryotic members, which enlightened our evolutionary understanding of these fascinating and functionally important proteins. Surprisingly, the prokaryotic receptors lack the Cys-loop characteristic of eukaryote receptors which once gave the name to the superfamily. This review focuses on the history of the evolutionary processes of these receptors involved in very important physiological functions as well as in many diseases such as tobacco addiction and schizophrenia, to name just a few.

Keywords: Assembly; Desensitization; Nicotinic receptors; Phylogenetics; pLGICs; Receptor evolution; Structure evolution

To cite this article: Marcelo O. Ortells. Structure and function development during evolution of pentameric ligand gated ion channels. Neurotransmitter 2016; 3: e1273. doi: 10.14800/nt.1273.

Introduction

Ion flux through excitable cellular membranes is an ubiquitous mechanism used in cell to cell signalling, widely utilized by members of all major superkingdoms. Ion flux is mediated by a broad variety of ion channels, gated by a diversity of signals including voltage changes, neurotransmitters, and mechanical forces. This receptor signalling mechanism is essential for the functioning of animal nervous systems. The fast propagation of the signal from neuron to neuron or to the target cells (i.e. muscle cells) is mediated by the activation of ion channels.

The Ligand Gated Ion Channel (LGIC) superfamily is the best known receptor superfamily, predominantly due to the comprehensive characterisation of the nicotinic acetylcholine (ACh) receptor (nAChR), which is the paradigm for the whole LGIC superfamily [1]. In vertebrates, nAChRs are

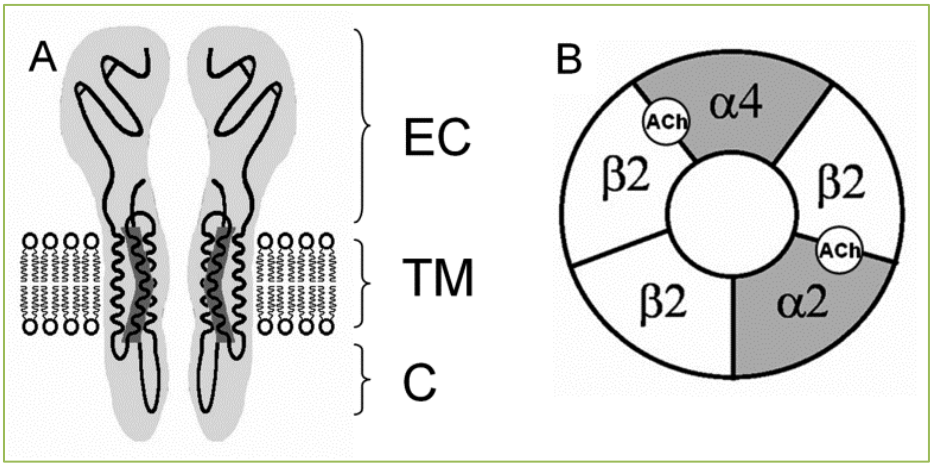

Figure 1. Schematic representation of the overall structure of pLGICs. A. View of "two" subunits parallel to the membrane. The M2 segment which delineates the pore is greyed. ECD: Extracellular domain; TMD: Transmembrane domain; CL: Cytoplasmic loop. B. Schematic view of the ECD from the α4β2 nAChR, as an example, showing the two acetylcholine (ACh) binding sites. This particular stoichiometry, $(\alpha4)_2(\beta2)_3$, has two α subunits and therefore two agonist binding sites.

present mainly in the nervous system but they are also expressed in non-neuronal tissues, where, for example, modulate the angiogenesis process. There are two nAChR types, neuronal and muscle receptors, the latter are also expressed in electric organs from fish such as *Torpedo sp.* and *Electrophorus electricus*.

LGICs combine the property of selectively binding a neurotransmitter with the subsequent opening of an ion channel. When a neurotransmitter is released from the presynaptic cell, it binds to the extracellular domain (ECD) present in all LGICs, and causes the resting (closed) ion channel embedded in these receptors to open for a short period of time. At the ECD is located the "Ligand Binding Domain" (LBD) where neurotransmitters selectively bind. After the neurotransmitter-induced activation, ions flux through the channel. Depending on the receptor subtype, anion influx (e.g., Cl^-) induces membrane hyperpolarization and consequently neuronal inhibition, whereas cation influx (e.g., Na^+ and Ca^{2+}) induces membrane depolarization and consequently neuronal stimulation. After a brief moment of activation, the channel becomes "desensitized", a closed state different (i.e., closed but not activatable by agonists) to that for the resting state (i.e., closed but activatable by agonists). Afterwards, the neurotransmitter is released from its binding sites or breakdown by enzymes (e.g., the acetylcholinesterase metabolizes ACh), and the receptor goes back to the "resting" closed state waiting for the next activation. This is the classic scenario seen in vertebrate receptors. A clear knowledge of how the action of an agonist ends after receptor activation is still missing, but a good review on this topic is presented by Corringer *et al*. [5].

The LGICs is now referred as pentameric LGICs (pLGICs) to differentiate them from the glutamate (Glu) receptor family which are tetramers and with the trimeric receptor channels for ATP. The pLGIC receptors are not only widely distributed in phylogenetic terms, but they also exist in metazoans, from nematodes and insects to all vertebrates, and, as we shall see, in prokaryotes. Interestingly, these receptors have a great variability of subunits, which is a key element for the study of these receptors from an evolutionary point of view. Sequence information reveals that the subunits (~400-500 amino acids) for any given pentameric member are themselves homologous originated from gene duplications. These subunits have an N-terminal ECD, bearing the ligand binding sites (Figure 1A). The binding pocket for the agonist is situated between two subunits. The transmembrane domain (TMD) is composed of four transmembrane α-helical segments (M1-M4), and finally a short extracellular C-terminus. Frequently in animals, a large cytoplasmic loop is present between transmembrane regions M3 and M4. The M2 transmembrane helix is important as the main structural component of the ion channel itself, which is composed by five M2s, one from each monomer (see below for more details).

The neurotransmitters that activate the members of this receptor superfamily include those found in vertebrates such as ACh, serotonin [i.e., 5-hydroxytryptamine (5-HT)], glycine (Gly), and γ-aminobutyric acid (GABA), as well as those found in non-vertebrates such as Glu and histamine. The general rule is that the nAChR and type 3 5-HT receptors (5-HT_3Rs) are selective for cations and hence excitatory. A new member of the cation family, the ZAC channel (i.e., Zinc activated ion-channel) was recently

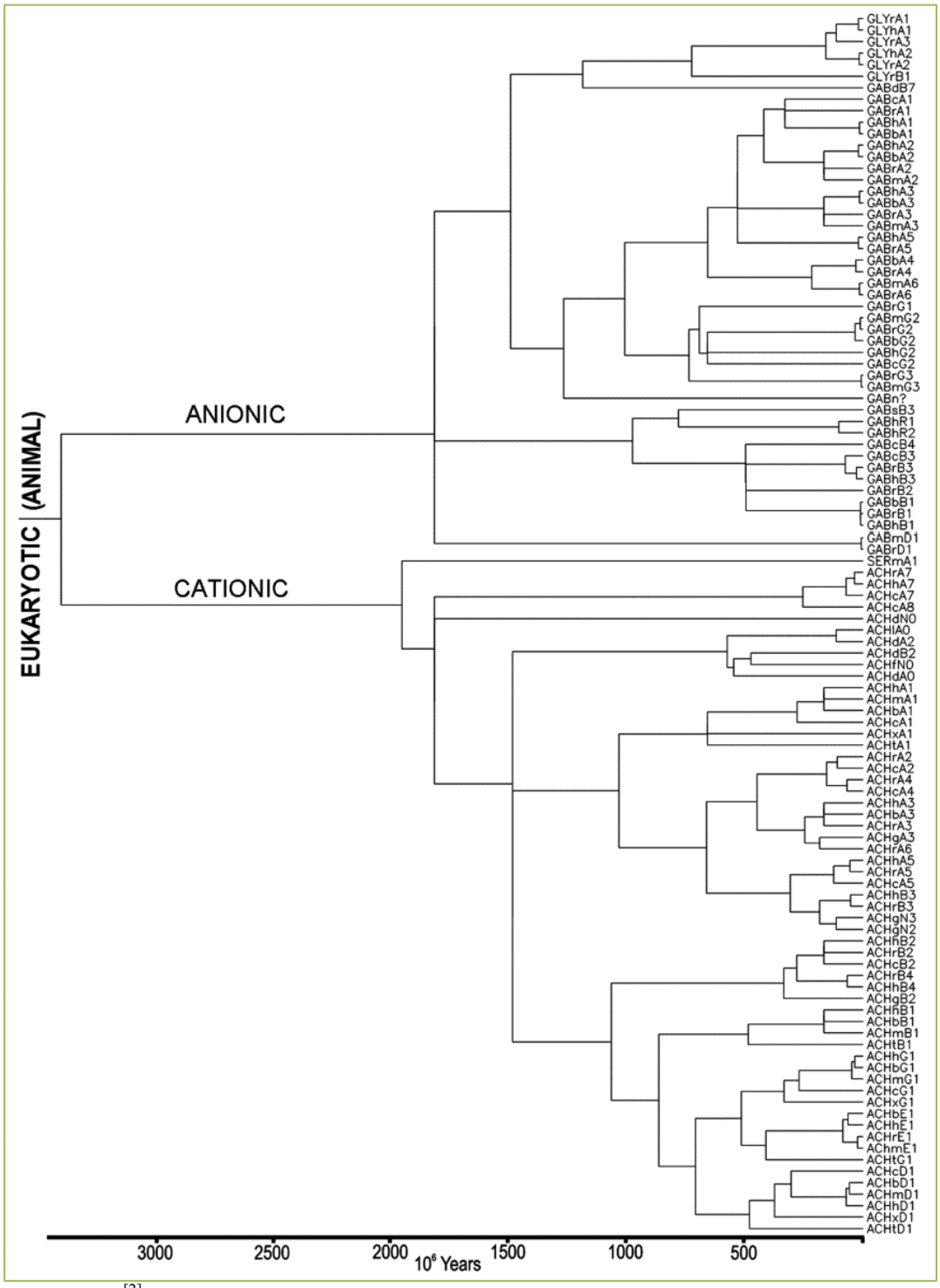

Figure 2. Phylogenetic tree of the animal pLGIC family assuming a molecular clock. The calibration (modified from Ortells and Lunt [1]) from relative to absolute time scaling was based on the fossil record for the average time of divergence of the lineages leading to: mammals and birds (approximately 300 My ago), mammals-birds and amphibians (approximately 350 My ago), and fish and the remaining vertebrates (approximately 430 My ago [74]. A six characters nomenclature was used to discriminate the receptor type and species of origin (i.e., RRRsS#), where "RRR" indicates the receptor type (i.e., ACH = nAChR; GAB = GABAR; GLY = GlyR; SER = 5-HT3R), "s" the organism (i.e., b = bovine; c = chicken; d = *Drosophila*; f = filaria; g = goldfish; h = human; l = locust; m = mouse; n = nematode; r = rat; s = snail; t = *Torpedo*; and x = *Xenopus*), "S" the subunit type (i.e., A = α; B = β; G = γ; D = δ; E = ε; R = ρ; N = non-α; ? = undetermined), and # the subunit number (where 0 is undetermined).

discovered [2]. It is present in human and dogs, but absent from mouse or rats and its physiological function is still unknown. In addition, $GABA_AR$ and $GABA_CR$, GluR s, and GlyRs are selective for anions, and consequently they are inhibitory. But there are some exceptions to this classification. For example, some invertebrate GABARs are excitatory, whereas some ACh/5-HT gated receptors permeate anions (see below). This different activity is because the type of ion that flows through the channel is not determined by the agonist type but it is mainly ruled by the charge distribution in the ion channel [3-4].

pLGICs are also targets for numerous toxins and venoms, such as nicotine, strychnine, snail conotoxins, coral lophotoxins, and many snake venoms [6]. Nicotine is a substance present in tobacco that acts as a toxin in insects, and as the main addictive compound in humans. In the brain, nicotine directly stimulates cholinergic (nicotinic) neurons, and indirectly, via presynaptic AChR-induced neurotransmitter release, many others circuits such as the

glutamatergic, GABAergic, and dopaminergic neurons, a characteristic by which tobacco addiction is based on [7-8].

pLGIC members not only have a high degree of amino acid sequence homology, but also share similar tertiary and quaternary structures. Within each receptor family, subunits are split in different types named α, β, γ, and so on. Particularly, the nAChR subunits are basically classified as α and non-α, based on the presence of a pair of adjacent cysteines situated in the so called ¨Loop C", an important tertiary substructure involved in ligand binding (see below). nAChR α subunits are necessary to bind the neurotransmitter ACh or other agonists such as nicotine (Figure 1B). A characteristic of the whole animal superfamily is a sequence motif of 15 residues named the Cys-loop in the LBD. However, recent discoveries on prokaryotic members of this superfamily (see below) showed exceptions to this paradigm [9].

Phylogenetic trees

The first studies on the phylogenetics of pLGIC receptors appeared 20 years ago. The work by Ortells and Lunt [1] presented, based on the knowledge at that time, an evolutionary tree of the whole pLGIC superfamily taking account their nucleotide sequences. The other work by Le Novére and Changeux [10] presented three different evolutionary trees for only the nAChR family; one tree based on amino acid sequence information, a second based on the structure of the nAChR genes, and the third as a tentative consensus between them.

Ortells and Lunt [1] employed an alignment of 106 amino acid sequences of pLGIC receptors known at that time as the starting point for the construction of their evolutionary tree (Figure 2). Branch lengths of Figure 2 were calculated assuming a molecular clock. The use and fundamentals of the molecular clock hypothesis has been and is still controversial. It was discovered that certain proteins have a constant rate of amino acid substitutions among several mammalian lineages. Consequently, it was suggested [11] that the rate of molecular evolution of a given protein is almost constant over time, hence the term "molecular clock". To estimate divergence times using molecular information, a known time scale is needed to convert relative molecular distances to real time. Fossil information was employed for this purpose.

The main difference between the DNA and amino acid tree, was the topology involving the muscle nAChRs. For Ortells and Lunt [1], the α1 subunit belongs to the α clade that includes other neuronal α subunits (i.e., α2-α10), consequently this clade is monophyletic (a clade is a group of organisms or genes which have their origin from a common ancestor, i.e. a monophyletic group). Likewise, the muscle-type β1 subunit, the remaining non-α muscle subunits (i.e., ε, γ, and δ), and additional neuronal β subunits (i.e., β2-β4) belong to the β clade (Figure 2). On the other hand, based on the gene exonic structure (which refers to the relationship between exons and introns within the complete gene, in terms of their positions and sizes) Le Novére and Changeux [10] concluded that all muscle subunits, α and non-α (i.e., α1, β1, γ, ε, and δ) have a common ancestor derived from a muscle α-like subunit. More recently, Dent [12] found, however, that the muscle and neuronal α subunits form a monophyletic group (see below), and the muscle and neuronal non-α subunits form another group, a conclusion similar to that arrived by Ortells and Lunt [1]. Nevertheless, within the neuronal α group there are reversions to non-α, but only for those that lost the pair of cysteines but still conserve a high degree of sequence similarity with the rest of the α subunits.

Figure 2 shows important fundaments on subunit variability: there is variability of the same subunit type among different species, which, together with fossil record ages, allowed to calibrate the times of split, and consequently the time length for the whole tree. When the DNA based tree was built, no outgroup was available. By rooting it in the middle of its length the pLGICs' ancestor was placed between animal cationic and anionic receptors. Assuming a molecular clock like this, the date estimated for this ancestor is at least 2,500 million years (My) ago [1]. The first impression was that this was a surprisingly remote origin, probably before the first eukaryotes [13]. However, at the time of publication, that seemed not to be an isolated case. In spite of the lack of sequence similarity, G-coupled protein receptors, another major group of cell surface signalling proteins, were known to have a tertiary structure similar to bacteriorhodopsin [14] and hence, probably homologous to what is clearly a prokaryotic protein. Such considerations suggested that these very important surface signalling molecules associated with present day nervous systems were readily available well before this novel signalling function made its appearance during evolution [1]. More recently the same was observed for the voltage-gated potassium channels of the Shaker-type superfamily and the voltage-gated sodium channels [15-18].

An earlier study [19] suggested that pLGIC like proteins may be widely represented in a variety of organisms and the ancestral role of primitive pLGIC receptors was discussed in the context of osmotic regulation and nutrient seeking, both of which may involve transmembrane ion fluxes and ligand recognition. Nutrient seeking in particular may relate to the

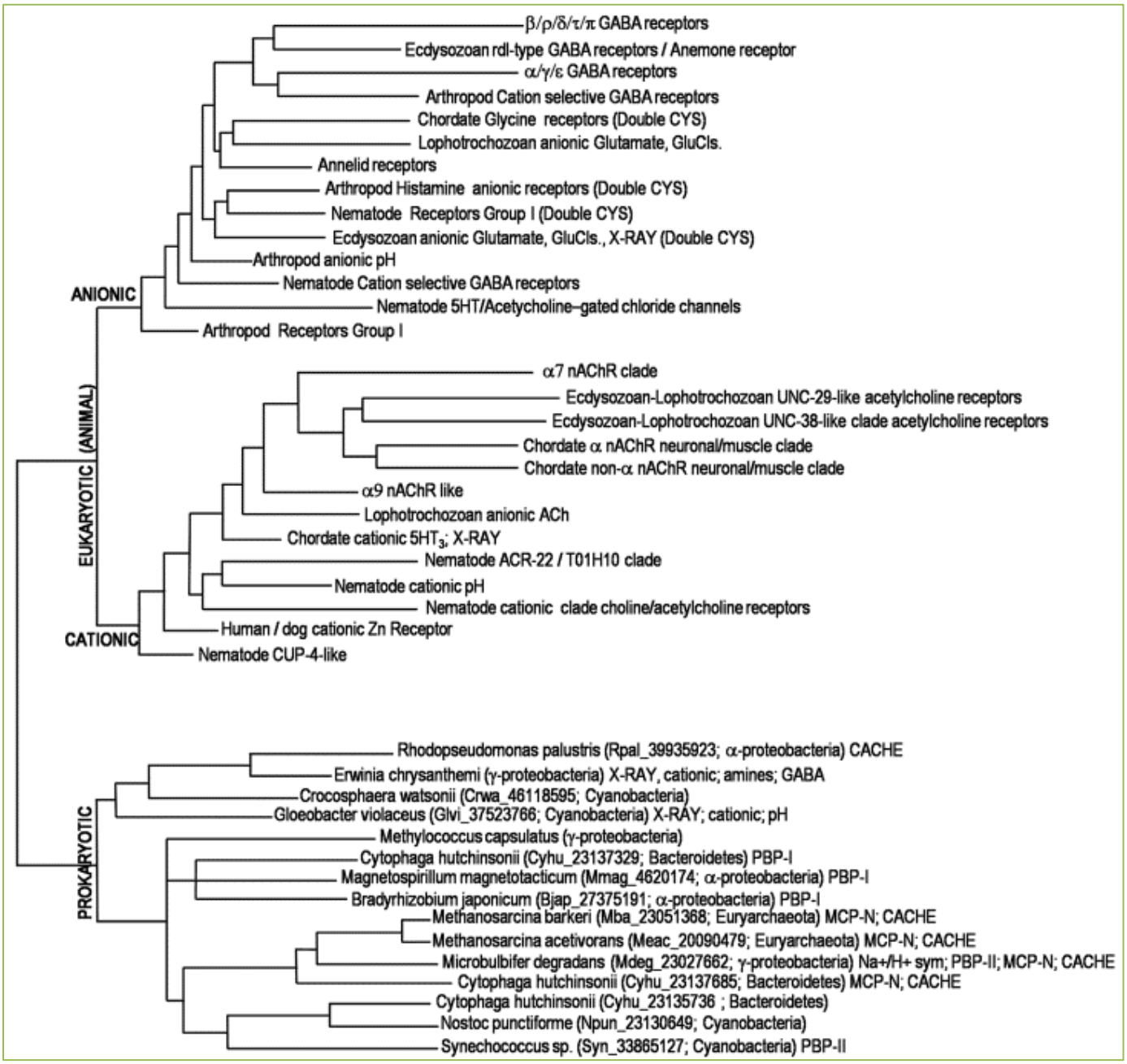

Figure 3. Phylognetic tree of the whole and known pLGIC superfamiliy, including eukaryotic and prokaryotic members. Each clade includes all species members, as a difference with Figure 2. When known or suspected, the agonist and type of ion that permeates the channel is indicated. Double-Cys: member of the double-Cys loop subfamily, according to Dent (2006). X-Ray indicates that there is at least a member of the clade with a high resolution tertiary and quaternary structure solved. Proteins linked to prokaryotic pLGICs N-terminal domains: PBP-I, Periplasmic binding protein type I; PBP-II, Periplasmic binding protein type II; MCP-N, Methyl-accepting chemotaxis protein-N; CACHE, Ca2+ channels and chemotaxis receptors; Sym: symporter. To simplify nomenclature "Prokaryotes" includes modern classification of Bacteria (Eubacteria) and Archea

present function of pLGICs since their ECM is a highly complex molecular recognition system that would not be needed for osmotic regulation [1], a function that might relate more probably to that of voltage-dependent ion channels.

The prediction of a pre-eukaryotic origin for this superfamily [1] was confirmed when sequences of putative pLGIC orthologs were identified in genomes of bacterial species and in the archaea *Methanosarcina* [9]. In his phylogenetic analysis of the pLGIC family, Dent [12, 20] included members from early Bilateria and many sequences not included in previous works [1, 9, 10], such as the sequences from the primitive chordate *Ciona intestinalis*, Ecdysozoans represented by arthropods (insects) and nematodes,

Lophotrochozoans including mollusks and annelids, and anemone. As mentioned, Dent [12, 20] agreed with the conclusions reached by Ortells and Lunt [1] regarding the monophyletic origins of the α and non-α clades. However, he also concluded that the anionic GlyRs do not belong to the GABAR clade (Figure 2), as suggested by Ortells and Lunt [1], but are sister branches. A synoptic tree including prokaryotic and eukaryotic pLGIC receptors is presented in Figure 3.

Dent [12, 20] included, in addition to the normal chordate excitatory nAChRs and 5-HT_3Rs as well as inhibitory GlyRs and GABARs, a wide array of invertebrate pLGIC receptors. There are also eukaryotic pLGIC receptors not activated by the traditional neurotransmitters (Figure 3), including inhibitory histamine receptors [21-22] and GluRs [23-26]. In both "cationic" and "anionic" branches there are pLGICs activated by pH changes. For human and mouse, a ZAC channel is also included. Interestingly, Figure 3 also shows that many clades have receptors for which the normal agonist or effect (excitatory or inhibitory) is still unknown.

In the traditional anionic receptor family appeared some members that, although genetically related, allow cations instead of anions influx, including the non-orthologous nematode [27] and arthropode [21] cation selective GABARs (Figure 3). Likewise in the cation family, Lophotrochozoans have an anionic receptor activated by ACh [28]. In the anion family there are receptors whose agonists usually activate cationic receptors, such as the "inhibitory" ACh [29] and 5-HT [30] gated chloride channels.

Dent [12] proposed for the eukaryotic anionic family, a subfamily of "double Cys-loop" receptors, which included the chordate GlyRs and three non-chordate clades. However, in the tree presented later [20], Dent did not mention it and according to his analysis it appears either not to be monophyletic or the double Cys-loop was lost in many clades (Figure 3).

Figure 3 also shows the phylogenetic relationships of the different pLGICs found in Eubacteria (Bacteria) and Archea (formerly both classified and merged as Prokaryotes). A very interesting fact is that in Eukaryota, pLGICs are only found in animals (Metazoans or Animalia), but they could not be detected in other branches for which genomic sequences are available such as fungi, plants, *Entamoeba*, nor in ancient organisms like Trichomonas. Different are the cases of the voltage-gated potassium and sodium channels already mentioned, where representatives are known from both animal and non-animal Eukaryota as well as from numerous prokaryotes [15-18].

In prokaryotes, pLGICs are spread in a very complex way. They appear in few species, and quite unrelated and distant taxa possess very similar receptors (i.e., *Cytophaga* and the archaeon *Methanosarcina*), whilst phylogenetically related taxa may have or not a pLGIC receptor. Tasneem *et al*. [9] suggested that this pattern, similar to those detected in other signalling proteins in prokaryotes, is derived from a high degree of mobility through lateral transfer and frequent gene loss [31]. The phylogenetic tree showed in Figure 3 is congruent with that displayed in Figure 2, which was lacking an outgroup, in that the animal branch is monophyletic and splits up in a similar fashion. According to the tree displayed in Figure 3, the group composed of *Crocosphaera*, *Gloeobacter*, *Erwinia*, and *Rhodopseudomonas* is clearly more similar in sequence to the eukaryotic pLGIC receptors. Thus, the ancestral gene of the animal branch for pLGICs might have been acquired only once by lateral transfer [9] from one of this prokaryotic forms.

Structure evolution

Although the sequence similarity among pLGICs can be as low as 10%, the analysis of the global alignment of the subunit sequences shows that there are several key residues that are extremely conserved in all pLGICs. Not only these conserved residues made possible the alignment but also suggested a common tertiary structure for the whole superfamily.

Subunit sequences from Eukaryotic pLGICs (i.e., Metazoans or Animalia, as it is the only kingdom where pLGICs were found so far) are much closer to each other than those from Eubacteria and Archea pLGICs. Although animal subunit sequences possess the characteristic Cys-loop which once gave the name to this superfamily, this important substructure is lacking in the prokaryotic counterparts. Interestingly, in replacement of the second Cys, prokaryotic subunits possess a conserved hydrophobic residue, which probably stabilizes in a similar fashion the region corresponding to the Cys-loop in eukaryotic subunits [9]. In addition, the large cytoplasmic region located between M3 and M4 observed in animal pLGICs is lacking in prokaryotes. In animal pLGICs, this region has been implicated in protein sorting, trafficking, cytoskeleton anchoring, and membrane insertion [32].

The TMDs are the most conserved receptor domains since they have strong structural restrictions: to anchor the receptor to the lipid membrane as well as to open and close the ion channel. Sequence analysis of this region predicted a typical four-helix bundle quaternary structure. The LBD is also well conserved since it has the important function of translating the binding of an agonist into conformational changes that are transferred to the TMD to finally open the pore.

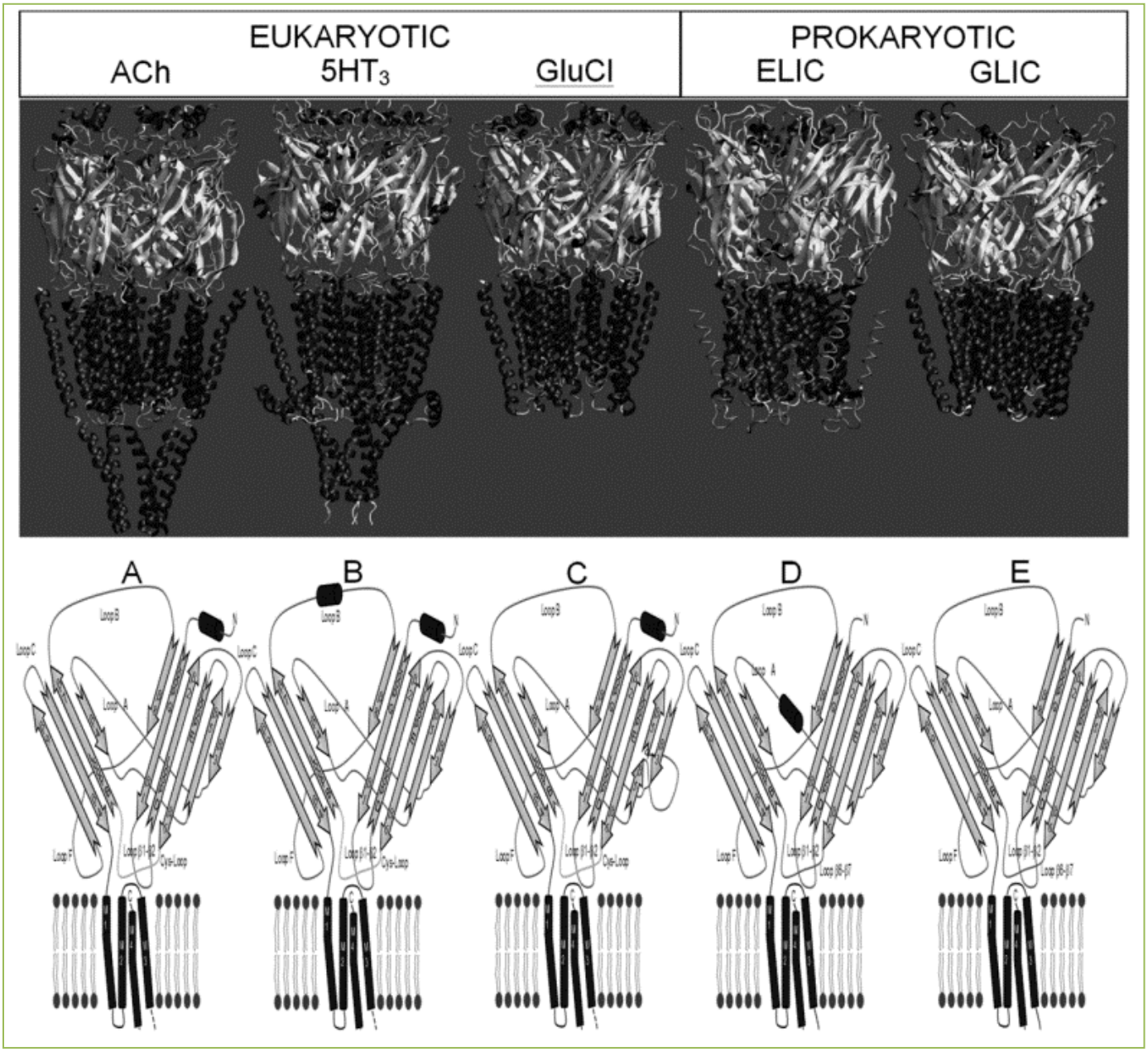

Figure 4. Tertiary and quaternary structures of pLGICs. A. *Torpedo* nAChR; B. Mouse 5-HT3R; C. *Caenorhabditis elegans* Glutamate chloride receptor. D. *Erwinia chrysanthemi* GABA/amine cationic receptor. E. *Gloeobacter violaceus* proton activated cationic receptor. Upper row: ribbon representation of medium or high resolution structures. Secondary structures are coloured in black (α-helices), white (β-sheets) and gray (loops). Lower row: cartoon representation of the subunit secondary structures. Cylinders represent α-helices and arrows β-sheets. The nomenclature of each secondary structure is specified. Some β-sheets where originally considered "loops", so both nomenclatures are shown.

Tertiary and quaternary structure images of members from this superfamily were obtained with increasing atomic resolution. The difficulty for years in obtaining crystals for X-ray analysis came from the fact that two of the three domains (extra and intracellular) are water soluble while the transmembrane region is hydrophobic. This problem wad solved, for example, thanks to the use of novel techniques based on llama-derived single-chain antibodies as crystallization chaperones. Currently, there are several solved structures which are shown in the upper row of Figure 4A-E.

The initial nAChR images (Figure 4A) at medium resolution (4-9 Å) were obtained from electron microscopy diffraction studies of quasi-crystalline nAChR specimens from the electric ray *Torpedo californica* [33-34]. The first high resolution images were obtained using soluble acetylcholine binding proteins (AChBPs) from the molluscs *Lymnea stagnalis* [35], *Aplysia californica* [36], and *Capitella teleta* [37],

which have structural resemblance to the extracellular domains of pLGICs. All AChBPs lack the transmembrane and intracellular regions. These proteins permitted to clarify the structural components of the agonist binding sites since Loop E and Loop D were actually made of residues from adjacent β strands and not from loop structures.

In 2007, it was discovered that the expression of the coded protein (GLIC, Figures 3 and 4E) from one of the prokaryotic species, *Gloeobacter violaceus*, produced functional cationic pLGICs activated by protons [38]. Short after, the first complete high-resolution image of a full-length pLGIC, from *Erwinia chrysanthemi* (ELIC, Figures 3 and 4D), was solved at 3.3 Å in the closed and putative open states [39-40]. The ECDs of these channels were similar to those found in AChBPs. These receptors have been shown to be activated by primary amines and by high concentrations of GABA [41]. These first studies were followed by the structure solution of GLIC at 2.9 Å resolution [42]. Regarding eukaryotic receptors, the earliest high resolution image was solved at 3.3 Å in 2011 from the GluCl channel (Figures 3 and 4C) from the Ecdysozoan nematode *Caenorhabditis elegans* [43]. In 2014, the first vertebrate pLGIC structure was solved [44], the mouse 5-HT_{3A}R, at an atomic resolution of 3.5 Å (Figures 3 and 4B).

The comparison of all these structures displayed in Figure 4 shows a striking similarity in spite of the low amino acid sequence identity among them. All these structures have a universal organization, comprising five homologous subunits (identical or not) symmetrically arranged around a central ionic channel that is perpendicular to the membrane. The structures show the ECD and TMD as predicted by the amino acid sequences. These structures showed other differences between the prokaryotic and eukaryotic N-terminal domains.

The ECD of all prokaryotic and eukaryotic pLGIC receptors are folded in an immunoglobulin-like β-sandwich stabilized by inner hydrophobic residues. This folding pattern is remarkably highly conserved among receptor subtypes. However, the connecting loops are variable in length and structure. Eukaryotic, but not prokaryotic, structures have an α-helix at the beginning of the ECD whose role is unknown, but it has been implicated in receptor expression and antibody recognition, especially those autoantibodies found in patients with *Myasthenia gravis* [45]. These differences can be observed in the cartoon structure representations included in Figure 4 (lower row). These cartoons also show some other differences in the structures. For example, the 5-HT_{3A}R has a small α-helical structure in Loop B (Figure 4B) not seen in other structures. In addition, the β5-strand of GluCl has an inserted loop (Figure 4C), and similarly the prokaryotic ELIC has a unique small α-helical structure but in Loop A (Figure 4D).

The TMD is consistent in all structures and with the sequence based predicted folding, a four α-helix compacted bundle. The M2 segments in each pentamer, line and shape the ion channel itself as predicted [3, 46-49]. The M2s are surrounded by the M1 and M3 helices. The M4 segment conforms the outer ring of this packet bundle and interacts with the surrounding lipid bilayer. In fact, the GLIC structure shows three lipid molecules bound to each subunit docked in the crevices formed by M4 and either M1 or M3 [42]. This was also early described and predicted by experimental and modeling studies [50].

The functional conservation of the ECD and TMD during evolution is demonstrated by the fact that many chimeras using different eukaryotic and prokaryotic domains have been successfully produced [51-52], especially important is the one constructed between the bacterial GLIC and α1 GlyR subunits [53].

In addition to the similarities in the tertiary and quaternary structures, the more basic comparative sequence analysis of the prokaryotic and eukaryotic pLGIC subunits allowed to pinpoint some key common residues that are probably essential for functional features and for the tertiary and quaternary structure determination: (1) The helix-bending position in helix M1 (P, G, or S; *T. californica* α1-P221) is critical for the conformational change during channel opening [33-34]. (2) A small residue in the middle of helix M2 (*T. californica* α1-S252) initiates a helix bend important for channel opening [33]. (3) A small residue in the M2-M3 loop (*T. californica* α1-G275) is critical for the rotational freedom of the M2 helix during the gating process [33]. (4) A polar residue just before the beginning of the segment M4 [33]. (5) Cation channels usually have a sequence motif of the form glutamate (E)-[arginine (R)/lysine (K)] with the E residue playing a role in cation selectivity [4, 47, 54]. An E residue is present in eight of the known bacterial sequences [9]. (6) Anion channels usually have a motif of the form alanine (A)-[RK] with the basic residue participating in anion selectivity [4, 49, 54]. A basic residue corresponding to the anion channel's motif is seen in six of the known bacterial sequences [9]. (7) Non-conserved polar or charged residues at the C-terminal end of M2 might be involved in fine-tuning ion selectivity [6, 49]. (8) In many animal pLGIC receptors, the ligand-binding pocket is surrounded by multiple aromatic residues that interact with the ligand. Eukaryotic or prokaryotic receptors, with the exception of the human ZAC, contain at least a single aromatic residue in the agonist binding pocket, suggesting that cation-π interactions with the bound ligand are widespread in the entire superfamily [9]. The average number of aromatic positions is 2.1 for prokaryotic

receptors and 2.6 for eukaryotic receptors, suggesting some differences in the type and nature of the ligand interactions [9]. (9) A highly conserved motif aPaD (where 'a' is any aromatic residue) is observed in the middle of the Cys-loop in animal pLGICs or in the homologous region in prokaryotic receptors [55]. (10) A tryptophan residue is located at the end of the β2-strand (W58 in AChBP) [9]. (11) An aromatic or hydrophobic residue is located in a homologous position to W82 (in AChBP) [9].

In spite of the similarities shown in Figure 4, the prokaryotic pLGICs can be more complex in than the eukaryotic ones. Several receptors have one or more amino-terminal fusions with other proteins [9]. For example, Figure 3 shows a domain from the periplasmic binding protein type I (PBP-I) superfamily. These bacterial proteins bind amino acids or small molecules in the extracellular space to facilitate their uptake. Other bacterial proteins are the CACHE (Ca^{2+} channels and chemotaxis receptors) and MCP-N (methyl-accepting chemotaxis protein-N) domains. These are common prokaryotic sensor domains that bind a variety of extracellular or periplasmic ligands and regulate signal transduction. In addition, some of the bacterial forms occur in predicted operons with the genes of the PBP-II superfamily and the above mentioned CACHE domains. These domains, fused or predicted, may be involved in sensing the concentration of small molecules (as proposed for the pLGIC alone by Ortells and Lunt [1]), which then bind to the LBD finally opening the channel. This mechanism might regulate the bacterium motility in response to the ligand [9].

Evolution of desensitization

Ortells and Lunt [9, 56], under the initial assumption that pLGICs might be present in ancient organisms such as prokaryotes, proposed that desensitization would have provided a mechanism to prevent long term opening of the channel that could be not only harmful but futile to the primitive cell. This proposal was made in the context that in animals, and especially in the muscle junction, these receptors never reach desensitization as the neurotransmitter is rapidly eliminated from the synapse by the action of hydrolytic enzymes (i.e., acetylcholinesterase). In other words, desensitization was only seen experimentally. Therefore, given that desensitization does exist in present day animal pLGICs, and that there is no selectivity pressure in animals to keep this feature, they stated that "this would indicate that desensitisation is intrinsically inherent in the very basics of the structure of these receptors and probably strongly selected for prior to the LGICs' assumption of their present function in the nervous system". With the discovery that prokaryotic pLGICs can be desensitized [41, 57], this assumption was confirmed. Thus, even without the selective pressure, this feature remains intact in vertebrate pLGICs. The available data show that there are many receptor regions involved in desensitization, including the ECD and TMD, but especially in the ECD-TMD interface [58-62]. The complex structures involved in the process of desensitization have low probability of being lost completely by genetic drift, without disrupting the entire receptor functions.

Assembly evolution

Ortells and Lunt [1] suggested that the first receptors of the eukaryotic family were homomeric in nature, a proposition in agreement with that made by Tasneem *et al.* [9], suggesting that the ancestral gene of the animal branch might have been acquired only once by lateral transfer from a prokaryotic form, which are known to be only homomeric. The eukaryotic receptors can be homomeric but also heteromeric structures formed by different types of homologous subunits. In vertebrates, ten nicotinic α (α1-α10) and seven non-α (β1 to β4, γ, δ, and ε) subunits have been identified, and thus the potential number of possible receptors is astronomic. If we assume a molecular clock, this multiplication took place very recently (Figure 2). It seems that in muscle nAChRs the non-α subunit proliferation (β1, γ, δ, and ε) may have been a means to "fine tune" a single role based on one receptor subtype, whilst in the brain it was used to expand beyond one role generating different receptor subtypes formed by the combination of α and β subunits. In spite of this possibility, relatively few combinations have been observed [63-64]. In fact, the amount of actual receptor variability is an unsolved problem that has been extensively analyzed. For example, the only outcome of the *in vivo* expression of the α1, β1, δ, and γ subunits found in muscle or in the electric ray *Torpedo* nAChRs is the β1δα1γα1 stoichiometry (and likewise, in the adult muscle nAChR, replacing the δ by the ε subunit). It is well known that neuronal nAChR subunits can be combined as heteromeric receptors, for example, α4β2 or α3β4, whereas the α7 subunit can form functional homomeric receptors. More recent data indicate, however, that the most probably endogenous α7 nAChR stoichiometry also contains β subunits (e.g., β2 subunits [65]). The structural information for subunit recognition during assembly is located in the ECD of nAChRs [66-68]. Based on this, Ortells and Barrantes [69] presented an assembly model based on the shape and hydrophobic and electrostatic interactions of the ECD of nAChRs and 5-HT_3Rs which could explain why some subunits may interact or not with other subunits, and given a particular pathway of assembly the stoichiometry of the receptor as the output. In an attempt to explain this subunit specificity, Ortells [70], after comparing the rates of substitutions in synonymous and non-synonymous nucleotide positions [71], could not find a trace of positive selection on

ECD amino acids in subunits that have a recent common ancestor. So far, how the affinity between subunits evolved is still unknown.

Conclusions

The possibility of studying the evolution of pLGICs began with the cloning of the α1 subunit from *Torpedo* electric organs in 1982 [72]. Thirty years later, we know that pLGICs are present in a wide variety of organisms, including prokaryotes, where they presumably have had their origin, and so far in eukaryotes only in animals, probably from a unique lateral transfer from the former. This restriction contrasts with the widespread presence in prokaryotes and eukaryotes of other types of homologous ion channels. The resolution of the three-dimensional structures of both prokaryotic and eukaryotic pLGICs made clear that the basic scaffold and functional organization of these proteins remained unchanged since they appeared. However, we also know that both the ECD [33] and TMD [72] can fold in isolation, and that many functional ECD-TMD chimeras can be constructed, even by mixing prokaryotic and eukaryotic domains [51,53,73]. Therefore, we can imagine a scenario where pLGICs had their origin from an earlier domain fusion from independent TMDs and ECDs.

Conflicting interests

The authors have declared that no conflict of interests exists.

Acknowledgments

M.O.O. is supported by the National Council for Science and Technology (CONICET), Argentina.

References

1. Ortells MO, Lunt GG. Evolutionary history of the ligand-gated ion-channel superfamily of receptors. Trends Neurosci 1995; 18:121-127.
2. Davies PA, Wang W, Hales TG, Kirkness EF. A novel class of ligand-gated ion channel is activated by Zn^{2+}. J Biol Chem 2003; 278:712-717.
3. Imoto K, Methfessel C, Sakmann B, Mishina M, Mori Y, Konno T, *et al*. Location of a delta-subunit region determining ion transport through the acetylcholine receptor channel. Nature 1986; 324:670-674.
4. Keramidas A, Moorhouse AJ, French CR, Schofield PR, Barry PH. M2 pore mutations convert the glycine receptor channel from being anion- to cation-selective. Biophys J 2000; 79:247-529.
5. Corringer PJ, Poitevin F, Prevost MS, Sauguet L, Delarue M, Changeux JP. Structure and Pharmacology of Pentameric Receptor Channels: From Bacteria to Brain. Structure 2012; 20:941-956.
6. Lester HA, Dibas MI, Dahan DS, Leite JF, Dougherty DA. Cys-loop receptors: new twists and turns. Trends Neurosci 2004; 27:329-936.
7. Ortells, MO, Barrantes, GE. Tobacco addiction: A biochemical model of nicotine dependence. Medical Hypotheses 2010; 74:884-994.
8. Ortells, MO, Arias, HA. Neurochemical mechanisms underlying nicotine dependence. In Pharmacology of Nicotinic Acetylcholine Receptors from the Basic and Therapeutic Perspectives. Edited by Arias H. Kerala: India, Research Signpost. 2011:199-221.
9. Tasneem A, Iyer LM, Jakobsson E, Aravind L. Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels. Genome Biol 2005; 6, R4.
10. Le Novère N, Changeux J-P. Molecular evolution of the nicotinic acetylcholine receptor: An example of multigene family in excitable cells. J Mol Evol 1995; 40:155-172.
11. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In Evolving Genes and Proteins. Edited by Bryson V, Vojel HJ. New York: Academic Press; 1965:97-166.
12. Dent JA. Evidence for a diverse cys-loop ligand-gated ion channel superfamily in early bilateria. J Mol Evol 2006; 62:523-535
13. Schopf JW. Earth's Earliest Biosphere: Its Origin and Evolution. Guildford N.J. Princeton University Press; 1983.
14. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J Mol Biol 1990; 213:899-929.
15. Goldstein SA, Wang KW, Ilan N, Pausch MH. Sequence and function of the two P domain potassium channels: implications of an emerging superfamily. J Mol Med 1998; 76:13-20.
16. Ito M, Xu H, Guffanti AA, Wei Y, Zvi L, Clapham DE, Krulwich TA. The voltage-gated Na+ channel NaVBP has a role in motility, chemotaxis, and pH homeostasis of an alkaliphilic *Bacillus*. Proc Natl Acad Sci USA 2004; 101:10566-10571.
17. Moulton G, Attwood TK, Parry-Smith DJ, Packer JC. Phylogenomic analysis and evolution of the potassium channel gene family. Receptors Channels 2003, 9:363-377.
18. Nelson RD, Kuan G, Saier MH Jr, Montal M. Modular assembly of voltage-gated channel proteins: a sequence analysis and phylogenetic study. J Mol Microbiol Biotechnol 1999; 1:281-287.
19. Cockcroft VB, Osguthorpe DJ, Barnard EA, Friday AE, Lunt GG. Ligand-gated ion channels: Homology and diversity. Mol Neurobiol 1990; 4:129-169.
20. Dent JA. The evolution of pentameric ligand-gated ion channels. Adv Exp Med Biol 2010; 683:11–23.
21. Gisselmann G, Plonka J, Pusch H, Hatt H. *Drosophila melanogaster* GRD and LCCH3 subunits form heteromultimeric GABA-gated cation channels. Br J Pharamcol 2004; 142:409-413.
22. Zheng Y, Hirschberg B, Yuan J, Wang AP, Hunt DC, Ludmerer SW, *et al*. Identification of two novel *Drosophila melanogaster* histamine-gated chloride channel subunits expressed in the eye. J Biol Chem 2002; 277:2000-2005.
23. Cully DF, Vassilatis DK, Liu KK, Paress PS, Van der Ploeg LH,

Schaeffer JM, *et al*. Cloning of an avermectin-sensitive glutamate-gated chloride channel from *Caenorhabditis elegans*. Nature 1994; 371:707-711.

24. Dent JA, Davis MW, Avery L. avr-15 encodes a chloride channel subunit that mediates inhibitory glutamatergic neurotransmission and ivermectin sensitivity in worms. EMBO J 1997; 16:5867-5879.
25. Vassilatis DK, Elliston KO, Paress PS, Hamelin M, Arena JP, Schaeffer JM, *et al*. Evolutionary relationship of the ligand-gated ion channels and the avermectin-sensitive, glutamate-gated chloride channels. J Mol Evol 1997; 44:50150-50158.
26. Horoszok L, Raymond V, Sattelle DB, Wolstenholme AJ. GLC-3:a novel fipronil and BIDN-sensitive, but picrotoxinininsensitive, L-glutamate-gated chloride channel subunit from *Caenorhabditis elegans*. Br J Pharmacol 2001; 132:1247-1254.
27. Beg AA, Jorgensen EM. EXP-1 is an excitatory GABA-gated cation channel. Nat Neurosci 2003; 6:1145-1152.
28. Van Nierop P, Keramidas A, Bertrand S, van Minnen J, Gouwenberg Y, Bertrand D, *et al*. Identification of molluscan nicotinic acetylcholine receptor (nAChR) subunits involved in formation of cation- and anion-selective nAChRs. J Neurosci 2005; 25:10617-10626.
29. Putrenko I, Zakikhani M, Dent JA. A family of acetylcholine-gated chloride channel subunits in *Caenorhabditis elegans*. J Biol Chem 2005; 280:6392-6398.
30. Ranganathan R, Cannon SC, Horvitz HR. MOD-1 is a serotonin-gated chloride channel that modulates locomotory behaviour in *C. elegans*. Nature 2000; 408:470-475.
31. Anantharaman V, Aravind L. Application of comparative genomics in the identification and analysis of novel families of membrane-associated receptors in bacteria. BMC Genomics 2003; 4:34.
32. Xu J, Zhu Y, Heinemann SF. Identification of Sequence Motifs That Target Neuronal Nicotinic Receptors to Dendrites and Axons. J Neurosci 2006; 26:9780-9793.
33. Miyazawa A, Fujiyoshi Y, Unwin N. Structure and gating mechanism of the acetylcholine receptor pore. Nature 2003; 423:949-955.
34. Unwin N, Miyazawa A, Li J, Fujiyoshi Y. Activation of the nicotinic acetylcholine receptor involves a switch in conformation of the α subunits. J Mol Biol 2002; 319:1165-1176.
35. Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, Smit AB, *et al*. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. Nature 2001; 411:269-276.
36. Hansen S, Sulzenbacher G, Huxford T, Marchot P, Taylor P, Bourne Y. Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. EMBO J 2005; 24:2625-2646.
37. Billen B, Spurny R, Brams M, van Elk R, Valera-Kummer S, Yakel JL, *et al*. Molecular actions of smoking cessation drugs at α4β2 nicotinic receptors defined in crystal structures of a homologous binding protein. Proc Natl Acad Sci USA 2012; 109:9173-9178.
38. Bocquet N, Prado de Carvalho L, Cartaud J, Neyton J, Le Poupon C, Taly A, *et al*. A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. Nature 2007; 445:116-119.
39. Hilf RJ, Dutzler R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. Nature 2008; 452:375-379.
40. Hilf RJ, Dutzler R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. Nature 2009; 457:115-118.
41. Zimmermann I, Dutzler, R. Ligand activation of the prokaryotic pentameric ligand-gated ion channel ELIC. PLoS Biol 2011; 9, e1001101.
42. Bocquet N, Nury H, Baaden M, Le Poupon C, Changeux JP, Delarue M, *et al*. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. Nature 2009; 457:111-114.
43. Hibbs RE, Gouaux E. Principles of activation and permeation in an anion-selective Cys-loop receptor. Nature 2011:474:54–60.
44. Hassaine G, Deluz C, Grasso L, Wyss R, Tol MB, Hovius R, *et al*. X-ray structure of the mouse serotonin 5-HT3 receptor. Nature 2014; 512:276-281.
45. Luo J, Taylor P, Losen M, de Baets MH, Shelton GD, Lindstrom J. Main immunogenic region structure promotes binding of conformation dependent *Myasthenia gravis* autoantibodies, nicotinic acetylcholine receptor conformation maturation, and agonist sensitivity. J Neurosci 2009; 29:13898-13908.
46. Giraudat J, Dennis M, Heidmann T, Chang JY, Changeux JP. Structure of the high-affinity binding site for noncompetitive blockers of the acetylcholine receptor:serine-262 of the delta subunit is labeled by [3H]chlorpromazine. Proc Natl Acad Sci USA 1986; 83:2719-2773.
47. Giraudat J, Dennis M, Heidmann T, Chang JY, Changeux JP. Structure of the high-affinity binding site for noncompetitive blockers of the acetylcholine receptor: [3H]chlorpromazine labels homologous residues in the beta and delta chains. Biochemistry 1987; 26:2410-2418.
48. Hucho F, Oberthür W, Lottspeich F. The ion channel of the nicotinic acetylcholine receptor is formed by the homologous helices M II of the receptor subunits. FEBS Lett 1986; 205:137-142.
49. Imoto K, Busch C, Sakmann B, Mishina M, Konno T, Nakai J, *et al*. Rings of negatively charged amino acids determine the acetylcholine receptor channel conductance. Nature 1988; 335:645-648.
50. Ortells MO, Cockcroft VB, Lunt GG, Marsh D, Barrantes FJ. The nicotinic acetylcholine receptor and its lipid microenvironment. In Membrane proteins: Structures, interactions and models. Edition 25. Edited by Pullman, A, Jortner, J, Pullman, B. Dordrecht, Kluwer Academic Publishers. 1992:185-98.
51. Eiselé JL, Bertrand S, Galzi JL, Devillers-Thiéry A, Changeux JP, Bertrand D. Chimaeric nicotinic-serotonergic receptor combines distinct ligand binding and channel specificities. Nature 1993; 366:479-483.
52. Grutter T, de Carvalho LP, Dufresne V, Taly A, Edelstein SJ, Changeux JP. Molecular tuning of fast gating in pentameric ligand gated ion channels. Proc Natl Acad Sci USA 2005; 102:18207-18212.

53. Duret G, Van Renterghem C, Weng Y, Prevosta M, Moraga-Cida G, Huona C, *et al*. Functional prokaryotic-eukaryotic chimera from the pentameric ligand-gated ion channel family. Proc Natl Acad Sci USA 2011; 108:12143-12148.

54. Gunthorpe MJ, Lummis SC. Conversion of the ion selectivity of the 5-HT(3a) receptor from cationic to anionic reveals a conserved feature of the ligand-gated ion channel superfamily. J Biol Chem 2001; 276:10977-10983.

55. Connolly CN, Wafford KA. The Cys-loop superfamily of ligand gated ion channels: the impact of receptor structure on function. Biochem Soc Trans 2004; 32:529-534.

56. Ortells MO, Lunt GG. Evolution and desensitization of LGIC receptors. Trends Neurosci 1995; 18:297-299.

57. Gonzalez-Gutierrez G, Grosman C. Bridging the gap between structural models of nicotinic receptor superfamily ion channels and their corresponding functional states. J Mol Biol 2010; 403:693-705.

58. Bohler S, Gay S, Bertrand S, Corringer PJ, Edelstein SJ, Changeux JP, *et al*. Desensitization of neuronal nicotinic acetylcholine receptors conferred by N-terminal segments of the β2 subunit. Biochemistry 2001; 40:2066-2074.

59. Bouzat C, Bartos M, Corradi J, Sine SM. The interface between extracellular and transmembrane domains of homomeric Cys loop receptors governs open-channel lifetime and rate of desensitization. J Neurosci 2008; 28:7808-7819.

60. Forman SA, Miller K. Anesthetic sites and allosteric mechanisms of action on Cys-loop ligand-gated ion channels. Can J Anaesth 2011; 58:191-205.

61. Wang Q, Lynch JW. Activation and desensitization induce distinct conformational changes at the extracellular-transmembrane domain interface of the glycine receptor. J Biol Chem 2011; 286:38814-38824.

62. Yamodo IH, Chiara DC, Cohen JB, Miller KW. Conformational changes in the nicotinic acetylcholine receptor during gating and desensitization. Biochemistry 2010; 49:156-165.

63. Lukas RJ, Changeux JP, Le Novère N, Albuquerque EX, Balfour DJ, Berg DK, *et al*. International union of pharmacology XX Current status of the nomenclature for nicotinic acetylcholine receptors and their subunits. Pharmacological Rev 1999; 51:397-401.

64. Millar NS. Assembly and subunit diversity of nicotinic acetylcholine receptors. Biochem Soc Transactions 2003; 31:869-874.

65. Khiroug SS, Harkness PC, Lamb PW, Sudweeks SN, Khiroug L, Millar NS, *et al*. Rat nicotinic ACh receptor α7 and β2 subunits co-assemble to form functional heteromeric nicotinic receptor channels. J Physiol 2002; 540:425-434.

66. Xu XM, Hall ZW. Extracellular domains mediating ε subunit interactions of muscle acetylcholine receptor. Nature 1991; 352:64-67.

67. Xu XM, Hall ZW. A sequence in the main cytoplasmic loop of the α subunit is required for assembly of mouse muscle nicotinic acetylcholine- receptor. Neuron 1994; 13:247-255.

68. Sumikawa K. Sequences on the N-terminus of ACh receptor subunits regulate their assembly. Mol Brain Res 1992; 13:349-353.

69. Ortells MO, Barrantes GE. A model for the assembly of nicotinic receptors based on subunit-subunit interactions. Proteins: Struc Func Bioinf 2008; 70:473-488.

70. Ortells MO. Evolution of the AChR and Other Ligand-Gated Ion Channels. In The nicotinic acetylcholine receptor: Current views and future trends. Edited by F. Barrantes. Berlin/Heidelberg Springer Verlag, and Georgetown, Landes Publishing Co TX; 1998:11-30.

71. Li WH. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 1993; 36:96-99.

72. Noda M, Takahashi H, Tanabe T, Toyosato M, Furutani Y, Hirose T, *et al*. Primary structure of alpha-subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence. Nature 1982; 299:793-797.

73. Bondarenko V, Tillman T, Xu Y, Tang P. NMR structure of the transmembrane domain of the n-acetylcholine receptor β2 subunit. Biochim Biophys Acta 2010; 1798:1608-1614.

74. Cooper ST, Harkness PC, Baker ER, Millar NS. Up-regulation of cell-surface α4β2 neuronal nicotinic receptors by lower temperature and expression of chimeric subunits. J Biol Chem 1999; 274:27145-27152.

75. Lvtrup S. The phylogeny of vertebrata. John Wiley & Sons 1977.