340–342 *Nucleic Acids Research, 1999, Vol. 27, No. 1*

The Ligand Gated Ion Channel Database

Nicolas Le Novère* and Jean-Pierre Changeux

Neurobiologie Moléculaire, URA CNRS 1281, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15, France

Received August 11, 1998; Revised September 7, 1998; Accepted September 30, 1998

ABSTRACT

The ligand gated ion channels (LGICs) are ionotropic receptors to neurotransmitters. Their physiological effect is carried out by the opening of an ionic channel upon binding of a particular neurotransmitter. These LGICs constitute superfamilies of receptors formed by homologous subunits. A database has been developed to handle the growing wealth of cloned subunits. This database contains nucleic acid sequences, protein sequences, as well as multiple sequence alignments and phylogenetic studies. This database is accessible via the worldwide web (http://www.pasteur.fr/units/neubiomol/LGIC.html), where it is continuously updated. A downloadable version is also available [currently v0.1 (98.06)].

INTRODUCTION

The term ligand gated ion channels (LGICs) may define three different groups of proteins which cross the cell membrane. In its narrow sense, this term is synonymous of the nicotinicoid superfamily of receptors (1,2). Thus, it includes the nicotinic receptors, the 5-HT_3 receptor, the $GABA_A$ (3) and $GABA_C$ receptors (4), the glycine receptors (5) and some invertebrate anionic glutamate receptors (6). All these receptors are considered as pentamers made with homologous subunits; in other words all the genes coding for these subunits are assumed to derive from a common ancestor gene (7,8). Each subunit can be subdivided into a large extracellular N-terminal domain, followed by three membrane spanning segments, an intracellular loop of variable length and a fourth membrane-spanning segment, the C-terminal end being then extracellular.

An intermediate meaning is equivalent to neurotransmitter-gated ion channel. This group comprises, in addition, the cationic glutamate receptors (9) and the ionotropic ATP receptors. All these receptors were initially thought to belong to the same superfamily and to possess the same global structure. It finally appeared that each of these groups in itself forms a superfamily of receptors made of homologous subunits. The glutamate cationic receptors include the NMDA receptors, the AMPA receptors, the kainate receptors, and the receptors formed by the δ subunits. This superfamily comprises also the kainate binding proteins, of a still enigmatic physiological function. Each subunit can be subdivided (10) into a large extracellular N-terminal domain, followed by a membrane spanning segment, a P-loop, another membrane spanning segment, a second extracellular domain, and a final membrane spanning segment, the C-terminal end being then intracellular. Each subunit of ionotropic ATP receptor is formed by a transmembrane segment followed by an extracellular N-terminal domain and a second transmembrane segment, the C-terminal end being then intracellular.

Finally, the widest meaning comprises all the ion channels which opening is controlled by a ligand (11). It then covers the inositol phosphate and cyclic nucleotides intracellularly gated channels.

In its current state, v0.1 (98.06), the Ligand Gated Ion Channel Database provides information regarding the intermediate meaning of LGICs with an emphasis on the members of the narrower ones. The nicotinicoid superfamily of receptors is extensively covered whereas the excitatory glutamate and ATP receptor superfamilies are less systematically covered.

STRUCTURE OF THE DATABASE

The Ligand Gated Ion Channel Database takes its origin in a survey of the nicotinic receptor subunits evolutionary history (7). The sequence material accumulated for this study constituted initially a home database for the Molecular Neurobiology Laboratory, that was placed on the World Wide Web by the end of 1995. The address of the Ligand Gated Ion Channel Database is: http://www.pasteur.fr/units/neubiomol/LGIC.html

The database comprises sequences [212 different entries in v0.1 (98.06) as illustrated in Table 1], alignments and phylogenetic information.

Table 1. Content of v0.1 (98.06)

Group	Number of entries
Nicotinicoid anionic receptor subunits (nucleic acid)	74
Nicotinicoid anionic receptor subunits (protein)	84
Nicotinicoid cationic receptor subunits (nucleic acid)	98
Nicotinicoid cationic receptor subunits (protein)	98
Glutamate cationic receptor subunits (nucleic acid)	23
ATP receptor subunits (protein)	7

*To whom correspondence should be addressed. Tel: +33 1 45 68 88 44; Fax: +33 1 45 68 88 36; Email: lenov@pasteur.fr

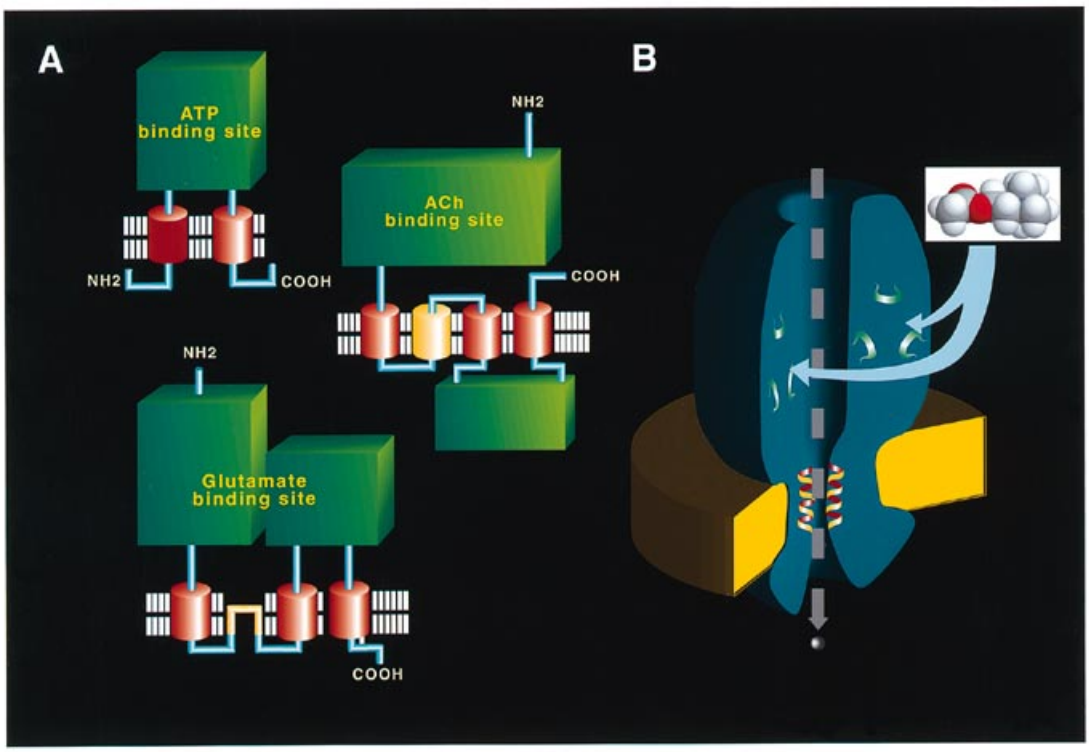

Figure 1. (**A**) Transmembrane topology of one subunit of each superfamily. The outlined membrane-spanning segment is probably lining the ion channel. (**B**) Quaternary structure of the nicotinic acetylcholine receptor in the open state. Two of the five subunits have been removed to expose the complementary faces of the acetylcholine binding site.

Table 2. Hierachy of the Ligand Gated Ion Channel Database

Superfamily of nicotinicoid receptors
family of nicotinic receptors
subfamily of epithelial receptors
subfamily of neuronal α-bungarotoxin sensitive receptors
subfamily of muscle receptors
subfamily of heteromeric neuronal receptors
subfamily of heteromeric protostomian receptors
family of serotonin receptors
family of GABA receptors
subfamily of $GABA_A$ receptors
subfamily of $GABA_C$ receptors
family of variable agonist receptors
subfamily of GABA receptors
subfamily of glutamate receptors
subfamily of glycine receptors
Superfamily of excitatory glutamate receptors
family of excitatory glutamate receptors
subfamily of NMDA receptor subunits
subfamily of AMPA receptor subunits
subfamily of kainate receptor subunits
subfamily of kainate binding proteins
subfamily of delta subunits
Superfamily of ATP receptors
family of ATP receptors

The sequences are most often originated from the GenBank and EMBL databases, although some of them were retrieved from SwissProt or directly copied from the original articles. The sequences of ATP receptor subunits were collected by Dr Ralf Schoepfer. The format is the one from GCG's Wisconsin Package (12). Some files containing cDNA were composed from the fusion of several files containing exon sequences and are then not directly processable as GCG sequence files. Some corrections have been made when the database sequence was found to be wrong (for instance when the published sequence was different from the one present in the databases). This is always clearly stated. The GenBank and EMBL databases are redundant. Therefore, instead of presenting every entry file, the Ligand Gated Ion Channel Database keeps only the longest sequence. The original work remains cited. When several different sequences exist for a particular subunit, without evidence of sequencing or typing errors, all the sequences are presented (for instance, this is the case of a gene sequenced in several different strains of rodent species).

The GenBank files with the formatted nucleotide sequence generally also contain the protein sequence in a raw format. However separate files, containing directly processable protein sequences under GCG format, have also been generated. For each group of sequences, the entries are ranked either by organisms or by sequence and function similarities (these two latter notions fortunately appear congruent in the case of LGIC subunits). The latter hierarchy required ranking thresholds (see Table 2). The first rank is the family, defined by the endogenous ligand responsible for the channel opening [in some cases the endogenous ligand differs according to the species. For instance the

invertebrate anionic glutamate receptors (6) are likely to be homologous to mammalian glycine or GABA receptors]. An exception is the nicotinic acetylcholine receptor family, classically characterised by the action of nicotine, an exogenous ligand. A second rank is the subfamily, defined by the ability to form endogenous receptors together (a family may be formed of only one subfamily as in the case of mammalian glycine receptors). Finally, a last rank is the tribe, sometimes defined by the similarity of function inside the oligomer (for instance the participation to the agonist binding site). It has to be noted that this classification is not totally congruent with the published phylogenetic inferences, although it is generally in global agreement. The online Ligand Gated Ion Channel Database is continuously updated. However, discrete versions can be downloaded from the WWW, in the form of a compressed archive.

A search system is available, allowing the user to scan the database with a nucleic acid or protein motif, by means of the program FASTA (13).

Users of the Ligand Gated Ion Channel Database are requested to cite the present article (exact format depending on the context of the citation).

ACKNOWLEDGEMENTS

The authors thank Ralf Schoepfer for the ATP receptor subunits and Yoav Paas for careful reading of the manuscript. They also thank all the users of the Ligand Gated Ion Channel Database who reported omissions, mistakes and suggested improvements. This work was supported by the Collège de France, the Centre National de la Recherche Scientifique; Biotech and Biomed contracts from the Commission of the European communities. N.L. was supported by the Ministère de l'Enseignement Supérieur et de la Recherche and by the Institut Pasteur.

REFERENCES

1 Cockroft,V., Osguthorpe,D., Barnard,E.A., Friday,A. and Lunt,G. (1992) *Mol. Neurobiol.*, **4**, 129–169.
2 Galzi,J.-L. and Changeux,J.-P. (1994) *Curr. Opin. Struct. Biol.*, **4**, 554–565.
3 Macdonald,R.L. and Olsen,R.W. (1994) *Annu. Rev. Neurosci.*, **17**, 569–602.
4 Bormann,J. and Feigenspan,A. (1995) *Trends Neurosci.*, **18**, 515–159.
5 Béchade,C., Sur,C. and Triller,A. (1994) *Bioessays*, **16**, 735–744.
6 Cully,D., Vassilatis,D., Liu,K., Paress,P., Van derPloeg,L., Schaeffer,J. and Arena,J. (1994) *Nature*, **371**, 707–711.
7 Le Novère,N. and Changeux,J.-P. (1995) *J. Mol. Evol.*, **40**, 155–172.
8 Ortells,M.O. and Lunt,G.G. (1995) *Trends Neurosci.*, **18**, 121–126.
9 Hollman,M. and Heinemann,S. (1994) *Annu. Rev. Neurosci.*, **17**, 31–108.
10 Paas,Y. (1998) *Trends Neurosci.*, **21**, 117–125.
11 Barnard,E.A. (1992) *Trends Biochem. Sci.*, **17**, 368–374.
12 Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
13 Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 2444–2448.