Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

Binding site prediction of galanin peptide using evolutionary trace method

Shanthi Nagarajan and Parthiban Marimuthu*

Bioinformatics Centre, Pondicherry University, Pondicherry, India; Parthiban Marimuthu* -
Email: ma_parthiban@yahoo.co.in; Phone: +91 413 2655212; Fax: +91 413 2655211; * Corresponding author
received June 08, 2006; revised July 17, 2006; accepted July 24, 2006; published online July 25, 2006

Abstract:

Galanin is a neuropeptide with aminoacid length ranging from 29 to 31 is widely distributed in central and peripheral nervous system. Galanin controls various psychological processes such as sensation of pain, learning, feeding, and sexual behaviour. The N-terminal region of this neuropeptide has highly conserved 15 amino acids, which is triggered by galanin receptors. We performed evolutionary trace analysis for galanin sequences to gather information about functional residues. The consensus pattern given by the evolutionary trace (ET) analysis is supported by CLUSTALW and WEBLOGO results. Our observations strongly suggest the presence of functional residues in the N-terminal region of galanin for agonist-receptor binding.

Keywords: Evolutionary trace; Binding site; Agonist- receptor

Background:

Galanin is a neuropeptide which mediates the function through its interaction with G-Protein Coupled Receptors and thereby controls various functions such as release of neurotransmitters or hormones. **[1]** Apart from this, its vital role in controlling various psychological processes such as sensation of pain, learning, feeding, sexual behavior makes it outstanding one. **[2]** The N-terminal region of galanin constitute about 15 amino acids which is highly conserved and act as the crucial region for agonist-receptor binding. **[1]**

There are about three Galanin receptors subtypes such as GalR1, GalR2 and GalR3, which belong to G-Protein coupled receptors, and the activity of galanin is modulated by its interaction with these receptors. **[1]** A detailed analysis on the evolutionary conservation information extracted from multiple sequence alignment would be used as important tool for prediction of functional properties as well as prediction of ligand binding sites, protein interface surfaces etc. The detection of conserved residues would be useful in identifying the functionally important residue even in the absence of structural information. **[3]** This work is an attempt to explore the information about functional residues of galanin through evolutionary conservation. Evolutionary Trace (ET Method) is a method in which protein sequences of a particular protein family is partitioned in to different groups, which originate from the common node in the phylogenetic tree and it also involves evolutionary time cut-off. The next step involves construction of consensus sequence for each group.

The invariant residues extracted from consensus sequences of each group were taken in to account and a comparison was done between invariant residues of consensus sequences of each group. According to the ET method, each residue is reported as either conserved or class-specific or variable based on the conservation properties. Later these residues are plotted on to the structure to get the information about three-dimensional structure. **[4]** The presence of common ancestral functional regions can be identified through the spatial arrangement of conserved and class-specific residues, which helps in the determination of the evolutionarily conserved functional sites.

Methodology:

A total of 13 non-redundant protein sequences of galanin from various sources were extracted from Swissprot database **[5]** and Genbank. **[6]** We performed a multiple sequence alignment for all galanin sequences using CLUSTALW. **[7]** ET studies for galanin sequences were carried out using ET server **[3, 4]**, which accepts multiple alignments as an input. The server calculated phylogenetic tree for galanin sequences, which were splited into two groups.

Results and discussions:

Galanin peptide sequence generally starts with amino acid sequence GWT. The absence of aminoacid GWT at N-terminal region was observed in case of galanin from *Canis familiaris* and is followed by 10 highly conserved residues LNSAGYLLGP sequence (Table 1). The ET Server generated 10 traces (Figure 1), which clearly represents the relationships among the sequences and conserved patterns across all the sequences. Trace 1 provides more conserved consensus pattern of "---LN-AGYLLGPH----HR----K-G----" whereas trace 10 exhibits lesser-conserved pattern with highest sequence similarity (Table 2). The consensus pattern given by the ET server was validated by using CLUSTALW and WEBLOGO **[8]** ((Figure 2). The CLUSTALW results indicated major proportion of conserved residues starting from 4 to 15 alignment positions in which the majority of the consensus pattern given by the ET server fits. Moreover the WEBLOGO results also support the same observation.

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

No.	Accession No.	Organism	Sequence	No. of Residues
1	NP_150240	*Rattus norvegicus* (Norway rat)	GWTLNSAGYLLGPHAIDNHRSFSDKHGLTG	29
2	P47212	*Mus musculus* (house mouse)	GWTLNSAGYLLGPHAIDNHRSFSDKHGLT	29
3	P22466	*Homo sapiens* (human)	GWTLNSAGYLLGPHAVGNHRSFSDKNGLTS	30
4	P33710	*Canis familiaris* (dog)	LNSAGYLLGPHAIDNHRSFHEKPGLT	26
5	NP_999399	*Sus scrofa* (pig)	GWTLNSAGYLLGPHAIDNHRSFHDKYGLAGK	31
6	P31234	*Ovis aries* (sheep)	GWTLNSAGYLLGPHAIDNHRSFHDKHGLA	29
7	P47215	*Alligator mississippiensis* (American alligator)	GWTLNSAGYLLGPHAIDNHRSFNEKHGIA	29
8	P47216	*Rana ridibunda* (marsh frog)	GWTLNSAGYLLGPHAIDNHRSFNDKHGLA	29
9	P47214	*Amia calva* (bowfin)	GWTLNSAGYLLGPHAVDNHRSLNDKHGLA	29
10	P11242	*Bos taurus* (cattle)	GWTLNSAGYLLGPHALDSHRSFQDKHGLA	29
11	P47213	*Oncorhynchus mykiss* (rainbow trout)	GWTLNSAGYLLGPHGIDGHRTLSDKHGLA	29
12	2102233A	*Thunnus albacares* (yellowfin tuna)	GWTLNAAGYLLGPHGIDGHRTLGDKPGLA	29
13	Q9W6M9	*Coturnix japonica* (Japanese quail)	GWTLNSAGYLLGPHAVDNHRSFNDKHGFT	29

Table 1: The list of galanin sequences with its database accession numbers and length of residues

Trace No.	Organism involved in group formation (ref. table)			Consensus Sequence
	Group 1	Group 2		
1	12, 11, 6, 5, 8, 2, 1, 9, 7, 10, 4, 3, 13		GROUP_1	---LN-AGYLLGP----------
			SUMMARY	---LN-AGYLLGP----------
2	12, 11, 6, 5, 8, 2, 1, 9, 7, 10, 4, 3		GROUP_1	---LN-AGYLLGPH----HR----K-G---
			SUMMARY	---LN-AGYLLGPH----HR----K-G---
3	12, 11, 6, 5, 8, 2, 1, 9, 7, 10, 4, 3		GROUP_1	---LN-AGYLLGPH----HR----K-G---
			SUMMARY	---LN-AGYLLGPH----HR----K-G---
4	12, 11, 6, 5, 8, 2, 1, 9, 7, 10, 4, 3		GROUP_1	---LN-AGYLLGPH----HR----K-G---
			SUMMARY	---LN-AGYLLGPH----HR----K-G---
5	12, 11, 6, 5, 8, 2, 1, 9, 7, 10, 4, 3		GROUP_1	---LN-AGYLLGPH----HR----K-G---
			SUMMARY	---LN-AGYLLGPH----HR----K-G---
6	12, 11	6, 5, 8, 2, 1, 9, 7, 10, 4, 3	GROUP_1	GWTLN-AGYLLGPHGIDGHRTL-DK-GLA.
			GROUP_2	---LNSAGYLLGPHA---HRS---K-G---
			SUMMARY	---LN-AGYLLGPHX---HRX---K-G---
7	12, 11	6, 5, 8, 2, 1, 9, 7, 10, 4, 3	GROUP_1	GWTLN-AGYLLGPHGIDGHRTL-DK-GLA.
			GROUP_2	---LNSAGYLLGPHA---HRS---K-G---
			SUMMARY	---LN-AGYLLGPHX---HRX---K-G---
8	12, 11	6, 5, 8, 2, 1, 9, 7, 10	GROUP_1	GWTLN-AGYLLGPHGIDGHRTL-DK-GLA.
			GROUP_2	GWTLNSAGYLLGPHA-D-HRS---K-G---
			SUMMARY	GWTLN-AGYLLGPHX-D-HRX---K-G---
9	6, 5, 8, 2, 1		GROUP_1	GWTLNSAGYLLGPHAIDNHRSF-DK-GL--
			SUMMARY	GWTLNSAGYLLGPHAIDNHRSF-DK-GL--
10	6, 5, 8	2, 1	GROUP_1	GWTLNSAGYLLGPHAIDNHRSF-DK-GLA.
			GROUP_2	GWTLNSAGYLLGPHAIDNHRSFSDKHGLT-
			SUMMARY	GWTLNSAGYLLGPHAIDNHRSF-DK-GLX-

Table 2: Evolutionary Trace Server provides 10 traces for 13 galanin, which is separated into two groups. The numbers in the each group corresponds to a different organism, which is mentioned in Table 1

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

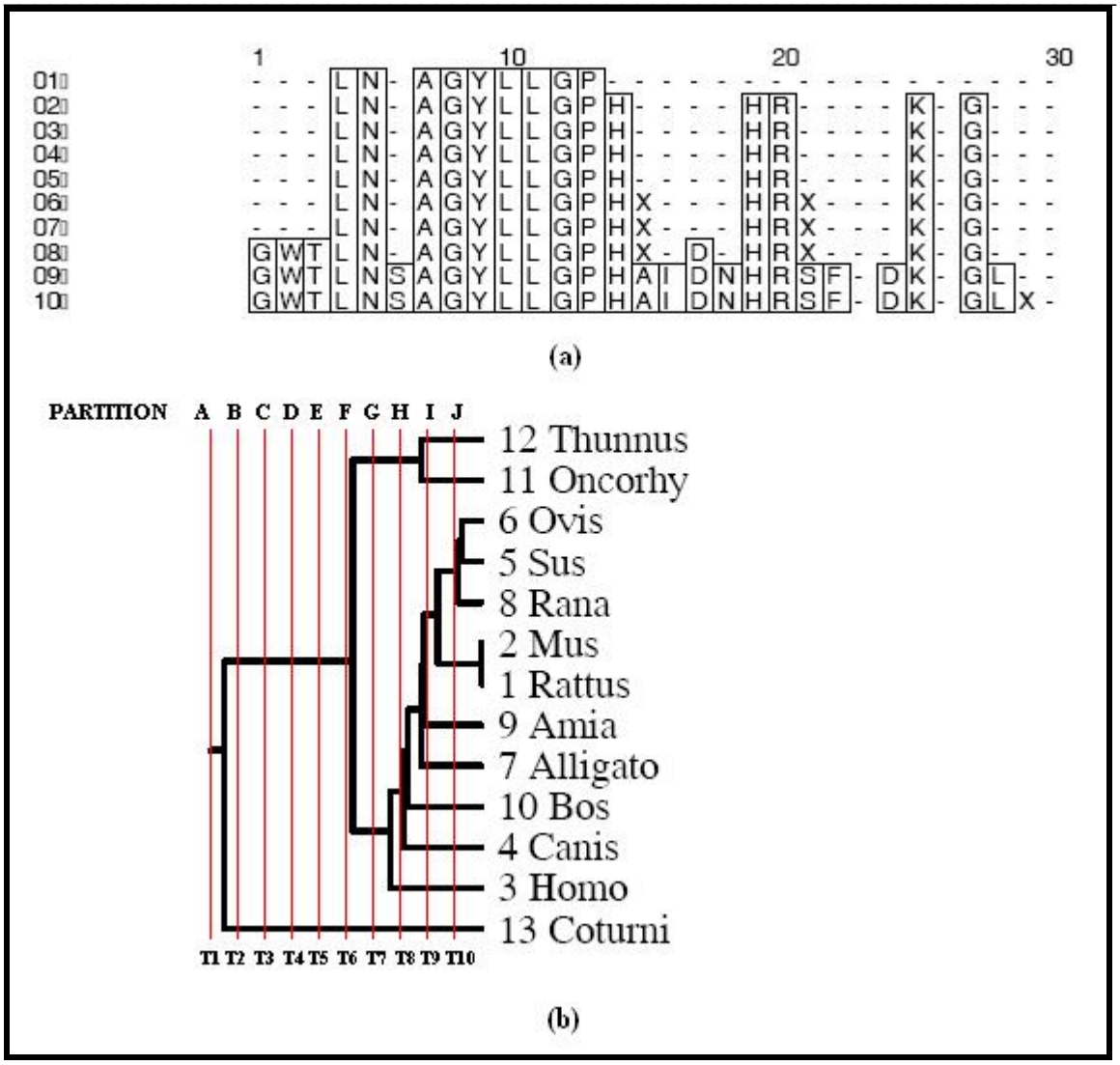

Figure 1: (a) Evolutionary Trace shows conserved consensus pattern, **(b)** Vertical lines in dendrogram A to J shows different Partition Identity Cutoffs (PICs). Each PIC represents an individual group; A represents the most conserved 10th trace. As PIC increases from A to J, partition comprises decreased group from 10 to 1

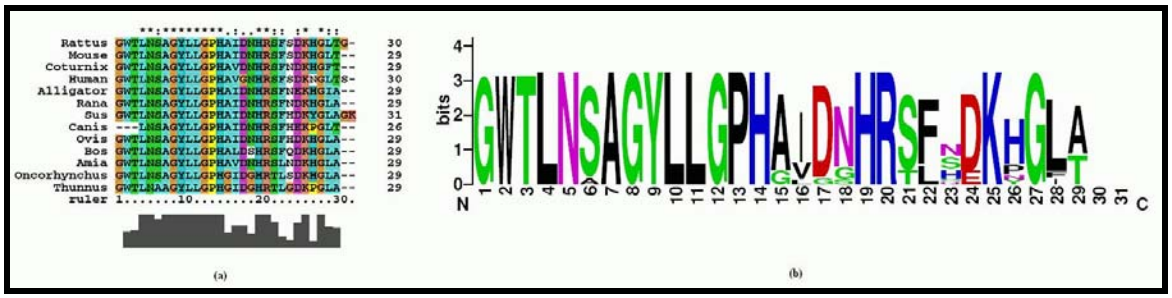

Figure 2: (**a)** Multiple sequence alignment of all galanin sequences using CLUSTALW. The conserved residues are indicated with *. (**b)** WEBLOGOI representation of 13 non redundant galanin sequences. Position 1 to 14 shows highly conserved consensus pattern

The galanin sequence was searched against PROSITE database **[9]** to extract information about the presence of functional motifs. The search results yielded a signature pattern for galanin which is at the N-terminal region (PROSITE accession number PS00861) and it picks up all galanin as well as a galanin-like peptide (GALP) that have been recently identified. Galanin-like peptide (GALP) shares sequence homology with galanin and binds to galanin receptors in vitro. **[10]** The output obtained from PROSITE was in concurrence with the results of our analysis that was done by using the three methods mentioned above.

The next step is to map the conserved residues in to the known three-dimensional structures to extract detailed information about the cluster of important amino acids, buried and exposed residues. Since there is no 3D structure for galanin, we performed Protein-protein Blast of galanin sequence from *Rattus norvegicus* against PDB database. **[11]** The search results showed similarity of galanin sequence with transportan (PDB code: 1SMZ). Transportan is a chimeric peptide constructed from 12 amino acid residues derived from the N-terminal part of the neuropeptide galanin linked with a lysine residue to the 14 amino acids of the wasp venom mastoparan. This structure of transportan reports N-terminal weak α-helix, which is a part of galanin and strong α-helix in C-terminal region, which is a part of mastoparan. Though N-terminal forms a weak helix in the transportan, it shows high conservation across galanin peptides and agonist property towards galanin receptor. The galanin part of transportan structure is not as well defined as compared with mastoparan part of transportan. **[12]** Our analysis of galanin sequences based on ET analysis, multiple sequence alignment, WEBLOGO and PROSITE results indicates a major conservation of residues in the N-terminal region. These observations lead us to strongly suggest this region as binding site for galanin.

Conclusion:

Galanin as a neuropeptide is a potential target for studying basic mechanisms of seizure initiation and arrest, and for the development of novel approaches for various neurodegenerative diseases. Our detailed sequence analysis on the galanin will be useful in exploring the structure-function relationships, crucial residues involved in binding process. Moreover, the studies on galanin using ET method would be useful in various processes such as prediction of theoretical models for galanin, studying its interaction through docking mechanisms etc. The availability of three-dimensional structure of galanin in the future would definitely support our present findings such that there would be a possibility of mapping the residues proposed by us for the prediction of the binding site of galanin.

Acknowledgement:

Authors acknowledge Prof. P. P. Mathur (Coordinator, Bioinformatics Centre, Pondicherry University, India) for constant encouragement and suggestion. Authors are also thankful to Department of Biotechnology, New Delhi, India for financial assistance.

References:

[01] A. Floren, *et al., Neuropeptides,* 34:331 (2000) [PMID: 11162289]

[02] A. Brewer, *et al., Neuropeptides,* 39 :323 (2005) [PMID: 15944029]

[03] A. J. Bordner & R. Abagyan, *Bioinformatics,* 21:2315 (2005) [PMID: 15749694]

[04] O. Lichtarge, *et al*., *J Mol Biol.,* 257:342 (1996) [PMID: 8609628

[05] A. Bairoch, *et al., Brief. Bioinform.,* 5:39 (2004) [PMID: 15153305]

[06] D. A. Benson, *et al., Nucleic Acids Res.,* 33:D34 (2005) [PMID: 15608212]

[07] J. D. Thompson, *et al., Nucleic Acids Res.* 22:4673 (1994) [PMID: 7984417]

[08] G. E. Crooks, *et al*., *Genome Research,* 14:1188 (2004) [PMID: 15173120]

[09] N. Hulo, *et al*., *Nucleic Acids Res.,* 32:D134 (2004) [PMID: 14681377]

[10] S. M. Krasnow, *et al., Endocrinology,* 144:813 (2003) [PMID: 12586757]

[11] H. M. Berman, *et al*., *Nucleic Acids Res.,* 28**:**235 (2000) [PMID: 10592235]

[12] E. Barany-Wallje, *et al., FEBS Lett.*, 567:265 (2004) [PMID: 15178334]

Edited by R. Sowdhamini

Citation: Nagarajan & Marimuthu, Bioinformation 1(5): 180-183 (2006)