ANALYSING AMINO ACIDS IN HUMAN GALANIN AND ITS RECEPTORS - GRAPH THEORETICAL APPROACH

Suresh Singh G.[a] and Akhil C. K.[b]
[a]Department of Mathematics, University of Kerala,
Kariavattom, Thiruvananthapuram – 695581, Kerala, India,

[b]Department of Mathematics, University of Kerala,
Kariavattom, Thiruvananthapuram – 695581, Kerala, India,

Acknoldgements

We wish to thank Dr Achuthsankar S. Nair, Head of the Department, Department of Computational Biology and Bioinformatics, University of Kerala, Thiruvananthapuram along with the faculty members, Vijayalakshmi B. and Biji C. L. for their fruitful suggestions and constant encouragements.

Abstract: *Graph theoretical analysis is an important area of research in biological networks. In this work we define some new graphs called bipartite Pt-graphs and their physicochemical subgraphs for peptides/proteins and their receptors based on the physicochemical properties of amino acids. Here we analyze bipartite Pt-graphs and their physicochemical subgraphs of human galanin and its three receptors graph theoretically. From the graph theoretical analysis of bipartite Pt-graphs and the physicochemical subgraphs we get some observations about the relations among the amino acids, physicochemical properties, galanin and its receptors. By a graph theoretical parameter of physicochemical subgraphs we get all the collections of maximum independent pairs of amino acids which connect the galanin and receptors by sharing exactly* $n\ (n = 1,2,3, \dots)$ *common physicochemical properties. These analyses can be used to study all the relationships between peptide/protein ligands and their receptors and this may help in the field of drug designing.*

Keywords: Amino acid, galanin, galanin receptor, bipartite Pt-graph, physicochemical subgraph.

1. Introduction

Proteins are polymers of amino acids, with each amino acid residue joined to its neighbour by a specific type of covalent bond [3]. Twenty different types of amino acids are commonly found in peptide/protein. The sequence of amino acids in a protein is characteristic of that protein and is called its primary structure [3]. Peptides/proteins are the compounds of amino acids in which a carboxyl group of one is united with an amino group of another. Neuropeptides are peptides formed and released by neurons. They are involved in a wide range of brain functions.

Galanin is a neuropeptide of 30 amino acids in humans and 29 amino acids in other species [4]. It is expressed in a wide range of tissues including the brain, spinal cord and gut. Its signaling occurs through three G protein-coupled receptors. It is linked to a number of diseases including Alzheimer's disease, epilepsy, depression, eating disorders, cancer, etc.

In [7], we can see so many graph theoretical applications in various fields. Amino acid network with in protein was studied by S. Kundu [5]. By using some physicochemical properties (Hydrophobicity, Hydrophilicity, Polarity, Non-polarity, Aliphaticity, Aromaticity and Charge (Positive and Negative)) of amino acids, the amino acid network was studied by Adil Akthar and Nisha Gohan graph theoretically [1]. The centralities in amino acid networks were used by Adil Akhtar and Tazid Ali [2]. By using the concept of amino acid network we defined and analysed the peptide/protein graph (Pt-graph) and species peptide/protein graph (SPt-graph) of galanin present in fourteen species of animals graph theoretically [6]. In this work we define and analyse new graphs - bipartite Pt-graphs and physicochemical subgraphs (physicochemical property-wise) - of human galanin and its three receptors on the basis of the physicochemical properties of amino acids. The maximum matching of physicochemical subgraphs is done to get all the collections of maximum independent pairs of amino acids which connect the galanin and its receptors by sharing exactly n $(n = 1,2,3, \dots)$ common physicochemical properties of amino acids. This method can be applied for all relationships between peptides/proteins and their receptors and this may help in the field of drug designing.

2. Basic Concepts of Graph Theory

Definition 2.1: A Graph [7] $\mathcal{G}$ is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of a finite set $\mathcal{V}$ and a set $\mathcal{E}$ of 2-element subsets of $\mathcal{V}$. The elements of $\mathcal{V}$ are called vertices and elements of $\mathcal{E}$ are called edges. The set $\mathcal{V}$ is known as the vertex set of $\mathcal{G}$ and $\mathcal{E}$ as the edge set of $\mathcal{G}$. Two vertices u and v of $\mathcal{G}$ are said to be adjacent, if an edge join u and v, and two edges are adjacent if they have common vertex. The number of vertices in a graph $\mathcal{G}$ is called its order and the number of edges is its size. A graph with p vertices and q edges is said to be a (p, q) graph.

Definition 2.2: Centrality measures in Graphs [2] are the vertex representation which gives the relative importance within the graph. A centrality is a real-valued function f which assigns every vertex $v \in \mathcal{V}$ of a given graph $\mathcal{G}$ a value $f(v) \in \mathbb{R}$.

Definition 2.3: Let $\mathcal{G}$ be an arbitrary (p, q) graph. $\mathcal{M} \subset \mathcal{E}(\mathcal{G})$ is said to be a matching in $\mathcal{G}$ if its elements are links in $\mathcal{G}$ and no two elements of $\mathcal{M}$ are adjacent in $\mathcal{G}$. $\mathcal{M}$ is said to be a maximal matching if there exists no matching $\mathcal{M}'$ of $\mathcal{G}$ with $|\mathcal{M}'| > |\mathcal{M}|$. An edge $e \in \mathcal{E}(v)$ is said to be matched under $\mathcal{M}$ (resp. unmatched under $\mathcal{M}$) if $e \in \mathcal{M}$ (resp. if $e \notin \mathcal{M}$). A vertex v is said to

be saturated by a matching $\mathcal{M}$ ($\mathcal{M}$- saturated) or matched vertex with respect to $\mathcal{M}$ if v is incident with an edge of $\mathcal{M}$. Otherwise, the vertex is said to be unsaturated by $\mathcal{M}$ ($\mathcal{M}$-unsaturated) or a single vertex with respect to $\mathcal{M}$. A matching $\mathcal{M}$ of a graph $\mathcal{G}$ is said to be a perfect matching if all the vertices of $\mathcal{G}$ are saturated by $\mathcal{M}$.

Definition 2.4: A Pt-graph is defined as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a peptide/protein in which the vertex set, $\mathcal{V}$ is the collection of all different amino acids presented in the peptide/protein and weight of a vertex in $\mathcal{G}$ is the number of times it appears in the sequence of the peptide/protein. Two vertices are said to be adjacent in $\mathcal{G}$ if they are consecutive elements in the sequence and also have at least one common physicochemical property.

For all terminologies and notations not mentioned in this work, we follow [7] (related to graph theory) and [3] (related to biology).

Remark: Weight of a vertex implies the frequency of occurrence of a specific amino acid in a sequence. Obviously greater the weight of a vertex of a Pt-graph implies greater the characteristics of those particular amino acid can be attributed to the peptide/protein. Also the centrality measures of a Pt-graph help us to determine the number of amino acids possess inter-relationships with each other.

3. Bipartite Pt-graphs and physicochemical subgraphs of human galanin and its receptors

In this section we define some new graphs called bipartite Pt-graphs and physicochemical subgraphs for peptides/proteins and its receptors. Also we construct and analyze bipartite Pt-graphs and their physicochemical subgraphs of human galanin and its receptors using $\mathcal{C}$-sets of corresponding Pt-graphs.

efinition 3.1: A bipartite Pt-graph is defined as a simple bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a peptide/protein and its receptors with $\mathcal{X}$ and $\mathcal{Y}$ as the partitions of the vertex sets of the corresponding Pt-graphs of the peptide/protein and its receptors respectively. Two vertices $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ are said to be adjacent if they have atleast one common physicochemical property.

Definition 3.2: A $\mathcal{C}$-set of a Pt-graph of a peptide/protein is defined as the subset of the vertex set whose elements are the amino acids which recieve the highest centrality measures for each physicochemical properties of amino acids.

Pt-graphs of human galanin and its receptors

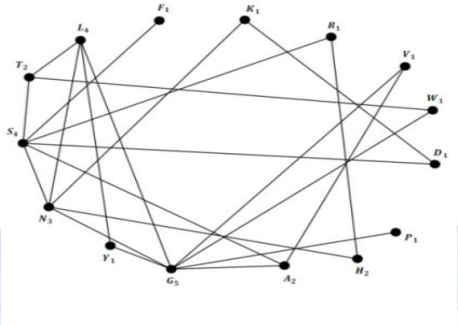

Figure 1: Pt-graph of human galanin

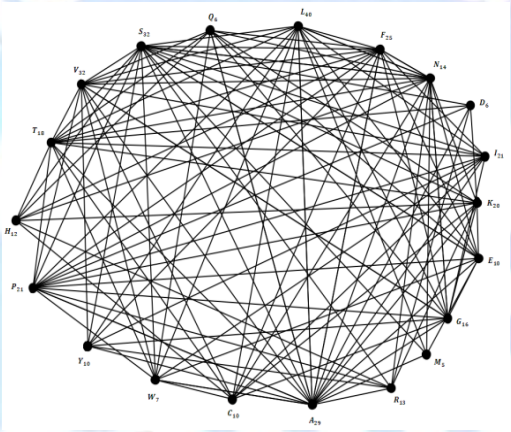

Figure 2: Pt-graph of receptor-1

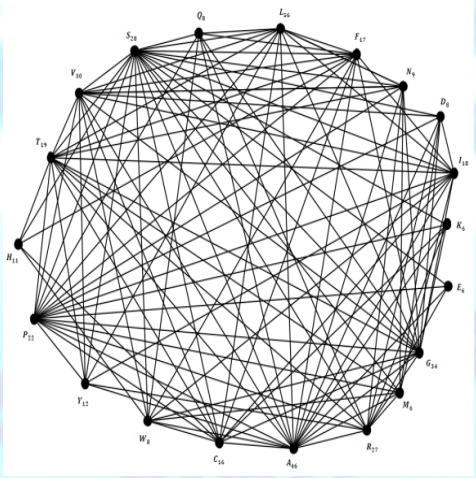

Figure 3: Pt-graph of receptor-2

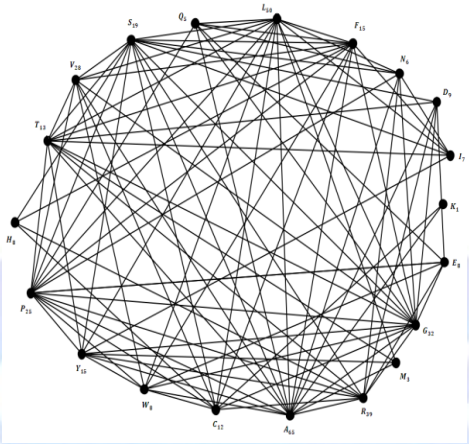

Figure 4: Pt-graph of receptor-3

From Pt-graphs we get the $\mathcal{C}$-set from the highest centralities of amino acids for each physicochemical property. For human galanin, G_5 (Hydrophoic and Non-polar), S_4 and N_3 (Hydrophilic and polar), L_4 (Aliphatic), Y_1 and W_1 (Aromatic), H_2, K_1 and R_1 (Positive), D_1 (Negative) are the amino acids which receive the highest centralities. Hence the $\mathcal{C}$-set for the galanin is $\{G_5, S_4, N_3, L_4, Y_1, W_1, H_2, K_1, R_1, D_1\}$. Similarly we get the $\mathcal{C}$-sets for the Pt-graphs of receptor-1, receptor-2 and receptor-3. Then the $\mathcal{C}$-set for receptor-1 is $\{A_{29}, N_{14}, S_{32}, L_{40}, F_{25}, K_{20}, E_{10}\}$, $\mathcal{C}$-set for receptor-2 is $\{A_{46}, S_{28}, V_{30}, F_{17}, R_{27}, D_8, I_{18}, W_8\}$ and $\mathcal{C}$ -set for receptor-3 is $\{A_{65}, G_{32}, P_{25}, S_{19}, L_{50}, F_8, R_{39}, E_8, D_9, V_{28}\}$.

Next we analyse the bipartite Pt-graphs of human galanin and its receptors.

Bipartite Pt-graphs of human galanin and its receptors

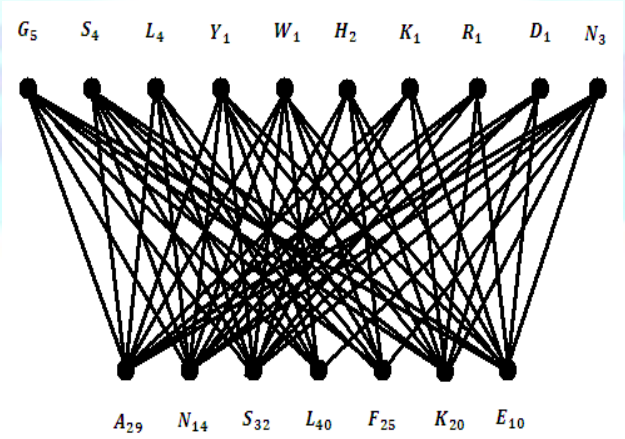

Figure 5: Bipartite Pt-graph of galanin and receptor-1

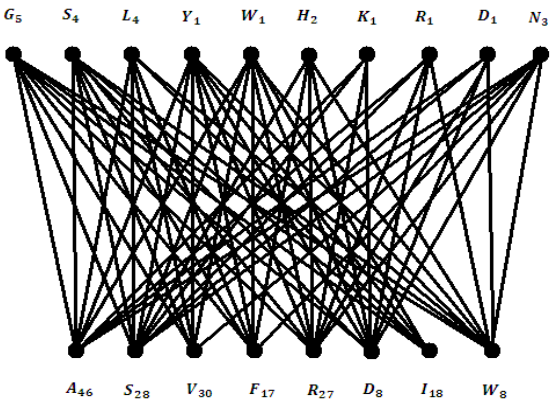

Figure 6: Bipartite Pt-graph of galanin and receptor-2

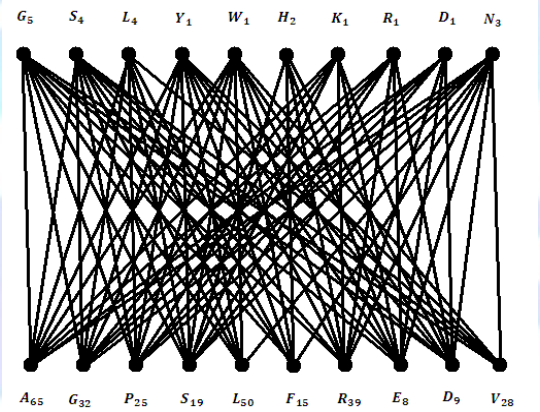

Figure7: Bipartite Pt-graph of galanin and receptor-3

Definition 3.3: Physicochemical subgraphs ${\mathcal{H}_k}^i (\text{for } i = 1,2,3,\dots)$ of a bipartite Pt-graph $\mathcal{G}$ of a peptide/protein and k receptors is defined as a subgraph whose vertex sets are same as that of $\mathcal{G}$ and two vertices in the different partitions are adjacent if they have exactly i common physicochemical properties.

Remark: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite Pt-graph of a peptide/protein and its k receptors with $\mathcal{X}$ and $\mathcal{Y}$ as the partitions of the vertex sets and let ${\mathcal{H}_k}^i(\mathcal{X}, \mathcal{Y}_i)$, where $\mathcal{Y}_i \subseteq \mathcal{Y}$ (for $i = 1,2,\dots,n$) be n physicochemical subgraphs ,
Then,

$$\bigcap_{i=1,2,3,\dots n} \mathcal{Y}_i = \emptyset \text{ and } \bigcup_{i=1,2,3,\dots n} \mathcal{Y}_i = \mathcal{Y}$$

Next we analyse physicochemical subgraphs of bipartite Pt-graphs of human galanin and its receptors. The dark edges indicate the maximum matching for physicochemical subgraphs as given below.

Physicochemical subgraphs of galanin and receptor-1

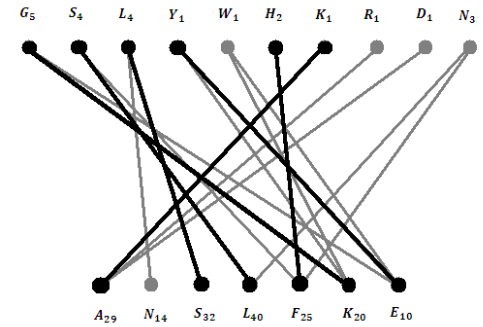

Figure 8: $\mathcal{H}_1{}^1$ **subgraph**

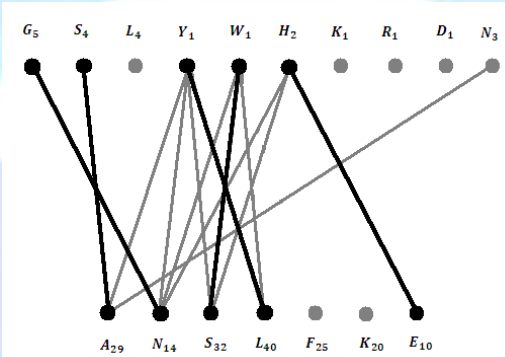

Figure 9: $\mathcal{H}_1{}^2$ **subgraph**

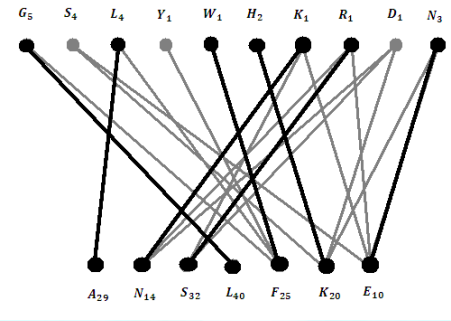

Figure 10: $\mathcal{H}_1{}^3$ **subgraph**

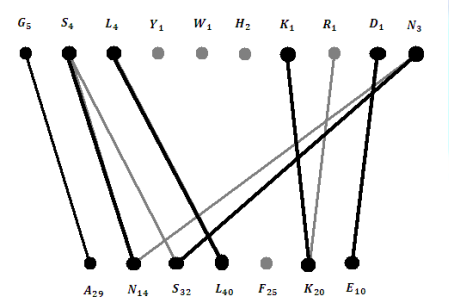

Figure 11: $\mathcal{H}_1{}^4$ **subgraph**

Physicochemical subgraphs of galanin and receptor-2

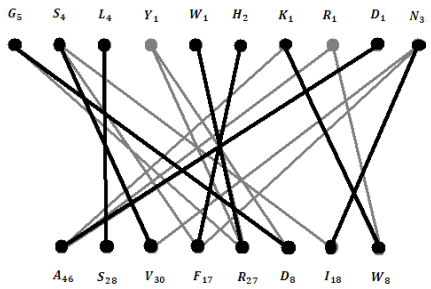

Figure 12: $\mathcal{H}_2{}^1$ **subgraph**

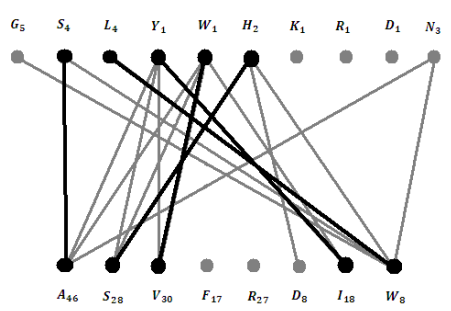

Figure 13:$\mathcal{H}_2{}^2$ **subgraph**

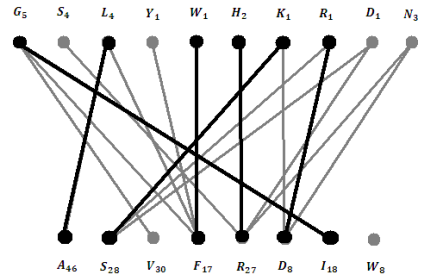

Figure 14: $\mathcal{H}_2{}^3$**subgraph**

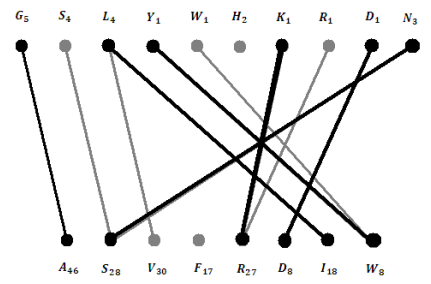

Figure 15: $\mathcal{H}_2{}^4$ **subgraph**

Physicochemical subgraphs of galanin and receptor-3

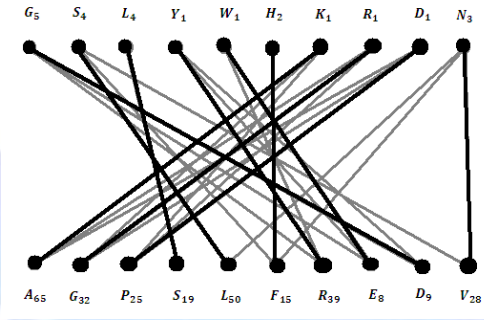

Figure 16: $\mathcal{H}_3{}^1$ subgraph

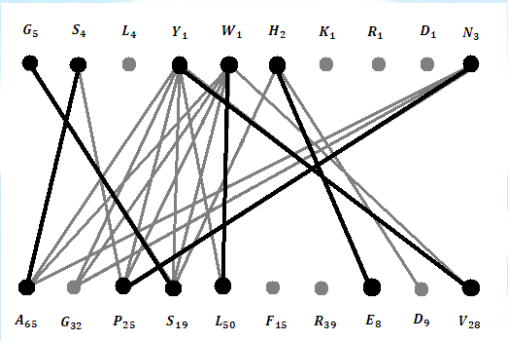

Figure 17:$\mathcal{H}_3{}^2$ subgraph

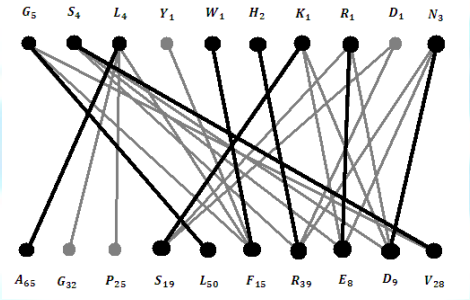

Figure 18: $\mathcal{H}_3{}^3$ subgraph

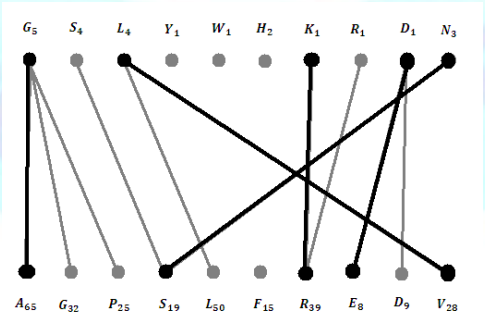

Figure 19: $\mathcal{H}_3{}^4$ subgraph

Next we obtain the maximum matching of $\mathcal{H}_k{}^i$ subgraphs of bipartite Pt-graphs of human galanin and its receptors. Table 1 represents all the collections of maximum independent pairs of amino acids which connecting the human galanin and its receptors by sharing exactly i $(i = 1,2,3,4)$ common physicochemical properties.

Receptors k	**Maximum matching of $H_k{}^i$ subgraphs**			
	$\mathcal{H}_k{}^1$ subgraphs	**$\mathcal{H}_k{}^2$ subgraphs**	**$\mathcal{H}_k{}^3$ subgraphs**	**$\mathcal{H}_k{}^4$ subgraphs**
$k = 1$	$(G_5, K_{20}), (S_4, L_{40}), (L_4, S_{32}), (Y_1, E_{10}), (H_2, F_{25}), (K_1, A_{29})$	$(G_5, N_{14}), (S_4, A_{29}), (Y_1, L_{40}), (W_1, S_{32}), (H_2, E_{10})$	$(G_5, L_{40}), (L_4, A_{29}), (W_1, F_{25}), (H_2, K_{20}), (K_1, N_{14}), (R_1, S_{32}), (N_3, E_{10})$	$(G_5, A_{29}), (S_4, N_{14}), (L_4, L_{40}), (K_1, K_{20}), (D_1, E_{10}), (N_3, S_{32})$
$k = 2$	$(G_5, D_8), (S_4, V_{30}), (L_4, S_{28}), (W_1, R_{27}), (H_2, F_{17}), (K_1, W_8), (D_1, A_{46}), (N_3, I_{18})$	$(S_4, A_{46}), (L_4, W_8), (Y_1, I_{18}), (W_1, V_{30}), (H_2, S_{28})$	$(G_5, I_{18}), (L_4, A_{46}), (W_1, F_{17}), (H_2, R_{27}), (K_1, S_{28}), (R_1, D_8)$	$(G_5, A_{46}), (L_4, I_{18}), (Y_1, W_8), (K_1, R_{27}), (D_1, D_8), (N_3, S_{28})$
$k = 3$	$(G_5, D_9), (S_4, L_{50}), (L_4, S_{19}), (Y_1, R_{39}), (W_1, E_8), (H_2, F_{15}), (K_1, A_{65}), (R_1, G_{32}), (D_1, P_{25}), (N_3, V_{28})$	$(G_5, S_{19}), (S_4, A_{65}), (Y_1, V_{28}), (W_1, L_{50}), (H_2, E_8), (N_3, P_{25})$	$(G_5, L_{50}), (S_4, V_{28}), (L_4, A_{65}), (W_1, F_{15}), (H_2, R_{39}), (K_1, S_{19}), (R_1, E_8), (N_3, D_9)$	$(G_5, A_{65}), (L_4, V_{28}), (K_1, R_{39}), (D_1, E_8), (N_3, S_{19})$

Table 1: Maximum matching of $\mathcal{H}_k{}^i$ subgraphs of human galanin and its receptors

Let $\mathcal{G}_k(\mathcal{X}, \mathcal{Y}_k)$ (for receptors $k = 1,2,3$) be three bipartite Pt-graphs of human galanin and its receptors, where $\mathcal{X}$ is the vertex set of Pt-graph of human galanin and $\mathcal{Y}_1, \mathcal{Y}_2$ and $\mathcal{Y}_3$ are the vertex sets of Pt-graphs of receptor-1, receptor-2 and receptor-3 respectively. Also let $\mathcal{H}_k{}^i(\mathcal{X}, \mathcal{Y}_k)$ be the physicochemical subgraphs of $\mathcal{G}_k(\mathcal{X}, \mathcal{Y}_k)$, where i indicates the number of common physicochemical properties of amino acids. Also let we denote the amino acids with physicochemical properties as

$P_1{}^+ =$ Hydrophobic, $P_1{}^- =$ Hydrophilic,
$P_2{}^+ =$Polar, $P_2{}^- =$ Non-polar,
$P_3{}^+ =$ Aliphatic, $P_3{}^- =$ Aromatic, $P_3{}^0 =$ Neutral (aliphatic nor aromatic),
$P_4{}^+ =$ Positive charge, $P_4{}^- =$ Negative charge, $P_4{}^0 =$ Neutral in charge.

By analysing ${\mathcal{H}_k}^i$ subgraphs, we get some observations about the physicochmical property-wise connections of amino acids of galanin and its receptors.

Observation 3.1: In the analysis of the bipartite Pt-graphs of human galanin and receptors, we obtain N_{14}, S_{32} (in receptor 1), S_{28}, W_8 (in receptor 2) and S_{19} (in receptor 3) are the amino acids which receiving all highest centralities. Also we obtain G_5, S_4, Y_1, W_1 and N_3 are the common amino acids of human galanin which receive the highest centralities in all the bipartite Pt-graphs.

Observation 3.2: In the physicochemical subgraphs with amino acids sharing exactly one common property, the connections of amino acids of galanin to receptor-1, receptor-2 and receptor-3 are

(1) ${P_1}^+$ is not connected with ${P_1}^+$ and ${P_1}^-$ is not connected with ${P_1}^-$
(2) ${P_2}^-$ is not connected with ${P_2}^-$.
(3) ${P_3}^+$ is not connected with both ${P_3}^+$ and ${P_3}^-$ but ${P_3}^-$ is not connected with ${P_3}^+$.
(4) ${P_4}^+$ and ${P_4}^-$ are not connected with both ${P_4}^+$ and ${P_4}^-$.

Observation 3.3: In the physicochemical subgraphs with amino acids sharing exactly two common properties, the connections of amino acids of galanin to receptor-1 and receptor-3 are

(1) ${P_1}^+$ and ${P_1}^-$ are connected with both ${P_1}^+$ and ${P_1}^-$.
(2) ${P_2}^-$ is not connected with ${P_2}^-$.
(3) ${P_3}^+$ is not connected with, ${P_3}^+$, ${P_3}^-$ and ${P_3}^0$
${P_3}^-$ is not connected with ${P_3}^-$
${P_3}^0$ is not connected with ${P_3}^+$ and ${P_3}^-$.
(4) ${P_4}^+$ is not connected with ${P_4}^-$
${P_4}^-$ is not connected with ${P_4}^+$, ${P_4}^-$ and ${P_4}^0$
${P_4}^0$ is not connected with ${P_4}^+$ and ${P_4}^-$.

Observation 3.4: In the physicochemical subgraphs with amino acids sharing exactly two common properties, the connections of amino acids of galanin to receptor-2 are

(1) ${P_1}^+$ and ${P_1}^-$ are conected with both ${P_1}^+$ and ${P_1}^-$.
(2) ${P_2}^-$ is not connected with ${P_2}^-$.
(3) ${P_3}^+$ is not connected with ${P_3}^+$ and ${P_3}^0$
${P_3}^-$ is not connected with ${P_3}^-$
${P_3}^0$ is not connected with ${P_3}^+$.
(4) ${P_4}^+$ is not connected with ${P_4}^-$
${P_4}^-$ is not connected with ${P_4}^+$, ${P_4}^-$ and ${P_4}^0$
${P_4}^0$ is not connected with ${P_4}^+$ and ${P_4}^-$.

Remark: In the physicochemical subgraphs with amino acids sharing exactly two common properties, the amino acids ${P_3}^+$ and ${P_3}^0$ of galanin are conncted with the amino acids ${P_3}^-$ of receptor-2 and not with receptor-1 and receptor-3. The only aromatic amino acid Tryptophan (W) of receptor-2 is conneceted to the aliphatic and neutral (neither aliphatic nor aromatic) amino acids of galanin.

Observation 3.5: In the physicochemical subgraphs with amino acids sharing exactly three common properties, the connections of amino acids of galanin to receptor-1, receptor-2 and receptor-3 are

(1) P_1^+ is not connected with P_1^- and P_1^- is not connected with P_1^+.
(2) P_2^- is not connected with P_2^+.
(3) P_3^+ and P_3^- are not connected with P_3^+.
(4) P_4^+ is not connected with P_4^+.

Observation 3.6: In the physicochemical subgraphs with amino acids sharing exactly four common properties, the connections of amino acids of galanin to receptor-1, receptor-2 and receptor-3 are

(1) P_1^- have more neighbours than P_1^+ of X.
(2) P_2^- have more neighbours than P_2^+ of X.
(3) P_4^- have more neighbours than P_4^+ of X.

Observation 3.7: There is no neighbours for P_3^- in receptor-2 of the physicochemical subgraph with amino acids sharing exactly four common properties.

Observation 3.8: The physicochemical subgraph of galanin and receptor-3 with amino acids sharing exactly one common property (figure 16) is the subgraph having a maximum matching which is the only perfect matching among all the physicochemical subgraphs of galanin and its receptors.

Conclusion

Here we analysed the amino acids and some of their physichochemical properties which involved in the human galanin neuropeptide and its receptors graph theoretically. We have constructed and analysed some newly defined graphs - bipartite Pt-graphs and their physicochemical subgraphs - of galanin and its receptors using $\mathcal{C}$-sets of the corresponding peptide/protein graphs (Pt-graphs). We observed from the analysis of the bipartite Pt-graphs of galanin and its receptors that G_5, S_4, Y_1, W_1 and N_3 (in galanin), N_{14}, S_{32} (in receptor 1), S_{28}, W_8 (in receptor 2) and S_{19} (in receptor 3) are the amino acids which receive all the highest centralities. The analysis of physicochemical subgraphs shows that, (i) if the amino acids of galanin and its receptors share exactly two common properties, the aliphatic and neutral (neither aliphatic nor aromatic) amino acids of galanin are connected with an aromatic amino acids (ie., Tryptophan (W)) only in receptor 2 (figure 13). (ii) if the amino acids of galanin and its receptors share exactly three common properties, (a) hydrophilic amino acids of galanin are more connected than hydrophobic amino acids, (b) non-polar amino acids of galanin are more connected than polar amino acids and (c) negatively charged amino acids of galanin are more connected than positively chareged amino acids. The maximum matching of physicochemical subgraphs shows that Leucine (L), Glutamate (E) and Alanine (A) (in receptor 1), Alanine (A) and Isoleucine (I) (in receptor 2), Serine (S), Alanine (A) and Valine (V) (in receptor 3) and Glycine (G), Leucine (L) and Aspargine (N) (in galanin) are the most repeated amino acids in the independent pairs. The physicochemical subgraph of galanin and receptor 3 with amino acids

sharing exactly one common property (figure 16) is the subgraph having a maximum matching which which is the only perfect matching among all the physicochemical subgraphs of galanin and its receptors. These analyses can be used to study all the relationships between peptide/protein ligands and their receptors and this may help in the field of drug designing.

References

[1] Adil Akhtar and Nisha Gohain, Graph theoretic approach to analyze amino acid network, Int. J. Adv. Appl. Math. And Mech. 2(3) (2015) 31-37.

[2] Adil Akhtar and Tazid Ali, Analysis of Unweighted Amino Acids Network, Hindawi Publishing Corporation, Vol. 2014, Article ID 350276 (2014) 6 pages.

[3] David L. Nelson and Michael M. Cox, Lehninger Principles of Biochemistry, W. H Freeman and Company (2008).

[4] Mitsukawa K., Lu X., Bartfai T., Galanin, Galanin receptors and drug targets. Cell. Mol. Life. Sci. 65, 1976-180510.1007/s00018-008-8153-8 (2008)

[5] S. Kundu, Amino acid network with in protein, Physica A, 346 (2005) 104-109.

[6] Suresh Singh G., Akhil C. K., Analysing Amino Acids in Galanin – Graph Theoretical Appraoch, International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), June 17 Volume 5 Issue 6, ISSN: 2321-8169, PP: 342-346 (2017).

[7] Suresh Singh G., Graph Theory, PHI Learning Private Limited (2010).

[8] Yang Wang, Mingnia Wang, Sanwen Yin, Richard Jang, Jian Wang, Zhidong Xue and Tao Xu, Neoropep: a comprehensive Resource of Neuropeptides, Database (Oxford) 2015 Apr 29. doi:10.1093/ database/ bar038.