International Standard Serial Number (ISSN): 2249-6807

*International Journal of Institutional Pharmacy and Life Sciences 4(5): September-October 2014*

*INTERNATIONAL JOURNAL OF INSTITUTIONAL PHARMACY AND LIFE SCIENCES*

**Life Sciences** **Original Article……!!!**

Received: 15-09-2014; Revised; Accepted: 07-10-2014

# STUDY OF CERVICAL CANCER HOX-B4 PROTEIN WITH THE SELECTED FIVE SPECIES

P. Ruba Glory and K. Palanivelu*

P.G. and Research Dept. of Zoology, Govt. Arts College (Autonomous), Kumbakonam - 612 001, Tamil Nadu, India.

**Keywords:**

Hox-B4, Sequence, Species. LMICs

**For Correspondence:**

**Dr. K. Palanivelu**

P.G. and Research Dept. of Zoology, Govt. Arts College (Autonomous), Kumbakonam - 612 001, Tamil Nadu, India.

**E-mail:**

palanivelu59@gmail.com

## ABSTRACT

Hence the present study is undertaken to analyze the Hox-B4 protein of Human (*Homo sapiens*), with the selected different species, such as Mouse (*Mus musculus*), Chick (*Gallus gallus*), Dove (*Columba livia*), and Fish (*Salmo salar*). Homeobox protein Hox-B4 is a protein that in humans is encoded by the HOXB4 gene. The Cervical cancer is one of the most common cancers among women worldwide (WHO, 2009). Its mortality exemplifies health inequity, as its rates are higher in low and middle income countries (LMICs), WHO and in low socio-economic groups within countries (Kurkure and Yeole, 2006). Around 80% of global cervical cancer cases are in LMICs (Waggoner, 2003). Cancer of the cervix is primarily caused by human papillomavirus (HPV) infection, for which there is a vaccination now available (Miller *et al*, 1990). The selected five species sequence was collected from Fasta. Human HoxB4 Protein sequence was retrieved from the SWISS-PROT protein sequence data base. Protparam to used the Compute molecular weight, theoretical pI, amino acid composition, atomic composition, extinction co-efficients, estimated half-like, instability index, aliphatic index and grand average of hydropathicity. RasMol to used Molecular graphics program intended for the visualization of proteins, nucleic acids and small molecules.

## INTRODUCTION

Bioinformatics in 1970 to refer to the study of information processes in biotic systems [Hogeweg *et.al.*, 2011]. This definition placed bioinformatics as a field parallel to biophysics or biochemistry [Hogeweg *et.al.*, 1978].

**Cervical Cancer:** It is a major health issue, ranking as the second most frequent cancer among women throughout the world [Goggin and Mayrand, 2009]. It remains the second most frequent cancer among Canadian women 25 to 44 years of age [Parkin and Bray, 2006]. In Quebec, it is estimated that 270 women were diagnosed with cervical cancer in 2010 and that 65 died of it [Marrett *et.al.*, 2002]. Cancer refers to a class of diseases in which a cell or a group of cells divide and replicate uncontrollably, intrude into adjacent cells and tissues and ultimately spread to other parts of the body than the location at which they arose (National Cancer Registry Programme, 2002). In cervical cancer, cancer develops in the tissues of the cervix, which is a part of the female reproductive system (Canadian Cancer Society, 2010). The cervix connects the upper body of the uterus to the vagina. The endocervix is covered by glandular cells, and the ectocervix is covered by squamous cells. The transformation zone refers to the place where these two regions of the cervix meet (American Cancer Society, 2009).

**Homeobox protein Hox-B4:** It is a protein that in humans is encoded by the HOXB4 gene (McAlpine and Shows, 1990). It is included in a cluster of homeobox B genes located on chromosome 17. The encoded protein functions as a sequence-specific transcription factor that is involved in development (Scott, 1992).

**Five Species:** Hox-B4 protein of Human (*Homo sapiens*), with the selected different species, such as Mouse (*Mus musculus*), Chick (*Gallus gallus*), Dove (*Columba livia*), and Fish (*Salmo salar*).

## MATERIALS AND METHODS

**Swiss-Prot:** The following materials were used. Swiss-Prot is an annotated protein sequence database. The Hox-B4 protein was collected from Swiss-Prot. **Fasta:** DNA and protein sequence alignment software package. The selected five species sequence was collected from Fasta.

**Protparam:** Used the Compute molecular weight, theoretical pI, amino acid composition, atomic composition, extinction co-efficients, estimated half-like, instability index, aliphatic index and grand average of hydropathicity. (i) Physico-Chemical characterization of Homeobox protein, (ii) Calculation of extinction coefficient, (iii) Calculation of half-life of a protein.

**RasMol**: Molecular graphics program intended for the visualization of proteins, nucleic acids and small molecules.

**RESULTS**

The number of different amino acids and their percentage composition are determined. The number of amino acids composition are more less equal in human, mouse and fish and their molecular weight is different from one another (Table: 1). The number of amino acids are equal in chick and mouse and its molecular weight is different from one another are determined. (Table: 2). The total number of atoms gradually reduced from human > mouse = fish > chick > dove (Table: 4). The estimated half-life of reticulocytes invitro are more or less equal in human, mouse and fish. But the estimated half-life of reticulocytes invitro are more or less reduced in mouse and chick (Table: 3). Estimated half-life of human, mouse, fish, chick and dove are similar (invitro). The percentage of amino acids composition of glutamic, alanine, leucine and valaine are higher than the other amino acids in human, mouse and fish. The percentage composition of sulphur containing amino acids methionine and cysteine are observed to be low in human, mouse and fish. In mouse and chick the percentage of amino acids glutamic, valine, alanine, leucine are gradually reduced when compared with human, mouse and fish.

**Table:1 - Number Of Amino Acids Present In The Hox-B4 Protein**

| S. NO | NAME OF THE AMINO ACIDS | NUMBER OF AMINO ACIDS | | | | |
|---|---|---|---|---|---|---|
| | | ***Homo sapiens*** | ***Mus musculus*** | ***Gallus gallus*** | ***Columba livia*** | ***Salmo salar*** |
| 1 | Alanine | **22** | **21** | **12** | **12** | **10** |
| 2 | Ariginine | **23** | **22** | **17** | **18** | **18** |
| 3 | Asparagines | **9** | **9** | **13** | **11** | **19** |
| 4 | Aspartic acid | **5** | **5** | **3** | **3** | **7** |
| 5 | Cysteine | **7** | **6** | **7** | **7** | **5** |
| 6 | Glutamine | **10** | **11** | **15** | **10** | **14** |
| 7 | Glutamic acid | **13** | **13** | **15** | **15** | **11** |
| 8 | Glycine | **19** | **19** | **7** | **5** | **11** |
| 9 | Histidine | **7** | **7** | **12** | **11** | **8** |
| 10 | Isoleucine | **5** | **5** | **6** | **5** | **8** |
| 11 | Leucine | **13** | **13** | **15** | **13** | **15** |
| 12 | Lysine | **11** | **11** | **12** | **13** | **12** |
| 13 | Methionine | **4** | **4** | **5** | **5** | **4** |
| 14 | Phenyl alanine | **6** | **6** | **5** | **5** | **4** |
| 15 | Proline | **46** | **45** | **27** | **25** | **30** |
| 16 | Serine | **21** | **20** | **42** | **35** | **30** |
| 17 | Threonine | **6** | **8** | **7** | **7** | **14** |
| 18 | Tryptophan | **3** | **3** | **3** | **3** | **3** |
| 19 | Tyrosine | **11** | **11** | **12** | **12** | **13** |
| 20 | Valine | **10** | **11** | **10** | **10** | **14** |
| **Total** | | **251** | **250** | **245** | **225** | **250** |

**Table:2 - Percentage Composition Of Amino Acids Of Hox-B4 Protein**

| S. NO | NAME OF THE AMINO ACID | PERCENTAGE OF AMINO ACIDS % | | | | | Total |
|---|---|---|---|---|---|---|---|
| | | *Homo sapiens* | *Mus musculus* | *Gallus gallus* | *Columba livia* | *Salmo salar* | |
| 1 | Alanine | **8.8%** | **8.4%** | **4.9%** | **5.3%** | **4.0%** | **31.4** |
| 2 | Ariginine | **9.2%** | **8.8%** | **6.9%** | **8.0%** | **7.2%** | **40.1** |
| 3 | Asparagines | **3.6%** | **3.6%** | **5.3%** | **4.9%** | **7.6%** | **25.0** |
| 4 | Aspartic acid | **2.0%** | **2.0%** | **1.2%** | **1.3%** | **2.8%** | **9.3** |
| 5 | Cysteine | **2.8%** | **2.4%** | **2.9%** | **3.1%** | **2.0%** | **13.2** |
| 6 | Glutamine | **4.0%** | **4.4%** | **6.1%** | **4.4%** | **5.6%** | **24.5** |
| 7 | Glutamic acid | **5.2%** | **5.2%** | **6.1%** | **6.7%** | **4.4%** | **27.6** |
| 8 | Glycine | **7.6%** | **7.6%** | **2.9%** | **2.2%** | **4.4%** | **24.7** |
| 9 | Histidine | **2.8%** | **2.8%** | **4.9%** | **4.9%** | **3.2%** | **18.6** |
| 10 | Isoleucine | **2.0%** | **2.0%** | **2.4%** | **2.2%** | **3.2%** | **11.8** |
| 11 | Leucine | **5.2%** | **5.2%** | **6.1%** | **5.8%** | **6.0%** | **28.3** |
| 12 | Lysine | **4.4%** | **4.4%** | **4.9%** | **5.8%** | **4.8%** | **24.3** |
| 13 | Methionine | **1.6%** | **1.6%** | **2.0%** | **2.2%** | **1.6%** | **9.0** |
| 14 | Phenyl alanine | **2.4%** | **2.4%** | **2.0%** | **2.2%** | **1.6%** | **10.6** |
| 15 | Proline | **18.3%** | **18.0%** | **11.0%** | **11.1%** | **12.0%** | **70.4** |
| 16 | Serine | **8.4%** | **8.0%** | **17.1%** | **15.6%** | **12.0%** | **61.1** |
| 17 | Threonine | **2.4%** | **3.2%** | **2.9%** | **3.1%** | **5.6%** | **17.2** |
| 18 | Tryptophan | **1.2%** | **1.2%** | **1.2%** | **1.3%** | **1.2%** | **6.1** |
| 19 | Tyrosine | **4.4%** | **4.4%** | **4.9%** | **5.3%** | **5.2%** | **24.2** |
| 20 | Valine | **4.0%** | **4.4%** | **4.1%** | **4.4%** | **5.6%** | **22.5** |

**Table:3 -Positive And Negative Charged Amino Acid Of Hox-B4 Protein**

| S. NO | SPECIES NAME | MOLECULAR WEIGHT | NO OF + VE CHARGED AMINO ACIDS | | | NO OF –VE CHARGED AMINO ACIDS | | |
|---|---|---|---|---|---|---|---|---|
| | | | ARG | LYS | TOTAL | ASP | GLU | TOTAL |
| 1 | *Homo sapiens* | **27604.3** | 23 | 11 | **34** | 5 | 13 | **18** |
| 2 | *Mus musculus* | **27519.2** | 22 | 11 | **33** | 5 | 13 | **18** |
| 3 | *Gallus gallus* | **27782.0** | 17 | 12 | **29** | 3 | 15 | **18** |
| 4 | *Columba livia* | **25803.0** | 18 | 13 | **31** | 3 | 15 | **18** |
| 5 | *Salmo salar* | **28231.6** | 18 | 12 | **30** | 7 | 11 | **18** |

**Table:4 - Atomic Composition Of Hox-B4 Protein**

| S.NO | SPECIES NAME | NUMBER OF ATOMS | | | | | TOTAL |
|---|---|---|---|---|---|---|---|
| | | C | H | N | O | S | |
| 1 | *Homo sapiens* | 1223 | 1887 | 367 | 345 | 11 | **3833** |
| 2 | *Mus musculus* | 1221 | 1884 | 364 | 346 | 10 | **3825** |
| 3 | *Gallus gallus* | 1207 | 1865 | 363 | 371 | 12 | **3818** |
| 4 | *Columba livia* | 1127 | 1742 | 338 | 337 | 12 | **3556** |
| 5 | *Salmo salar* | 1234 | 1920 | 368 | 377 | 9 | **3908** |

**DISCUSSION AND SUMMARY**

The sequence of Hox-B4 protein in Fasta format was obtained from Swiss-Prot using online services (www.expasy.ch). The molecular weight, number of amino acids, percentage of amino acids and atomic composition was analyzed by Protparam. The molecular structure of Hox-B4 protein was identifies with Rasmol tool. The number of different amino acids and their percentage composition are determined. The number of different amino acids and their percentage composition were determined. Further, the amino acids were grouped according to their number of types and their percentage composition. The number of different amino acids and their percentage composition of Hox-B4 precursor protein in human, mouse and fish are more or less equal. An analysis showed that the percentage composition of Proline (46, 45, 30), Arginine (23, 22, 18), Alanine (22, 21, 10), and Serine (21, 20, 30), are higher than the other amino acids. The Proline (46, 45, 30), Ariginine (23, 21, 10), Alanine (22, 21, 10), and Serine (21, 20, 30) are aliphatic amino acids and glutamic acid (13, 13, 11) is basic amino acids. It can be noted that these are hydrophobic amino acids. Moreover, the percentage composition of sulphur containing amino acids the tryptophan and methionine are observed to be low. Among the aromatic amino acids the occurrence of isoleucine, threonine and cysteine are less compared to phenyl alanine. Both the negatively charged amino acids aspartic acids and glutamic are approximately equal composition, where as among the positively charged amino acids these are acidic amino acids. The composition of arginine is higher than lysine. These kinds of amino acids are called basic amino acids. The number of amino acids and the percentage composition of hydrophobic amino acids are higher than the hydrophilic amino acids.

The number of different amino acids and their percentage composition of Hox-B4 protein in human, mouse and fish are more or less equal. An analysis showed that the percentage composition of Proline, Serine, Arginine and Alanine are higher than the amino acids. It can be

noted that all these hydrophobic amino acids. Moreover, percentage composition of Tryptophan, Methionine and Aspartic acid are observed to be low. Here all these are Hydrophilic amino acids. Hence the number of amino acids and the percentage composition of hydrophobic amino acids are higher than the Hydrophilic amino acids. The estimated half-life of reticulocytes invitro are more or less equal in human, mouse and fish. But the estimate half- life of reticulocytes invitro are more or less reduced in chick and dove. The estimated half-life of human, mouse, chick, dove and fish are similar (*invivo*). Total number of atom is gradually reduced as human > mouse = fish > chick > dove > The number of amino acids composition of Hox-B4 protein are more or less equal in human, mouse, and fish but molecular weight is different from each other. The number of amino acids composition are equal in chick and mouse but molecular weight is different from one another.

**REFERENCES**


1. American Cancer Society (2009).“Detailed guide: Cervical Cancer” ACS website. http://www.cancer.org/docroot/CRI/content/CRI 241X what is cervical cancer8/asp.
2. Canadian Cancer Society and National Cancer Institute of Canada (2010). Canadian Cancer Statistics. Toronto, Canada.
3. Goggin P, Mayrand M-H (2009). Recommendations on optimizing cervical cancer screening in Quebec. Institute national de sante publique du Quebec.
4. Hogeweg, P (1978). "Simulating the growth of cellular forms". Simulation 31 (3): 90–96. doi:10.1177/003754977803100305.
5. Hogeweg, P (2011). "The Roots of Bioinformatics in Theoretical Biology". In Computational Biology (3) :e1002021. Bibcode:2011 PLSCB.7E0020H. doi:10.1371 /journal. pcbi.1002021. PMC 3068925. PMID 21483479.
6. Kurkure AP, and Yeole BB (2006). “Social inequalities in cancer with special reference to South Asian countries”, Asian Pacific Journal of Cancer Prevention, 7(1) : 36-40.
7. Marrett LD, Frood J, Nishri D, Ugnat AM (2002). Cancer incidence in young adults in Canada: preliminary results of a cancer surveillance project. Chronic Dis Can;23:58-64.
8. McAlpine PJ, Shows TB (1990). "Nomenclature for human homeobox genes". *Genomics* 7 (3): 460. doi:10.1016/0888-7543(90)90186-X. PMID 1973146.
9. Miller AB., Chamberlain J, Day, NE, Hakama, M, and Prorok PC (1990). “Report on a workshop of the UICC project on evaluation of screening for cancer” International Journal of Cancer, 46: 761-9.

10. National Cancer Registry Programme and World Health Organisation (2002). Development of an atlas of cancer in India, “First All India Report, Mapping Patterns of Cancer in India”.
11. Parkin DM, Bray F (2006). Chapter 2: The burden of HPV-related cancers. Vaccine Aug 21;24 Suppl 3:S11-25.:S11-S25.
12. Scott MP (1992). "Vertebrate homeobox gene nomenclature". *Cell* 71 (4): 551–3. doi: 10.1016 / 0092 - 8674(92)90588-4. PMID 1358459.
13. Waggoner SE (2003). “Cervical Cancer,” Lancet, 361: 2217- 25.
14. Waterman, Michael S (1995). Introduction to Computational Biology: Sequences, Maps and Genomes. CRC Press. ISBN 0-412-99391-0.
15. World Health Organisation (2009). Country Office for India. “Core Programme Clusters, Non Communicable Diseases & Mental Health (NMH): Major Non Communicable Diseases Cancer”.