Journal of PharmaSciTech
ISSN: 2231 3788 (Print)
2321 4376 (Online)
Research Article

Structural Insight of Homeobox DNA Binding Domain of Hox-B7A Protein of *Esox lucius*

Subhamay Panda[1, 2]*, Leena Kumari[1], Subhra Prakash Hui[3] and Santamay Panda[2,4]

[1]*Department of Pharmacy, Gupta College of Technological Sciences, Ashram More, G.T. Road, Asansol-713301, India*
[2]*Indian Institute of Human and Social Sciences, Sitarampur, Asansol-713359, India*
[3]*Developmental and Stem Cell Biology Division, Victor Chang Cardiac Research Institute, Lowy Packer Building, 405 Liverpool St, Darlinghurst, New South Wales 2010, Australia*
[4]*Department of Physics, NSHM Faculty of Engineering & Technology, NSHM Knowledge Campus, Durgapur-713212, India*

Correspondence: subhamay_panda@rediffmail.com, (Tel. +91 9531600015)

Abstract

Homeobox (Hox) genes are important metazoan developmental genes they dictate the identity of embryonic regions along the anterio-posterior axis. A homeobox DNA binding region exists in homeobox protein Hox-B7a along the length of its sequence. The objective of the present study is to evaluate the DNA binding domain of Homeobox protein Hox-B7a of *Esox lucius* (Northern pike) with special reference to structure generation, validation of the generated models, distribution of secondary structural elements and positive charge distribution over the structure. With the use of comparative modeling approach to we propose the first 3D structure of DNA binding region of Hox-B7a *Esox lucius*. The current study focuses on the understanding of evolutionary structural enrichment strategy of DNA binding region of Hox-B7a. The appearance of different secondary structural element over structure provides for the molecule specific uniqueness of DNA binding region of Hox-B7a of *Esox lucius*.

Keywords: Homeobox genes, *Esox lucius*, DNA binding region

Introduction

Homeobox (Hox) genes are a set of transcription factors has been studied extensively in diverse fields of molecular and evolutionary biology. This protein family plays a very vital role in anterior-posterior axis patterning of animal embryos and in the development of tetrapod limb [1- 2]. The common feature of Hox proteins is the presence of 60 amino acids motif called homeodomain and Hox genes belong to the extensive superfamily of homeobox transcription factors [3].

The organisation of Hox genes within the genome in clusters is commonly found in several animals. A cluster of ancestral Hox gene, which is known to have originated from tandem duplications in early eukaryotic species, has been found in all bilaterian animals. Although Hox genes have diverged in various species but the homeodomain protein motif has remained highly conserved. Thus a given Hox gene can be easily assigned by means of homology to one of the genes along the cluster. Hox genes come under the category of one of the 14 known Paralogous Groups (PG). The duplication of ancestral cluster has been accomplished early in the vertebrate lineage [4-5]. In case of mammals, Hox genes are arranged in four clusters whereas teleost Hox genes are usually organised on 7 clusters, which resulted because of an additional duplication specific to teleost fishes [6-7]. The subsequent occurrence of lineage-specific gene has been affected, followed by diverse presence/absence combinations of Hox genes [8].

Hox genes encode a major group of evolutionarily conserved transcription factors controlling various functions such as axis specification and patterning of the central nervous system during embryonic development [9]. The Hox genes in vertebrates comprises of 13 paralog groups, expressed in the form of clusters on different chromosomes. In tetrapods, there are at least 39 genes are organized in 4 clusters, HoxA, HoxB, HoxC, and HoxD, whereas in teleosts, 4 clusters are found which resulted due to duplication of a whole genome early in their lineage [10-11].

The duplicated chromosomal regions are best exemplified by the occurrence of clusters of Hox genes [12-14]. Hox transcription factors are generally characterized by the presence of their DNA binding domain, the homeodomain. Hox genes were first discovered in the fruit fly Drosophila as the target gene for undergoing homeotic mutation, in which alteration in the segmental identity takes place, as in case of bithorax phenotype [15]. Hox genes are especially featured by their arrangement in the form of genomic clusters. A single cluster is found in all invertebrates that are either interrupted as in Drosophila species [16] or are found dispersed through the genome as in urochordates [17-18] and nematodes [19], whereas four clusters are a unique feature of the tetrapods such as human or frogs [8, 20], as akin to cartilaginous fish [21]. The invertebrates which are in close relation to vertebrates, such as the cephalochordate Branchiostoma [13, 22] possess a single cluster, which is also rearranged in case of the sea urchin [23]. The genome duplication in fish-specific genome resulted in the occurrence of seven Hox clusters in extant fish, with alternate cluster loss in Ostariophysi (HoxDb in zebrafish) [6] and Acanthopterygii (HoxCb in pufferfish, medaka, cichlid) [11, 24-25]. The additional clusters are not exactly alike to the homologous genes of tetrapods, but they usually undergoes independent losses of genes [8], thereby rendering much more variations in the gene content in teleost clusters than those of tetrapods. All of the fish species investigated till date experiences differences in gene content among their Hox clusters [8, 26-27].

The morphological variation and functional innovation relies upon the duplication of genes and entire genomes [12, 28-29]. During the evolution of the vertebrates, the clusters of Hox genes have experienced several rounds of duplication. Two rounds of genome duplication took place in case of jawed vertebrates, resulting in the production of four canonical Hox clusters of most gnathostomes designated as "HoxA", "HoxB", "HoxC", and "HoxD" clusters. In the case of subset of ray-finned fishes, a third round of Hox cluster duplication have been affected [6, 30-31] and forms seven to eight clusters referred to as "Aa," "Ab," etc.

The objective of the present study is to evaluate the DNA binding domain of Homeobox protein Hox-B7a of *Esox lucius* (Northern pike) with reference to structure generation, validation of the generated models, distribution of secondary structural elements and positive charge distribution over the structure with the support of various bioinformatical algorithms.

Materials and Methods

Amino acid sequence of DNA binding domain of Hox-B7a of *Esox lucius* (Northern pike) was obtained from National Centre for Biotechnology Information (http://ncbi.nlm.nih.gov) [32-33]. Comparative model of DNA binding domain of Hox-B7a of *Esox lucius* was generated by the comparative modeling tool Swiss-PDB Viewer [34]. After generation of molecular model of DNA binding domain of Hox-B7a of *Esox lucius* was subjected to further structural improvements by energy minimization step by Swiss-PDB Viewer [35]. The validation for three-dimensional structural model obtained by molecular modeling approach was assessed by PROCHECK algorithm and ProSA tool [36-37]. Distribution of secondary structural elements and positive charge distribution over the structure was performed with the assistance of UCSF Chimera package [38-40].

Results and Discussion

Hox genes are important metazoan developmental genes they dictate the identity of embryonic regions along the anterio-posterior axis [41, 25]. A homeobox DNA binding region exists in homeobox protein Hox-B7a along the length of its sequence. The homeobox Hox gene encodes a characteristic homeodomain DNA binding protein fold that comprises of a 60-amino acid helix-turn-helix (HTH) structure where three alpha helices are joined together by short loop regions. The orientations of N-terminal two helices are antiparallel while the longer C-terminal helix arranges itself almost perpendicular to the axes of the first two helices. The third C-terminal helix undergoes a number of hydrogen bonding and hydrophobic interactions with DNA, generally occurring between specific side chains and the exposed bases and thymine methyl groups residing in the major groove of the DNA [42, 8].

The 3D structure of proteins gives us a clear idea about the various interactions and stabilizations of proteins in their stable conformation. Structural genomics and proteomics involve a comparative molecular modeling approach, which is regarded as one of the most common methods for the prediction of the structures.

Ramachandran plot analysis using overall score of G-factor (PROCHECK analysis) is a good standard for validation purpose. Ramachandran plot for DNA binding domain of Hox-B7a of *Esox lucius* has been illustrated in Figure 1. Altogether 100% of the residues were detected in allowed and favored regions, which in turn validate the quality of generated protein structural model. The overall score of G-factor for Hox-B7a of *Esox lucius* was -0.03 which was greater than the acceptable value of -0.50. PROCHECK algorithm also displayed 83.9% of residues in the most favored regions, with 16.1% residues in additionally allowed regions, respectively (Figure 1). This demonstrated that the backbone dihedral angles, phi and psi, in the DNA binding region of Hox-B7a 3D model, were reasonably accurate. This proposes that the modeled structure of Hox-B7a of *Esox lucius* is satisfactory and acceptable (Figure 2).

ProSA tool was employed to investigate three-dimensional model of Hox-B7a protein for possible errors. As shown in Figure 3 the Z-score of CX1.BEC was - 3.33. The score was well inside the range of scores usually observed for proteins of matching size indicating highly reliable structures.

The manual inspection of Hox-B7a postulates that the total protein is composed by 215 number of amino acids. The DNA binding domain of Hox-B7a of *Esox lucius* is consist of 61 amino acids. Interestingly this region populated with positively charged amino acid. The presence of total number of positively charged amino acids is 18. In contrast to that the total number of negatively charged amino acids is only 7 (Figure 4). The presence of high number of positively charged amino acids in DNA binding region of Hox-B7a due to its direct attachment with negatively charged DNA molecule in the physiological milieu. For proper structural interaction this charge complementarity gives an added advantage to the DNA and Hox-B7a complex. The molecular model of DNA binding domain possesses helix-turn-helix (HTH) structure depicting three alpha helices that are joined via short loop regions (Figure 2).

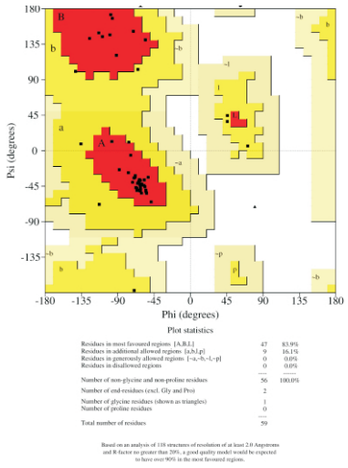

Figure 1: Ramachandran plot of molecular model of DNA binding domain (Hox-B7a of *Esox lucius*)

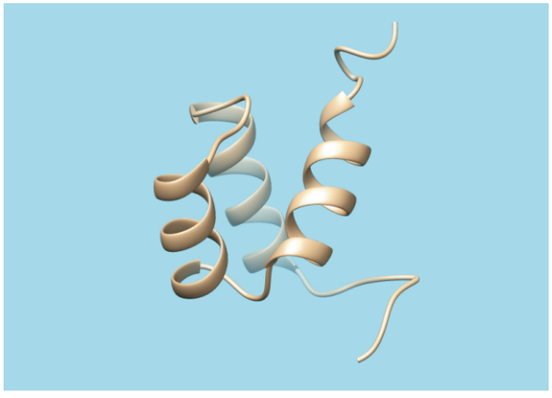

Figure 2: Three-dimensional modeled structure of DNA binding domain (Hox-B7a of *Esox lucius*).

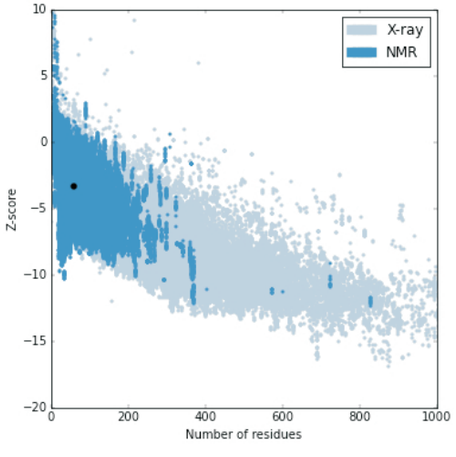

Figure 3: Stereo-chemical validation of modeled structure of DNA binding domain (Hox-B7a of *Esox lucius*).

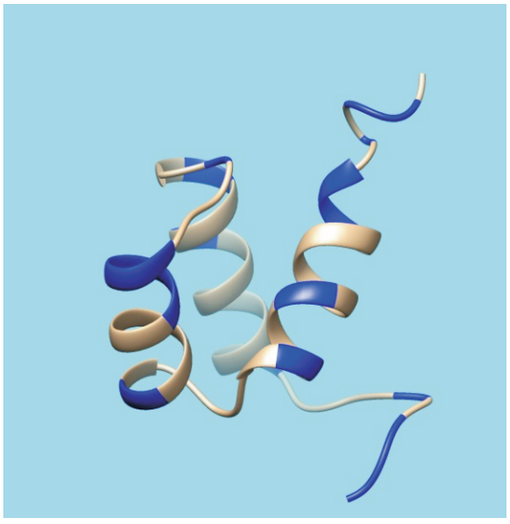

Figure 4: Positive charge distribution over the modeled structure of DNA binding domain (Hox-B7a of *Esox lucius*).

Conclusion

In the present investigation, we have used comparative modeling approach to propose the first 3D structure of DNA binding region of Hox-B7a of *Esox lucius*. The current study focuses on the understanding of evolutionary structural enrichment strategy of DNA binding region of Hox-B7a. The appearance of different secondary structural element over structure provides for the molecule specific uniqueness of DNA binding region of Hox-B7a. This reason for this structural uniqueness relies on the system specific functional necessities which drive new inventions at the structural context.

Acknowledgement

We are very thankful to Prof. Debesh Chandra Majumder, Chairman, Trinity Trust, Asansol, West Bengal, Prof. Kalyan Kumar Sen, Principal , Gupta College of Technological Sciences, Aasnsol, West Bengal for providing infrastructure facilities for carrying out the research work.

Declaration of Interest

The authors do not have any conflict of interest.

References

1. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell 1992; 68:283-302.

2. Nelson CE, Morgan BA, Burke AC et al. Analysis of Hox gene expression in the chick limb bud. Development 1996; 122:1449-66.

3. Gehring WJ, Muller M, Affolter M et al. The structure of the homeodomain and its functional implications. Trends Genet 1990; 6:323-29.

4. Finnerty JR, Martindale MQ. The evolution of the Hox cluster: insights from outgroups. Curr Opin Genet Dev 1998; 8:681-87.

5. Prince V. The Hox Paradox: More complex(es) than imagined. Dev Biol 2002; 249:1-15.

6. Amores A, Force A, Yan YL et al. Zebrafish hox clusters and vertebrate genome evolution. Science 1998; 282:1711-14.

7. Giddings SJ, Chirgwin J, Permutt MA. The effects of fasting and feeding on preproinsulin messenger RNA in rats. J Clin Invest 1981; 67:952-60.

8. Hoegg S, Meyer A. Hox clusters as models for vertebrate genome evolution. Trends Genet 2005; 21:421-24.

9. Hadrys T, Punnamoottil B, Pieper M et al. Conserved co-regulation and promoter sharing of hoxb3a and hoxb4a in zebrafish. Dev Biol 2006; 297:26-43.

10. Hurley I, Hale ME, Prince VE. Duplication events and the evolution of segmental identity. Evol Dev 2005; 7:556-67.

11. Jaillon O, Aury JM, Brunet F et al. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 2004; 431:946-57.

12. Crow KD, Stadler PF, Lynch VJ et al. The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. Mol Biol Evol 2006; 23:121-36.

13. Garcia-Fernandez J, Holland PW. Archetypal organization of the amphioxus Hox gene cluster. Nature 1994; 370:563-66.

14. Hoegg S, Boore JL, Kuehl JV et al. Comparative phylogenomic analyses of teleost fish Hox gene clusters: lessons from the cichlid fish Astatotilapia burtoni. BMC Genomics 2007; 8:317.

15. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature 1978; 276:565-70.

16. Negre B, Casillas S, Suzanne M et al. Conservation of regulatory sequences and gene expression patterns in the disintegrating Drosophila Hox gene complex. Genome Res 2005; 15:692-700.

17. Seo HC, Edvardsen RB, Maeland AD et al. Hox cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica. Nature 2004; 431:67-71.

18. Ikuta T, Yoshida N, Satoh N et al. Ciona intestinalis Hox gene cluster: Its dispersed structure and residual colinear expression in development. Proc Natl Acad Sci U S A 2004; 101:15118-23.

19. Aboobaker A, Blaxter M. Hox gene evolution in nematodes: novelty conserved. Curr Opin Genet Dev 2003; 13:593-98.

20. Meyer A, Van de Peer Y. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 2005; 27:937-45.

21. Powers TP, Amemiya CT. Evolutionary plasticity of vertebrate Hox genes. Current Genomics 2004; 5:459-72.

22. Mahadevan NR, Horton AC, Gibson-Brown JJ. Developmental expression of the amphioxus Tbx1/ 10 gene illuminates the evolution of vertebrate branchial arches and sclerotome. Dev Genes Evol 2004; 214:559-66.

23. Cameron RA, Rowen L, Nesbitt R et al. Unusual gene order and organization of the sea urchin hox cluster. J Exp Zool B Mol Dev Evol 2006; 306:45-58.

24. Santini S, Bernardi G. Organization and base composition of tilapia Hox genes: implications for the evolution of Hox clusters in fish. Gene 2005; 346:51-61.

25. Lee AP, Koh EG, Tay A et al. Highly conserved syntenic blocks at the vertebrate Hox loci and conserved regulatory elements within and outside Hox gene clusters. Proc Natl Acad Sci U S A 2006; 103:6994-99.

26. Jaillon O, Aury JM, Brunet F et al. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 2004; 431:946-57.

27. Amores A, Suzuki T, Yan YL et al. Developmental roles of pufferfish Hox clusters and genome evolution in ray-fin fish. Genome Res 2004; 14:1-10.

28. Taylor JS, Van de Peer Y, Meyer A. Genome duplication, divergent resolution and speciation. Trends Genet 2001; 17:299-301.

29. Wagner A. Birth and death of duplicated genes in completely sequenced eukaryotes. Trends Genet 2001; 17:237-39.

30. Postlethwait J, Amores A, Force A et al. The zebrafish genome. Methods Cell Biol 1999; 60:149-63.

31. Naruse K, Fukamachi S, Mitani H et al. A detailed linkage map of medaka, Oryzias latipes: comparative genomics and genome evolution. Genetics 2000; 154:1773-84.

32. Sayers EW, Barrett T, Benson DA et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2012; 40:D13-25.

33. Sayers EW, Barrett T, Benson DA et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2011; 39:D38-51.

34. Guex N, Peitsch M. Swiss-PdbViewer: a fast and easy-to-use PDB viewer for Macintosh and PC. Protein Data Bank Quaterly Newsletter 1996; 77.

35. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997; 18:2714-23.

36. Laskowski RA, Rullmannn JA, MacArthur MW et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996; 8:477-86.

37. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins 1993; 17:355-62.

38. Pettersen EF, Goddard TD, Huang CC et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004; 25:1605-12.

39. Pintilie GD, Zhang J, Goddard TD et al. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J Struct Biol 2010; 170:427-38.

40. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. J Struct Biol 2007; 157:281-87.

41. Alonso CR. Hox proteins: sculpting body parts by activating localized cell death. Curr Biol 2002; 12:R776-78.

42. Schofield PN. Patterns, puzzles and paradigms: the riddle of the homeobox. Trends in Neurosciences 1987; 10:3-6.