Structural and Sequence Conservation of the CtBP Corepressor C-Terminal Domain Across Vertebrate Species

Madeline Niblock[1], Kalynn Bird[1] Dhruva Kadiyala[2], Ana-Maria Raicu[3], David N. Arnosti[2]

[1]Michigan State University Lyman Briggs and Honors Colleges, Department of Biochemistry and Molecular Biology [2]Michigan State University Department of Biochemistry and Molecular Biology

[3]Michigan State University Graduate Program in Cell and Molecular Biology

Abstract

The C-terminal Binding Protein (CtBP) is a transcriptional corepressor that regulates gene expression and functions as a tumor suppressor. The CtBP catalytic core resembles an NAD(H)-dependent dehydrogenase, and binding to the NAD(H) cofactor enables formation of dimers and tetramers. The less conserved C-terminal domain (CTD) has not been structurally characterized and a functional role remains elusive. Some organisms, like Drosophila, have a single CtBP gene with multiple splice isoforms that encode variant "short" and "long" CTDs. Vertebrates encode two paralogous genes, CtBP1 and CtBP2. To uncover the relevance and conservation of the CTD, we performed an evolutionary comparison of protein sequences from vertebrates. We find that the CtBP1 primary structure is highly conserved, with only subtle differences observed in the CTD when comparing mammals to amphibians and fish. Conversely, the CtBP2 CTD is less conserved across the vertebrate phylogeny. Still, primary structure similarities can be identified between the CTD of CtBP1 and CtBP2. Secondary structure predictions using PSIPRED software indicate that in both CtBP1 and CtBP2, most of the CTD is predicted to be disordered, but an alpha-helix may form. CtBP2 is predicted to have more variable disordered regions varying in location and length, depending on the organism analyzed. The high level of primary and secondary structure conservation of the CtBP CTD among vertebrates suggests that this poorly characterized domain is critical for function. Extending our evolutionary approach to other metazoans can reveal possible regulatory and functional significance of the CTD of this metazoan transcriptional corepressor.

Methods

CtBP protein sequences from mammalia, sauria, amphibia, and actinopterygii (fish) were downloaded from NCBI. Here, we specifically chose the following species because they are representatives of these lineages of interest: *Homo sapiens* (Humans), *Delphinapterus leucas* (Beluga Whale), *Thamnophis elegans* (Western Terrestrial Gartner Snake), *Numida meleagris* (Helmeted Guineafowl), *Nanorana parkeri* (High Himalaya Frog), *Geotrypetes seraphini* (Gaboon caecilian), *Megalops cyprinoides* (Indo-Pacific Tarpon), and *Danio rerio* (Zebrafish). When protein sequences were missing in NCBI, cDNA sequences were downloaded and translated. To create the interspecies alignments, we used the Clustal Omega Multiple Sequence Alignment tool. The phylogenetic tree was generated with TimeTree (timetree.org). Secondary structure predictions of CtBP CTDs were made using PSIPRED Workbench V3.2 from UCL Department of Computer Science Bioinformatics Group.

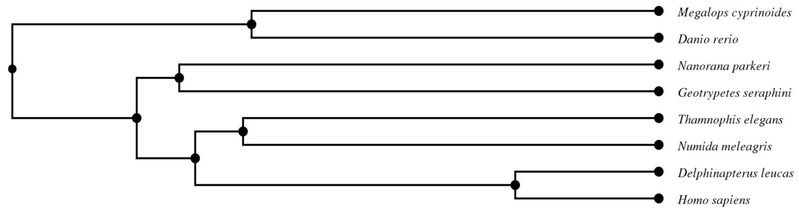

Figure 1. Phylogenetic Tree of vertebrate species of interest. 2 representative species were included from mammalia, sauria, amphibia, and actinopterygii

CtBP1

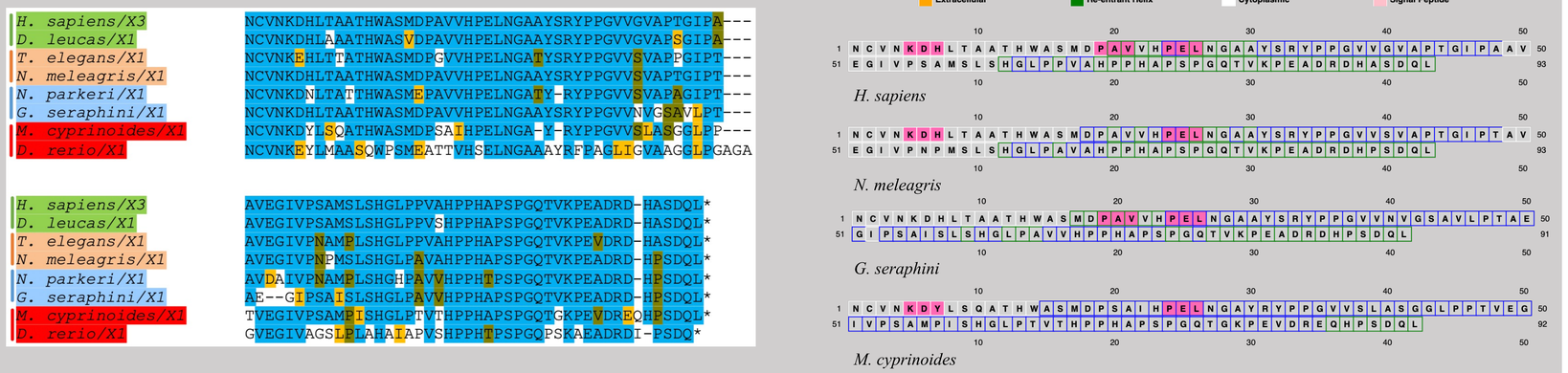

Figure 2. Conservation of CtBP1 C-terminal domain (A) Vertical lines and highlighting on the left indicate four groups: green (mammalia), orange (sauria), blue (amphibia), and red (actinopterygii). Amino acid residues were highlighted to indicate absolute conservation (blue), chemical conservation (gold), and conservation of an alternative amino acid (brown). A high level of conservation is observed with a slight decline in conservation of distant relatives of *H. sapiens*, such as in amphibians and actinopterygii. **(B)** Secondary structure predictions of the CTD of CtBP1 in *H. sapiens*, *N. meleagris*, *G. seraphini*, and *M. cyprinoides*. Conserved secondary structures include KDH, PAV, and PEL alpha helices. The CTD is predicted to be largely disordered near the middle, and disordered protein binding near the end of these sequences.

CtBP2

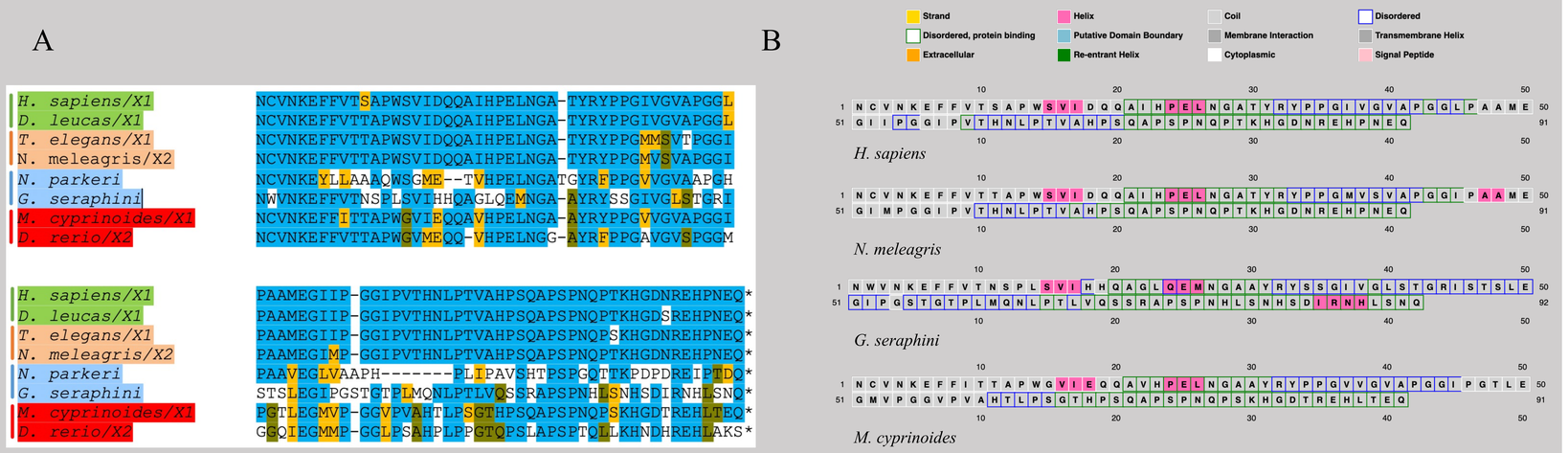

Figure 3. Conservation of CtBP2 C-terminal domain (A) Conservation of CtBP2 C-Terminal Domain. Color and vertical lines used correspond to that described in figure 2A. A high level of conservation of the CtBP2 CTD is visualized that is not as striking as that of CtBP1. Although there is a high amount of primary conservation between mammals and sauria, a divergence is seen compared to amphibia and actinopterygii. Additionally, there is more divergence in CtBP2 as compared to CtBP1 within amphibian and actinopterygii species chosen. **(B)** Secondary structure predictions of the CTD of CtBP2 in *H. sapiens*, *N. meleagris*, *G. seraphini*, and *M. cyprinoides*. Conserved secondary structures include SVI and PEL alpha helices. As with CtBP1, the middle of the sequences are disordered, and the ends are disordered protein binding.

Conclusions

Based on our peptide alignments and secondary structure predictions, we find that there is strong conservation of both the CtBP1 and CtBP2 CTDs within the vertebrate. Additionally, there are similarities in both sequence and predicted structure when comparing CtBP1 to 2. However, we do notice CtBP2 seems to show more variability than CtBP1. This may be due to CtBP2 having alternative functions or substrate interactions in different species. Besides vertebrates, we have compared this data to more distant relatives such as invertebrates, and we still see a high level of sequence and structural conservation of specific motifs. All of this leads to the notion that the CTD of CtBP is vital for function.

Acknowledgements

We would like to thank the MSU Department of Biochemistry and Molecular Biology for their support on this project. This research was funded by the NIH grant GM124137 to Dr. David N. Arnosti, the Biochemistry and Molecular Biology Undergraduate Research Fellowship to Madeline Niblock, and the Moore-Billman Scholarship to Kalynn Bird. We would like to thank the other members in the Arnosti Lab for their support.