

**This document provides details of Appendix files that are uploaded separately.**

Appendix files in the form of tables and figures pertaining to Chapters 2, 3, 5 and 6 are provided along with the main thesis. The name of the file starts with 'Appendix' followed by the nature of file (i.e., table or figure) and the chapter number and are separated by '\_' symbol. Chapter 6 also comprises four movie files.

File names and associated legends are provided below,

## **Chapter 2**

**Appendix table 2.1. Statistics on sequence conservation of orthologous proteins in Afrotheria.** Appendix table 2.1A in Sheet 1 provides the total number of most, moderate, and least conserved orthologous proteins in Afrotherian species. Appendix table 2.1B-G in Sheets 2-7 is the list of Afrotherian proteins in each species and their orthologues.

**Appendix table 2.2. The potential function of Afrotheria-specific proteins.** Appendix table 2.2A lists molecular features and inferred function of Afrotheria-specific uncharacterized proteins. Appendix 2.2B in Sheet 2 contains details on prediction of functions for species-unique proteins of Afrotheria.

**Appendix table 2.3. Distribution of functional domains in Afrotheria and reference species.** Appendix table 2.3A lists the normalized frequency of distribution of various functional domains in each species. Appendix tables 2.3B-D in Sheets 2-4 list proteins showing domain complexity index of 2, 3 and 4 or more.

**Appendix table 2.4. Functional domain architectures in Afrotheria.** Appendix table 2.4A lists unique functional domain architectures in Afrotheria. Appendix table 2.4B in Sheet 2 provides the list of domain repeats showing unique domain combinations in Afrotheria.

**Appendix table 2.5. Structural domain distributions in Afrotheria.** Appendix table 2.5A provides the list of structural domain superfamilies and their numbers in various Afrotherian proteomes. Appendix table 2.5B in Sheet 2 is a list of the most commonly occurring domains and the number of their species- unique architectures in various Afrotherian proteomes. Appendix table 2.5C in Sheet 3 lists species-unique architectures.

**Appendix table 2.6. Domains that occur in high numbers in Afrotheria.** The table shows the distribution of Ribosomal domains in Afrotheria and other daily torpor/hibernating species.

### Chapter 3

**Appendix Table 3.1.** PDB codes of 66 targets and corresponding templates identified for each of them. Columns D-G contain template PDB codes along with chain ID, percentage sequence identity and SCOPe family ID as PDB code\_chain ID (percentage sequence identity), SCOPe ID. Columns I-L provide details of template sequence length and its difference with the target sequence in bracket.

### Chapter 5

**Appendix table 5.1. The table provides details of the dataset used for sequence conservation analysis of SF3b proteins.** Appendix table 5.1A lists the species in which SF3b homologs are identified. Presence (Y) and absence (empty) of homologs of SF3b proteins in each species are indicated along with the status of species in which all SF3b homologs (SF3b1, SF3b2, SF3b3, SF3b4, SF3b14b and SF3b5) were identified. In addition, the status of availability of genome assembly for a species is mentioned. Appendix table 5.1B in Sheet 2 lists the taxonomy ID of species present in the dataset and customized name to indicate its eukaryotic lineage.

**Appendix table 5.2. A survey of C-terminal motifs in homologs of SF3b4.** Appendix table 5.2A in Sheet 1 provides 3 character codes used to indicate eukaryotic supergroup in the customized name for species. Appendix tables 5.2B-D in Sheets 2, 3 and 4 have the list of species in which SF3b4 homologs possess proline-rich motifs, disordered regions and no functional motifs at the C-terminus, respectively.

**Appendix table 5.3. List of species that lack the SF3b6 protein.**

**Appendix figures 5.1.** Multiple sequence alignments of SF3b proteins obtained from representative sequences of homologs clustered at 60% sequence identity. Alignments correspond to SF3b1 (5.1A), SF3b2 (5.1B), SF3b3 (5.1C), SF3b4 (5.1D), SF3b14b (5.1E), SF3b5 (5.1F) and SF3b6 (5.1G) are appended sequentially. Taxa-specific E/D rich regions are labelled in the SF3b3 alignment.

**Appendix figures 5.2.** Shown are phylogenetic trees generated for each protein of the SF3b complex. Phylogenetic tree corresponding to SF3b1, SF3b2, SF3b3, SF3b4, SF3b14b, SF3b5 and SF3b6 are appended sequentially along with the label of protein name. Branches are colored based on bootstrap support value. Nodes in the tree belonging to different kingdoms are highlighted with different colors in the background.

## Chapter 6

**Appendix table 6.1. Information on the sequence similarity and structural difference between orthologous yeast and human proteins of SF3b complex.**

### List of movies associated with Chapter 6 and their corresponding legends

**Movie 6.1** Top two lowest frequency global modes representing 'open-to-close' and 'twist' motions in SF3b1<sup>SF3b6-bound</sup> (right) are shown along with the top two lowest

frequency global modes experienced by SF3b1<sup>iso</sup> (left). Color scale blue to red indicates the magnitude scale of residue motions from low to high. It can be observed that SF3b1<sup>iso</sup> shows large scale movements concentrated only at the N-terminal region and have weak collective residue motions with the rest of the protein. In the case of SF3b1<sup>SF3b6-bound</sup>, large scale movements are exhibited by the N-terminal region and N- and C-terminal ends of superhelical HEAT repeats with high collectivity in residue motions.

**Movie 6.2** Shown is the 6<sup>th</sup> lowest frequency global mode of human SF3b1<sup>iso</sup> exhibiting 'twist' motion. It has the highest collectivity in residue motions among the top 20 lowest frequency global modes of SF3b1<sup>iso</sup> (Figure 1B, left panel) and similar to the second lowest frequency global mode of SF3b1<sup>SF3b6-bound</sup>. The color scale from blue to red represents the magnitude scale of residue motions from low to high.

**Movie 6.3** The top two lowest frequency modes of yeast Hsh155 representing 'open-to-close' and 'twist' motions. These motions are similar to the top two lowest frequency motions experienced by human SF3b1 in SF3b6 bound form (SF3b1<sup>SF3b6-bound</sup>). Color scale green to red represents the magnitude scale of residue motions from low to high.

**Movie 6.4** Morph of human SF3b3 and yeast Rse1 structure (of B<sup>act</sup> assembly) to showcase structural variability observed between them. 15aa long helix insertion is highlighted in yellow. From the movie, it can be inferred that all three  $\beta$ -propeller domains possess conformational differences between SF3b3 and Rse1.