# Identification and *in silico* characterisation of differentially expressed fluoride induced proteins in the muscle tissue of Indian Major Carp – *Gibelion catla* (Hamilton, 1822)

**Baludu Ravi[1]*, Chintala Ramakrishna[1], Gudivada Mani[2] and Mokka Jagannadharao[3]**
1. Dept. of Environmental Sciences, Gitam institute of Sciences, Visakhapatnam, INDIA
2. Dept. of Zoology, College of Science and Technology, Andhra University, Visakhapatnam, INDIA
3. Dept. of Geology, College of Science and Technology, Andhra University, Visakhapatnam, INDIA
*ravibaludu74@gmail.com

## *Abstract*

*The detrimental effects of fluoride on aquatic life and humans have made the toxicity of this substance a worldwide concern. Fluoride impacts fish survival, growth and reproduction by affecting the levels of biomolecules, as well as enzyme inhibition, collagen degradation, gastrointestinal injury and immune system disruption. Thus, it is essential to comprehend the molecular mechanisms influenced by fluoride to devise strategies for the protection and preservation of aquatic systems. The current investigation focuses on identifying and characterising differentially expressed proteins in Gibelion catla with respect to fluoride toxicity. Protein measurement indicated that the protein concentrations in the muscular tissues of control and NaF-treated fish were 17.8±1.4 and 12.3±2.2 mg/g.*

*Additionally, comparative proteome analysis through SDS-PAGE demonstrated that two peptide bands with molecular weights of 31 and 18 kDa were differentially expressed in fish exposed to NaF. These protein bands were identified by PMF as caspase-3 and Cu/Zn SOD. The physico-chemical characterisation indicated that the caspase-3 and Cu/Zn SOD proteins exhibited acidic characteristics, with pI values of 6.07 and 5.69. The secondary structure revealed that caspase-3 and Cu/Zn SOD exhibit a greater fraction of random coils, with respective percentages of 53.11 and 55.19%. Furthermore, the homology modelling indicated that the C5jftB and d1q0ea are the most suitable templates for caspase-3 and Cu/Zn SOD. The present research demonstrated that the predicted 3D models of caspase-3 and Cu/Zn SOD were reliable and coherent.*

**Keywords:** Fluoride, *Gibelion catla*, Proteome, SDS-PAGE, Caspase-3, Cu/Zn SOD.

## Introduction

Aquatic fluoride contamination has become a global concern due to its detrimental effects on both aquatic organisms and humans[40]. Fluorides are naturally occurring substances that are part of the halogen group of minerals. Fluoride is mostly found in groundwater, where it is formed by the solvent action of water on the rocks and dirt of the earth's crust. The elevated fluoride levels in groundwater and surface water in numerous regions globally are a significant problem. Fluoride-bearing rocks such as cryolite, fluorite, fluorspar, fluorapatite and hydroxylapatite are the primary sources of fluoride in groundwater[27].

According to WHO[41] statistics, fluoride enters the piscivorine population through the food chain. The gills and skin of fish allow fluoride to enter its body, where it typically accumulates in the muscles, liver, gut, skeletal structure and gills[37]. Fluoride primarily impacts the levels of protein, fat, glucose, cholesterol and glycogen, which are essential to fish survival, growth and reproduction[22]. Furthermore, fluoride toxicity results in enzyme inhibition, collagen degradation, gastrointestinal injury and immune system alteration[26].

Bajpai and Tripathi[3] assert that lipids and proteins function as growth bioindicators for fluoride pollution in *Heteropneustis fossilis*, with a prolonged fluoride presence leading to a reduction of these biomolecules in body tissues, ultimately impairing the fish's growth and development. In addition, exposure to NaF in experimental fish resulted in behavioural anomalies including alterations in body position, habits, food sensitivity, operculum opening rate and swimming movements[28]. However, the buildup and elevated release of mucus in fish exposed to fluoride could constitute an adaptive and defensive reaction to prevent the toxicant's absorption by the entire body surface of fish[43].

Fish is a significant component of the human diet in numerous regions globally, owing to its high protein content, low saturated fat and adequate omega fatty acids recognised for promoting great health[15]. Over the last two decades, Indian aquaculture has expanded by 6.5 times, with freshwater fish farming constituting over 95% of the overall production. The freshwater aquaculture sector, which represented 34% of inland fisheries in the mid-1980s, is now expanding to over 80%[10].

Hence, the toxicity assessment of fluoride on fish anticipates the health condition of the aquatic environment and serves as an important screening tool for devising protection and conservation strategies for aquatic systems. Therefore, understanding the molecular pathways influenced by

fluoride toxicity is crucial for understanding its impacts and filling the gaps in current studies. The current investigation focuses on identifying and characterising differentially expressed proteins in *Gibelion catla* muscle with respect to fluoride toxicity using *in silico* because of high productivity and economic significance.

One of the rapidly expanding species that has attracted market interest is the *G. catla*, which is a member of the carp family (Cyprinidae) and is often referred to as the Indian carp or *Catla*[18]. It can be found in Bangladesh, Nepal, India, Pakistan, Myanmar, Sri Lanka and China. The species is indigenous to North India and has recently been introduced to Peninsular India. The *G. catla* fish species possesses significant ecological importance by primarily consuming algae and plankton, hence regulating the population density of these minute creatures[12]. By preventing severe eutrophication and improving water quality indirectly, *G. catla* benefits the environment as a whole and boosts biodiversity[14]. *Catla* holds considerable economic and cultural significance as well as an exceptional option for commercial fishing in South Asia due to its substantial size and superior flesh quality[16].

Furthermore, *G. catla* is a rich source of beneficial lipids including polyunsaturated omega fatty acids like docosahexaenoic acid, arachidonic acid and eicosapentaenoic acid, which the human body cannot synthesise and requires dietary supplementation. Moreover, the literature search indicated a deficiency of complete evidence concerning the impact of fluoride poisoning on proteome expression. Hence, this study concentrated on the identification and *in silico* characterisation of differentially expressed proteins in the muscles of *G. catla* in relation to fluoride toxicity.

## Material and Methods

**Experimental design:** Healthy *G. catla* samples (weight approx. 150-250 g, length 20-25 cm) were sourced from local fish farming ponds of the Visakhapatnam district and maintained under stable laboratory conditions for 30 days. The study comprised of two groups: Group I, consisting of control fish and Group II, consisting of NaF-treated fish. The control group (Group I) consists of fish that are not exposed to NaF-toxicity, while group II is subjected to a NaF concentration of 35 mg/L. The investigation was conducted using glass aquarium measuring 60x40x45 cm for a period of 30 days during which *G. catla* were fed three times daily with pellet fish feed (Growell-Fishmax).

The water temperature was adjusted to align with the pond water temperature, ranging from 23-25ºC. The photoperiod was adjusted to 12 hours of light and 12 hours of darkness. Aerators were affixed to tubs to enhance aeration. Daily, excess waste was removed and 50% of the water in the aquarium was replaced with dechlorinated fresh water. Following a 30-day exposure, fish were dissected and the inedible organs such as offal, bones and scales were promptly removed to prevent decomposition. The excised muscle tissue was rinsed and preserved at -20ºC for subsequent analyses.

**Protein extraction and estimation:** The protein content of *G. catla* muscle tissue was extracted using PVP according to Ferreria et al[11] method. In order to extract protein, the muscle tissue was homogenised at 4ºC in 50 mM sodium phosphate buffer containing 10% PVP. The homogenate was subjected to centrifugation for 20 minutes at 4ºC and 14000 rpm, after which the supernatant was collected. Furthermore, the protein concentration in the supernatant was determined using the Lowry et al[24] method.

**Proteome profiling by SDS-PAGE and elution of peptide:** Proteins were separated based on molecular weight using SDS-PAGE in accordance with Sambrook and Russell[31] methodology. The crude protein sample was mixed in the sample loading buffer at a 1:3 v/v ratio and it was heated to 95ºC for 10 minutes and subsequently cooled. The sample was then subjected to SDS-PAGE alongside a parallel run of protein markers with molecular weights ranging from 250-10 KDa. Laemmli[23] method was utilised to prepare a 15% resolving gel and a 5% stacking gel. Using a vertical electrophoretic unit and running buffer that contained 25 mM Tris-base, 192 mM glycine and 3 mM SDS, the electrophoresis was initially carried out at 50 V for one hour and then it was run at 100 V.

Followed by electrophoresis, the gel was stained for overnight using 0.5% Coomassie Brilliant Blue R-250 and then destained until the background was clear. The gel was examined to identify the protein bands that were differentially expressed between the control and NaF-treated fish. After SDS-PAGE, the differentially expressed peptide gel bands from the NaF-treated fish were extracted from the gel by suspending them in elution buffer and precisely crushing the mixture. Then the solution was centrifuged at 10000 rpm for 10 minutes at 4ºC and the supernatant containing the eluted peptide fraction was collected.

**Protein identification by MALDI-TOF/MS:** For identification and molecular weight determination, the eluted peptide fraction was analysed through peptide mass fingerprinting using MALDI-TOF/MS. For PMF, 0.5 µL of the eluted peptide was applied to a matrix that contained saturated α-cyano-4-hydroxycinnamic acid that had been made with 50% acetonitrile and 5% trifluoroacetic acid. The mass spectrometric data of the digested protein sample was obtained using the ABI 4800 MALDI-TOF (Applied Biosystems, Foster City, CA). The spectral data were obtained in reflector mode across a mass range of 600-9000 Da. The generated mass spectrum was subjected to sequence database searches utilising MASCOT software. The MASCOT software analysis generated a score that indicated the likelihood of a true positive identification, with a minimum value of 50.

***In silico* characterisation:** The identified peptides were characterised by predicting their physico-chemical characteristics, secondary structure, tertiary structure, gene ontologies for their biological functions and PPI networks. The FASTA sequences of the identified proteins were obtained from UniProt, a publicly available protein database at www.uniprot.org[2]. The ProtParam server from Expasy was utilised to ascertain the physicochemical properties. The secondary structural components of the identified peptide were predicted utilising the SOPMA server. The 3D model was generated using the Phyre2 server[20]. Subsequently, the 3D model quality was assessed using the QMEAN scoring function and Ramachandran plot analysis using the PROCHECK server. Furthermore, the PPI networks were established using STRING analysis via PPI pairs with protein interaction values exceeding 0.4.

## Results and Discussion

**Estimation of protein:** Quantifying proteins in any organism is crucial for evaluating its physiological and metabolic condition, as protein levels directly correlate with health status. The protein concentrations in the muscle tissue of control fish and NaF-treated fish were measured as 17.8±1.4 and 12.3±2.2 mg/g respectively. These findings demonstrated a significant decrease in the protein content of fish exposed to fluoride. Kale and Muley[19] have demonstrated that fluoride exposure significantly reduced the total protein content in *Labeo rohita* muscle, liver, gills and kidney. Chitra et al[8] and Kumar et al[22] stated that fluoride impacts specific biomolecules and enzymes in various tissues of freshwater fish. Fluoride disrupts multiple enzymes, thereby inhibiting their function. Therefore, disruptions to several vital metabolic functions including glycolysis, transcription and others are necessary for maintaining fish physiology[4].

**Protein separation by SDS-PAGE:** The analysis of the muscle proteome in both control and NaF-exposed *G. catla* was conducted utilising SDS polyacrylamide gel electrophoresis, revealing that the molecular weights of the expressed proteome span from 10-200 KDa. Moreover, the majority of the protein bands exhibit high and medium molecular weights. The current study on comparative muscle proteome analysis showed that most of the protein bands were the same in both treatment groups (Figure 1). Nevertheless, two protein bands were additionally found in fish exposed to NaF, suggesting that these proteins were expressed at significantly high levels, with molecular weights recorded at 31 and 18 KDa. The protein bands derived from the NaF-exposed fish were chosen for MALDI-TOF analysis to facilitate the identification of the proteins.

Proteins serve as the prime regulators of metabolic processes, facilitating growth and development and their synthesis undergoes significant alterations in response to environmental stress. The proteins that exhibited differential expression were associated with various cellular and physiological processes, encompassing osmoregulation, energy and carbohydrate metabolism, redox homeostasis and ion balance. Aziz et al[1] discovered that fluoride elevated the levels of alkaline phosphatase (ALP), alanine aminotransferase (ALT) and aspartate aminotransferase (AST) in the gills of freshwater fish.

Consequently, the elevated levels of these enzymes can be attributed to disruptions in carbohydrate and protein metabolism. Yan et al[42] reports indicated that fluoride induces apoptosis through the elevation of oxidative stress-induced lipid peroxidation, which leads to mitochondrial dysfunction and subsequently activates caspase-9 and caspase-3.

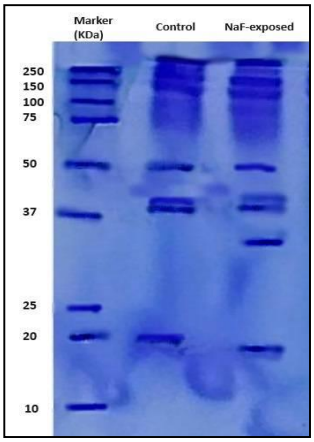


**Figure 1: Protein profiling of *Gibelion catla* muscle proteome under control and NaF-exposed conditions by SDS-polyacrylamide gel.**

**Protein identification:** Peptide mass fingerprinting represents a systematic approach for the identification of unknown proteins through the analysis of their peptide masses. This technique relies extensively on the availability of high-quality protein databases[25]. In this research, the PMF results showed that the protein bands at 31 and 18 KDa matched the caspase-3 protein and the Cu/Zn superoxide dismutase respectively. The mass spectrum data also show that *G. catla* caspase-3 protein and Cu/Zn superoxide dismutase are most similar to their corresponding proteins in *Cyprinus caprio* (Accession number XP_018965718.1) and *Channa argus* (Accession number AVX28209.1), with predicted identity scores of 76.3 and 83.4% respectively. Figures 2 and 3 present the mass spectra for the protein bands at 31 and 18 KDa respectively. Table 1 presents a summary of the protein identification results from peptide mass fingerprinting.

***In silico* characterisation**

**Physicochemical characterisation:** Caspase-3 and Cu/Zn SOD proteins consist of 271 and 154 amino acids respectively, with estimated molecular weights of 30400.43 and 16091.9 Da. Table 2 presents the physicochemical properties of caspase-3 and Cu/Zn SOD. The isoelectric points (pI) for caspase-3 and Cu/Zn SOD peptides were found to be 6.07 and 5.69 respectively. The pI represents the precise value at which a molecule achieves electrical

neutrality, which is characterised by an equal distribution of negative and positive charges. The protein's immobility at the pI makes it a useful component in the formulation of buffer system for isoelectric focus separation.

The extinction coefficients of the caspase-3 and Cu/Zn SOD are 29380 $M^{-1}cm^{-1}$ and 4595 $M^{-1}cm^{-1}$ when all pairs of cysteine residues are converted into cystines. Furthermore, upon reduction of all cystine residues, the extinction coefficients of caspase-3 and Cu/Zn SOD were 28880 $M^{-1}cm^{-1}$ and 4470 $M^{-1}cm^{-1}$ respectively. The highest extinction coefficient represents an elevated concentration of Trp and Tyr. Caspase-3 and Cu/Zn SOD were shown to have instability indices of 19.88 and 22.62 respectively. The protein instability index indicates the stability of a protein under laboratory conditions.

Guruprasad et al[17] stated that a protein is deemed stable if its instability index is less than 40 and that a number greater than 40 suggests possible protein instability. Zaccaria et al[44] demonstrated that small elements located near N-terminus of a protein affect its stability, thereby influencing its longevity. Rogers et al[30] demonstrated that proteins with a half-life of less than 5 hours exhibited an instability index exceeding 40, while those with a half-life greater than 16 hours had an instability index below 40.

**Table 1**
**Identification result of the differentially expressed proteins by peptide mass fingerprinting.**

| Source Organism | Protein Gel Band | Max. homology (Protein name) | Best match organism | Expt/Theor. Mw (KD) | Score (Ms/MS) | Accession No. |
|---|---|---|---|---|---|---|
| *G. catla* | 31 KDa | Caspase-3 | *Cyprinus caprio* | 31/31 | 76.3 | XP_018965718.1 |
| *G. catla* | 18 KDa | Cu/Zn superoxide dismutase | *Channa argus* | 16/18 | 83.4 | AVX28209.1 |

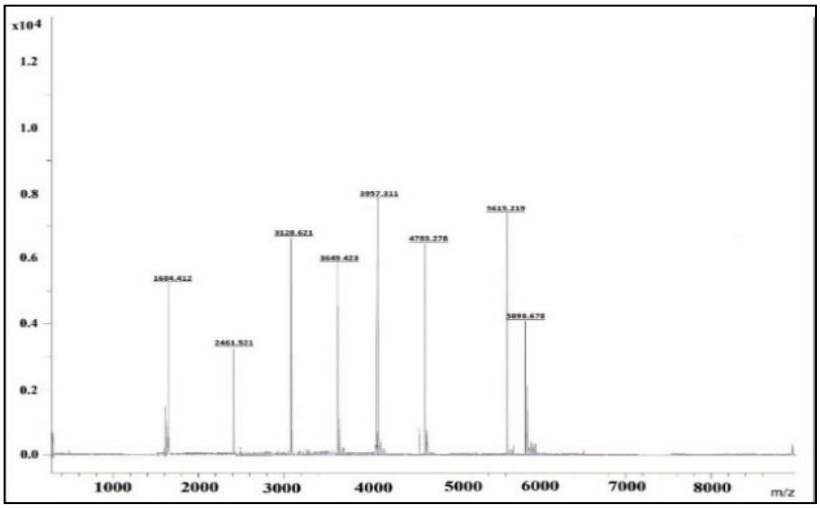


**Figure 2: MALDI-TOF/MS spectrum of 31 KDa protein**

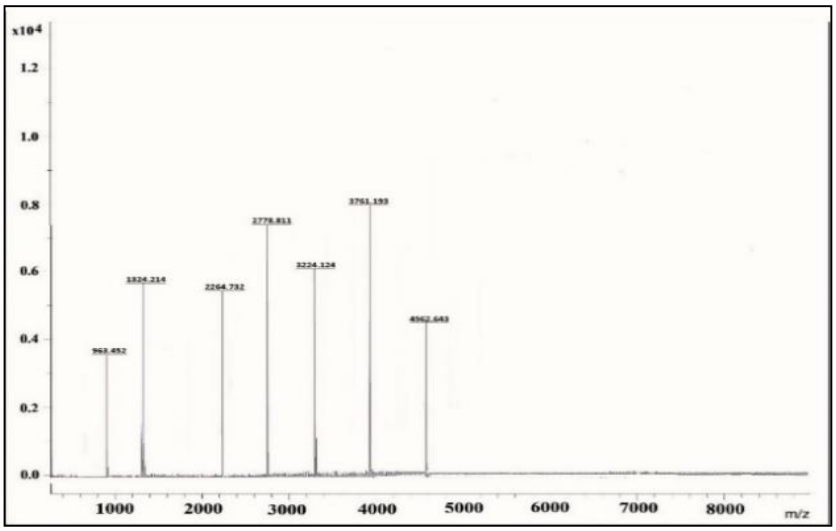


**Figure 3: MALDI-TOF/MS spectrum of 18 KDa protein**

Caspase-3 and Cu/Zn SOD exhibit a half-life of approximately 30 hours in mammalian reticulocytes under invitro conditions while their half-lives exceed 20 hours in yeast and 10 hours in *E. coli*. The protein sequences of caspase-3 and Cu/Zn SOD were found to have aliphatic indices of 74.62 and 76.56 respectively. The aliphatic index of these proteins reflects their stability across different temperatures. The aliphatic index measures the proportion of a protein's volume that is occupied by aliphatic side chains, particularly those derived from the amino acids such as alanine, valine, isoleucine and leucine. It is considered a valuable factor in improving the thermal stability of globular proteins.

Proteins with a high aliphatic index are anticipated to demonstrate stability over a wide range of temperatures. The diminished thermal stability over a wide range of temperatures suggests a structure characterised by increased flexibility. The GRAVY indices for caspase-3 and Cu/Zn SOD were calculated to be -0.279 and -0.392 respectively. The negative GRAVY suggests that proteins exhibit a polar and hydrophilic nature, resulting in enhanced interactions with water. The results indicate that ionisable amino acids located on the protein surface, accessible to water, predominantly affect the pI of proteins.

**Table 2**
**Physicochemical characteristics of fish Caspase-3 and Cu/Zn SOD**

| Parameter | | Caspase-3 | Cu/Zn SOD |
|---|---|---|---|
| Total no. of amino acids | | 273 | 154 |
| Molecular weight | | 30400.43 Da | 16091.9 Da |
| pI | | 6.07 | 5.69 |
| Positively charged residues | | 27 | 12 |
| Negatively charged residues | | 30 | 18 |
| Extinction coefficient | Cys oxidised | 29380 $M^{-1}cm^{-1}$ | 4595 $M^{-1}cm^{-1}$ |
| | Cys reduced | 28880 $M^{-1}cm^{-1}$ | 4470 $M^{-1}cm^{-1}$ |
| Instability index | | 19.88 | 22.62 |
| Aliphatic index | | 74.62 | 76.56 |
| GRAVY | | -0.279 | -0.392 |
| Half-life | In mammalian reticulocytes | 30 hours | 30 hours |
| | In yeast | >20 hours | >20 hours |
| | In *E. coli* | >10 hours | >10 hours |
| Formula | | $C_{1326}H_{2074}N_{370}O_{412}S_{19}$ | $C_{690}H_{1098}N_{204}O_{228}S_6$ |

**Table 3**
**Amino acid composition of fish Caspase-3 and Cu/Zn SOD**

| Amino acid | Caspase-3 | | Cu/Zn SOD | |
|---|---|---|---|---|
| | No. of Residues | % of Residues | No. of Residues | % of Residues |
| Ala (A) | 20 | 7.3 | 9 | 5.8 |
| Arg (R) | 14 | 5.1 | 3 | 1.9 |
| Asn (N) | 15 | 5.5 | 12 | 7.8 |
| Asp (D) | 20 | 7.3 | 7 | 4.5 |
| Cys (C) | 8 | 2.9 | 3 | 1.9 |
| Gln (Q) | 13 | 4.8 | 4 | 2.6 |
| Glu (E) | 10 | 3.7 | 11 | 7.1 |
| Gly (G) | 19 | 7.0 | 25 | 16.2 |
| His (H) | 6 | 2.2 | 8 | 5.2 |
| Ile (I) | 11 | 4.0 | 11 | 7.1 |
| Leu (L) | 19 | 7.0 | 8 | 5.2 |
| Lys (K) | 13 | 4.8 | 9 | 5.8 |
| Met (M) | 11 | 4.0 | 3 | 1.9 |
| Phe (F) | 11 | 4.0 | 3 | 1.9 |
| Pro (P) | 8 | 2.9 | 5 | 3.2 |
| Ser (S) | 20 | 7.3 | 7 | 4.5 |
| Thr (T) | 18 | 6.6 | 11 | 7.1 |
| Trp (W) | 2 | 0.7 | 0 | 0.0 |
| Tyr (Y) | 12 | 4.4 | 3 | 1.9 |
| Val (V) | 23 | 8.4 | 12 | 7.8 |

**Secondary structure:** The secondary structure of caspase-3 comprises of 27.11% alpha helices, 14.29% extended strands, 5.49% beta turns and 53.11% random coils. The secondary structure of Cu/Zn SOD protein is composed of 6.49% alpha helices, 32.47% extended strands, 5.84% beta turns and 55.19% random coils. Table 4 and figure 4 present the secondary structural components of caspase-3 and Cu/Zn SOD. The secondary structure of a protein elucidates the conformation of individual amino acids, signifying their arrangement within a helix, strand, or coil. The findings demonstrated that random coils represent the predominant secondary structural element.

Analysing a protein secondary structure facilitates the examination of hydrogen bonds within the protein, thereby yielding important insights into its structural and functional efficacy. Buxbaum[7] asserts that random coils play a crucial role in proteins by providing flexibility and facilitating conformational changes. Highly flexible glycine and hydrophobic proline amino acids are responsible for the high coil proportion. Proline uniquely induces bends in polypeptide chains, disrupting ordered secondary structures[38].

Neelamathi et al[29] indicate that the existence of coiled portions signifies a high degree of preservation and strength in protein structure. Hydrophobic residues create strong interactions with the hydrophobic lipid bilayer when they are added to a mixture[36]. Shelar et al[32] established a correlation between the structures of α-helical proteins and their diverse activities. These functions include signal detection, receptor activation, ion and chemical transport across membranes, energy transfer and preservation. Additionally, the elongated conformation of α-helical proteins may contribute to their dynamic behaviour and sliding motion, both of which are essential for their best performances.

**Table 4**
**Secondary structural elements of Caspase-3 and Cu/Zn SOD.**

| Structural elements | Caspase-3 | | Cu/Zn SOD | |
|---|---|---|---|---|
| | No. of residues | % of residues | No. of residues | % of residues |
| Alpha helices (h) | 74 | 27.11 | 10 | 6.49 |
| 310 helix (g) | 0 | 0.00 | 0 | 0.00 |
| Pi helix (i) | 0 | 0.00 | 0 | 0.00 |
| β-bridges (b) | 0 | 0.00 | 0 | 0.00 |
| Extended strand (e) | 39 | 14.29 | 50 | 31.47 |
| β-turn (t) | 15 | 5.49 | 9 | 5.84 |
| Bend region (s) | 0 | 0.00 | 0 | 0.00 |
| Random coil (c) | 145 | 53.11 | 85 | 55.19 |
| Ambiguous states | 0 | 0.00 | 0 | 0.00 |
| Other states | 0 | 0.00 | 0 | 0.00 |

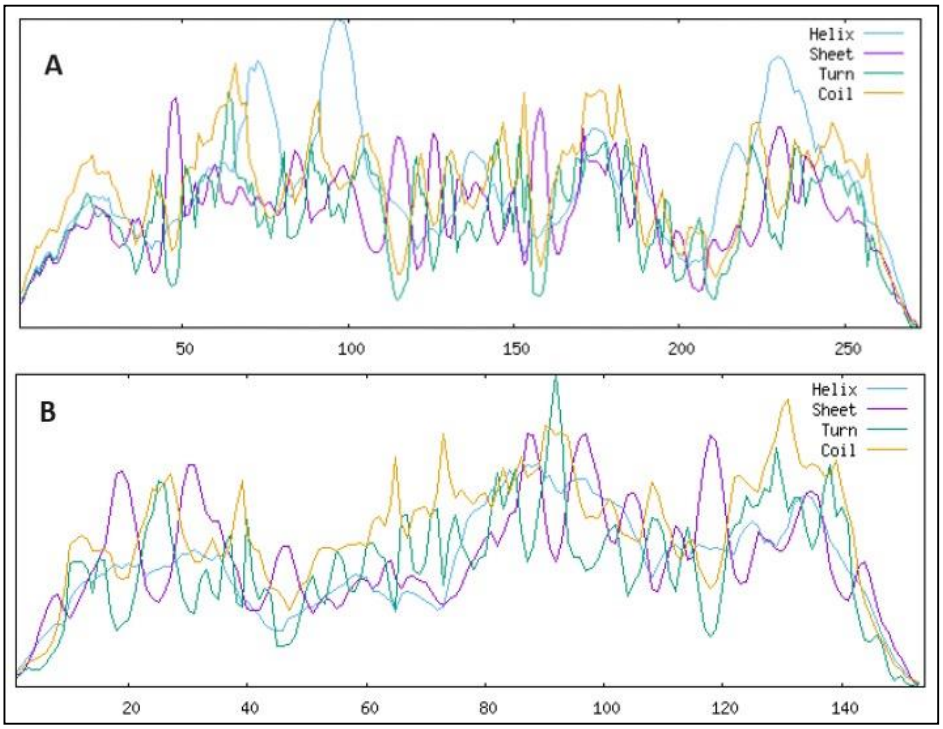


**Figure 4: Spectra of secondary structural elements. A) Caspase-3 B) Cu/Zn SOD.**

**3D structure:** The 3D models of caspase-3 and Cu/Zn SOD from *G. Catla* were predicted using protein IDs C5jftB and d1q0ea as templates, demonstrating identify percentages of 90 and 70% respectively, with a confidence score of 100. Figures 5 and 6 present the predicted 3D representations of caspase-3 and Cu/Zn SOD. The confidence score is used to evaluate the accuracy and reliability of predicted structural 3D models. The assessment involves evaluating the threading alignments and convergence parameters of structure assembly simulations.

Homology modelling is a prevalent technique frequently employed in life science research to create structural models of proteins in the absence of available realistic structures. The 3D structure of proteins offers detailed understanding of their interactions and localisation in a stable conformation and also crucial for their molecular functions. Biasini et al[6] asserted that modelling and assessment methods must account for protein flexibility, given that proteins can exist in structurally distinct functional states and are not static entities. The Phyre2 facilitated the generation of 3D structural models for target protein sequences. This method was subsequently utilised to investigate possible folds through threading, using template sequences from the PDB structure database as references.

**Quality assessment of predicted 3D models:** The caspase-3 and Cu/Zn SOD predicted models have QMEAN values of -0.26 and 0.70 respectively. According to the density map of the QMEAN score, the model's predicted consistency is between 0 and 1. The predicted models QMEAN score were near zero, signifying high model quality. Figures 7 and 8 present the density plots produced by the QMEAN server for the predicted 3D models of caspase-3 and Cu/Zn SOD. The QMEAN server assessed the accuracy of the predicted 3D models. QMEAN is an index that integrates statistical probabilities of average energy and model reliability with predicted structural properties based on the target protein sequence[33]. Benkert et al[5] asserted that the effectiveness of the predicted model is contingent upon the QMEAN score which has been standardised according to the number of interactions.

**Model validation:** The structural integrity of the caspase-3 and Cu/Zn SOD 3D models was evaluated using a Ramachandran plot assessment utilising the PROCHECK server, with a particular emphasis on the geometric features of the backbone conformations. The Ramachandran plot for the caspase-3 3D model indicates that 89.1% of residues are situated in the most favoured region, 10.4% in the additionally allowed region and 0.5% in the generously allowed regions.

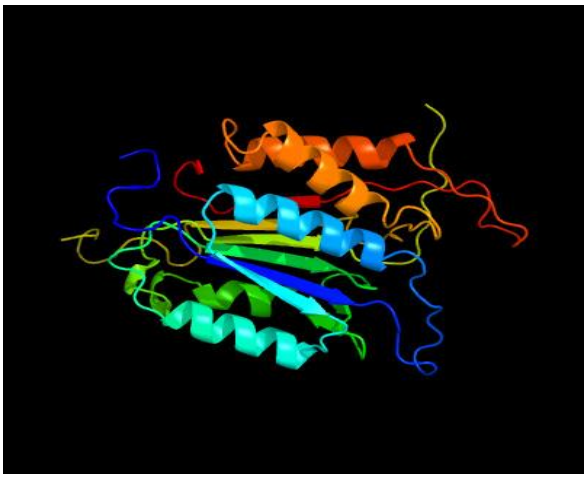

**Figure 5: Predicted 3D model of Caspase-3**

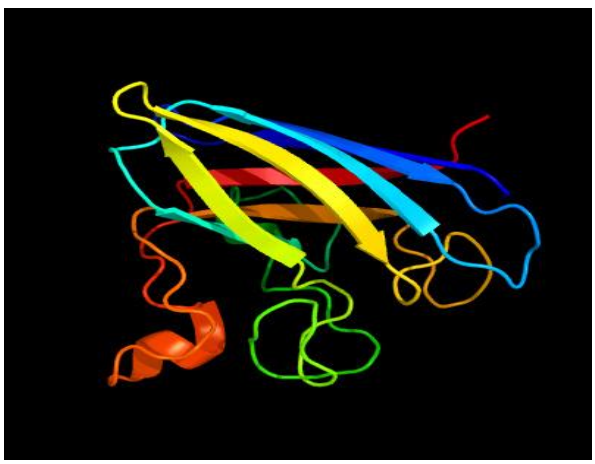

**Figure 6: Predicted 3D model of Cu/Zn SOD**

Figure 9 presents the Ramachandran plot for the predicted 3D model of caspase-3. The Ramachandran plot for the Cu/Zn SOD model indicates that 89.3% of residues are situated in the most favourable region, 9.9% in the additionally allowed region and 0.8% in the generously allowed regions. Figure 10 presents the Ramachandran plot for the predicted 3D model of Cu/Zn SOD. Table 5 displays the Ramachandran plot statistics for the caspase-3 and Cu/Zn SOD models. The results indicate that the majority of amino acids phi-psi conformations align with a right-handed helix. The predicted model demonstrates reliability and stability.

Based on the idea that the native structure of a protein molecule has the lowest free energy of all its potential conformations, several methods are employed. Kiran et al[21] employed PROCHECK to assess the validity of the 3D structure of ATP synthase β and glutamine dependent $NAD^+$ synthetase.

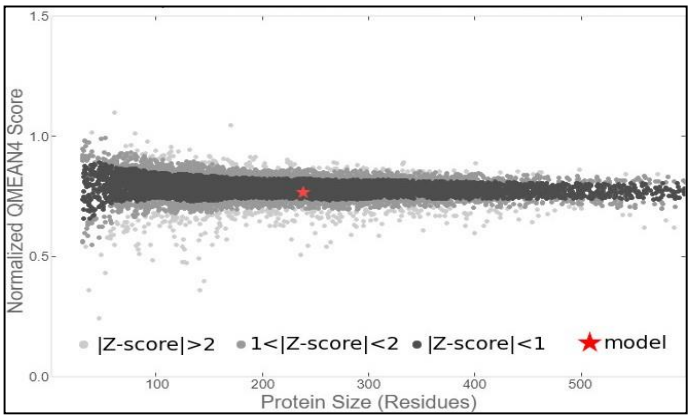


**Figure 7: Plot showing the QMEAN and Z-score of the predicted 3D model of Caspase-3**

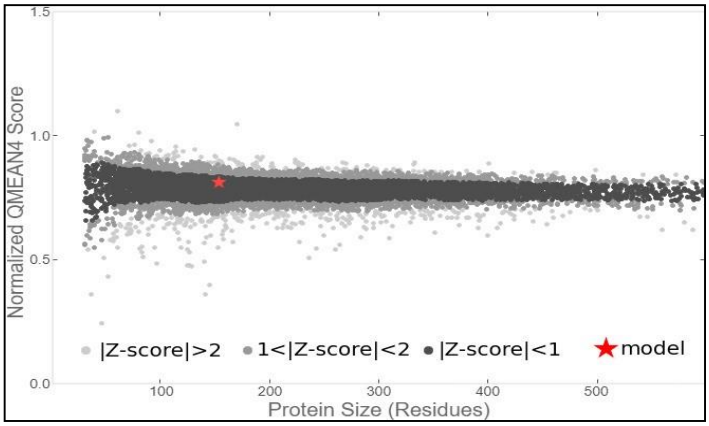


**Figure 8: Plot showing the QMEAN and Z-score of the predicted 3D model of Cu/Zn SOD**

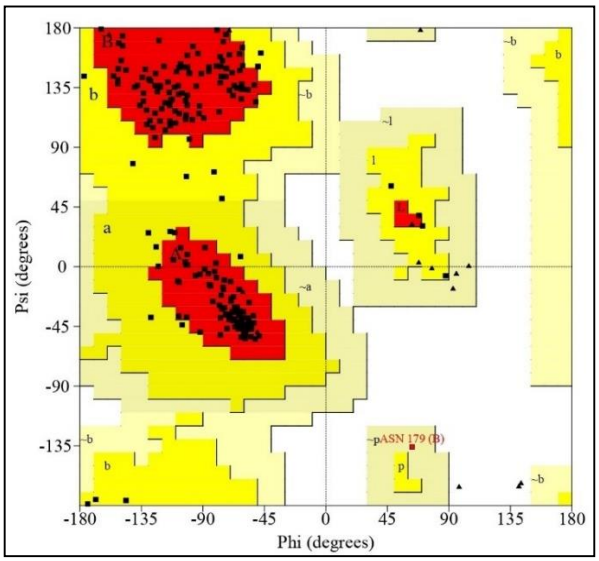


**Figure 9: Ramachandran plot statistics of predicted caspase-3 structural model**

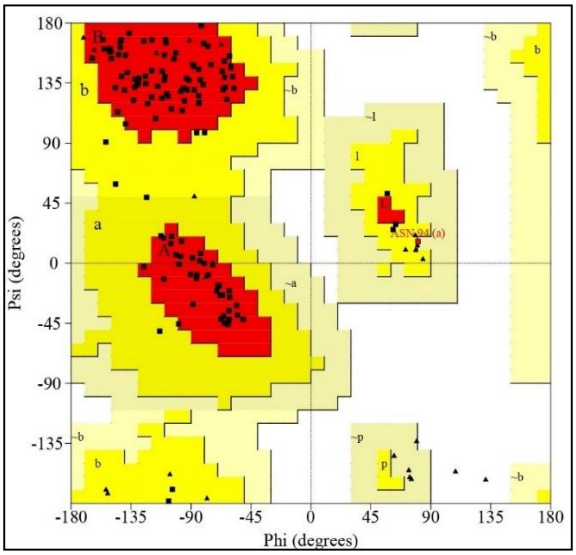


**Figure 10: Ramachandran plot statistics of predicted Cu/Zn SOD structural model**

**Table 5**
**Ramachandran plot statistics of predicted caspase-3 and Cu/Zn SOD models.**

| Residues | Caspase-3 | | Cu/Zn SOD | |
|---|---|---|---|---|
| | No. of Residues | % of Residues | No. of Residues | % of Residues |
| Most favoured regions | 202 | 80.5 | 137 | 95.1 |
| Additionally allowed regions | 30 | 12.0 | 7 | 4.9 |
| Generously allowed regions | 9 | 3.6 | 0 | 0.0 |
| Disallowed regions | 10 | 4.0 | 0 | 0.0 |
| Non-Glycine and Non-Proline | 251 | 100 | 144 | 100 |
| End Residues | 2 | - | 2 | --- |
| Glycine | 8 | - | 2 | --- |
| Proline | 10 | - | 2 | --- |

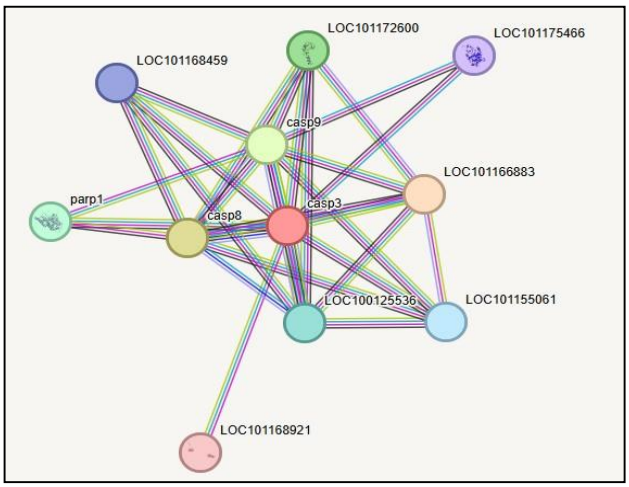


**Figure 11: The PPI network of Caspase-3 and the predicted functional partners**

**PPI network analysis:** The STRING database includes protein-protein interactions (PPIs) obtained from both experimental and computational approaches. It assigns a quality score to every interaction by combining data from many sources including literature and gene expression profiles[35]. The STRING database has comprehensive information on protein biology including data on their structure, gene sequence, homology, co expression and association. Researchers have processed this data using text mining, computational phylogeny, *in vitro*, *in vivo* and *in situ* analysis[34]. Analysing the caspase-3 and Cu/Zn SOD proteins with the STRING database predicted the functional partners of the protein with moderate confidence (score: 0.4). The interaction networks of caspase-3 and Cu/Zn SOD with their predicted functional partners is shown in figures 11 and 12.

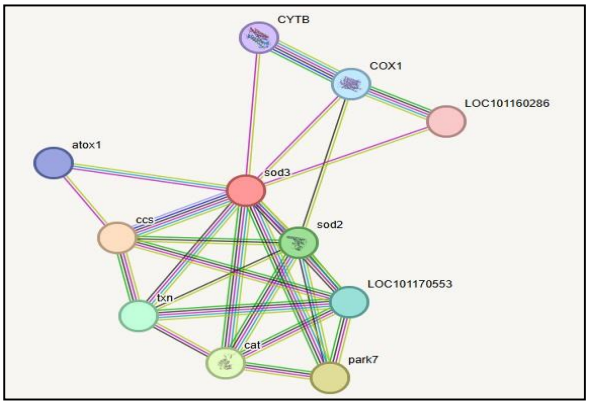


**Figure 12: The PPI network of Cu/Zn SOD and the predicted functional partners**

The functional partners of caspase-3 were found to be LOC10116683 (X-linked inhibitor of apoptosis), Caspase-8, Caspase-9, LOC101172600 (Baculoviral IAP repeat-2), Parp1 (Poly-ADP-ribose polymerase), LOC100125536 (C14 family peptidase), LOC100155061 (Baculoviral IAP repeat-5a), LOC101175466 (Unknown protein) and LOC101168921 (Actin). The statistical parameters of caspase-3 PPI network such as number of nodes, number of edges, average node degree, average local clustering coefficient and PPI enrichment P-value were found to be 11, 30, 5.45, 0.856 and 2.56e-06.

The functional partners of Cu/Zn SOD were found to be CCS (Cu-chaperon for SOD), park7, Cat (Catalase), Sod2 (superoxide dismutase), txn (Thioredoxin), LOC101170553 (NME/NM23 family member), Cox1 (Cytochrome-C oxidase subunit 1), atox1 (Antioxidant copper chaperone) and cytb (Cytochrome b) and LOC101160286 (Heme binding protein). The statistical parameters of Cu/Zn SOD PPI network such as number of nodes, number of edges, average node degree, average local clustering coefficient and PPI enrichment P-value were found to be 11, 26, 4.73, 0.798 and 0.00219.

## Conclusion

The research indicated a significant reduction in protein content in fish exposed to fluoride. Comparative muscle proteome analysis indicated that the majority of protein bands were similar across both treatment groups. But when exposed to NaF, two new protein bands with molecular weights of 31 and 18 KDa were discovered. These protein bands were recognised as caspase-3 and Cu/Zn SOD. The physicochemical characterisation indicated that the caspae-3 protein and Cu/Zn SOD in fish exhibited acidic characteristics, with pI values of 6.07 and 5.69 respectively.

The secondary structure analysis revealed that caspase-3 and Cu/Zn SOD exhibit a greater proportion of random coils. Furthermore, the homology modelling showed that the best templates for the caspase-3 and Cu/Zn SOD are C5jftB and d1q0ea. Caspase-3 and Cu/Zn SOD predicted models had QMEAN values of 0.7 and -0.26 respectively. The present study showed that the predicted structural models of caspase-3 and Cu/Zn SOD were accurate and coherent.

## Acknowledgement

The authors acknowledge the Department of Environmental Sciences, Gitam Institute of Sciences, Visakhapatnam, for providing us the laboratory facilities to carry out the research.

## References

1. Aziz F., Azmat R. and Jabeen F., Skin permeability induced absorption of metals under fluoridation in edible fish *Notopterus notopterus*, Keenjhar Lake, Thatta, Sindh, Pakistan, *International Journal of Environmental Sciences*, **3(6)**, 2339-2347, doi:10.6088/ijes.2013030600049 **(2013)**

2. Bairoch A. and Apweiler R., The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, *Nucleic Acids Research*, **28(1)**, 45-48, doi: 10.1093/nar/28.1.45 **(2000)**

3. Bajpai S. and Tripathi M., Retardation of growth after fluoride exposure in catfish, *Heteropneustis fossilis* (Bloch), *Bioresources for Rural Livelihood*, **1**, 167-173 **(2010)**

4. Barbier O., Arreola-Mendoza L. and Del Razo L.M., Molecular mechanisms of fluoride toxicity, *Chemico-biological Interactions*, **188(2)**, 319-333, doi: 10.1016/j.cbi.2010.07.011 **(2010)**

5. Benkert P., Biasini M. and Schwede T., Toward the estimation of the absolute quality of individual protein structure models, *Bioinformatics*, **27(3)**, 343-350, doi: 10.1093/ bioinformatics/btq662 **(2011)**

6. Biasini M., Bienert S., Waterhouse A., Arnold K., Studer G., Schmidt T., Kiefer F., Cassarino T.G., Bertoni M., Bordoli L. and Schwede T., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Research*, **42(W1)**, W252-W258, doi: 10.1093/nar/gku340 **(2014)**

7. Buxbaum E., Fundamentals of protein structure and function, Springer, New York, 31 **(2007)**

8. Chitra T., Reddy M.M. and Rao J.R., Levels of muscle and liver tissue enzymes in *Channa punctatus* Bloch exposed to NAF, *Fluoride*, **16(1)**, 48-51 **(1983)**

9. Czerwinska U., Calzone L., Barillot E. and Zinovyev A., DeDaL: Cytoscape 3 app for producing and morphing data-driven and structure-driven network layouts, *BMC Systems Biology*, **9**, 1-11, https://doi.org/10.1186/s12918-015-0189-4 **(2015)**

10. DADF, Annual report 2016-17, Department of Animal Husbandry, Dairying and Fisheries, Ministry of Agriculture, Government of India, 162 **(2017)**

11. Ferreira R.R., Fornazier R.F., Vitória A.P., Lea P.J. and Azevedo R.A., Changes in antioxidant enzyme activities in soybean under cadmium stress, *Journal of plant Nutrition*, **25(2)**, 327-342, https://doi.org/10.1081/PLN-100108839 **(2002)**

12. Foran C.M. and Bass A.H., Preoptic GnRH and AVT: axes for sexual plasticity in teleost fish, *General and Comparative Endocrinology*, **116(2)**, https://doi.org/10.1006/gcen.1999.7357, 141-152 **(1999)**

13. Franceschini A., Szklarczyk D., Frankild S., Kuhn M., Simonovic M., Roth A., Lin J., Minguez P., Bork P., Von Mering C. and Jensen L.J., STRING v9. 1: protein-protein interaction networks, with increased coverage and integration, *Nucleic Acids Research*, **41(D1)**, https://doi.org/10.1093/nar/gks1094, D808-D815 **(2012)**

14. Garlov P.E., Plasticity of nonapeptidergic neurosecretory cells in fish hypothalamus and neurohypophysis, *International Review of Cytology*, **245**, 123-170, https://doi.org/10.1016/S0074-7696(05)45005-6 **(2005)**

15. Geetha V., Chathur K.N., Ramkumar S. and Halami P.M., *In vitro* fermentation of glycosaminoglycans from mackerel fish waste and its role in modulating the antioxidant status and gut microbiota of high fat diet-fed C57BL/6 mice, *Food & Function*, **14(15)**, https://doi.org/10.1039/D2FO03603G, 7130-7145 **(2023)**

16. Godwin J. and Thompson R., Nonapeptides and social behavior in fishes, *Hormones and Behavior.*, **61(3)**, 230-238, https://doi.org/10.1016/j.yhbeh.2011.12.016 **(2012)**

17. Guruprasad K., Reddy B.B. and Pandit M.W., Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, *Protein Engineering, Design and Selection*, **4(2)**, 155-161 **(1990)**

18. Harikrishnan R., Balasundaram C. and Heo M.S., Fish health aspects in grouper aquaculture, *Aquaculture*, **320(1-2)**, 1-21, https://doi.org/10.1016/j.aquaculture.2011.07.022 **(2011)**

19. Kale M.D. and Muley D.V., Biochemical alternation in fresh water fish *Labeo rohita* exposed to the sodium fluoride (NAF), *IOSR Journal of Environmental Science, Toxicology and Food Technology*, **9**, 48-52, doi: 10.9790/2402-09134852 **(2015)**

20. Kelley L.A., Mezulis S., Yates C.M., Wass M.N. and Sternberg M.J., The Phyre2 web portal for protein modeling, prediction and analysis, *Nature Protocols*, **10(6)**, 845-858 **(2015)**

21. Kiran M.K., Sandeep B.V. and Sudhakar P., Identification and in-silico characterization of differentially expressed salt-induced proteins in the leaves of mangrove grass *Myriostachya wightiana*, *Journal of Applied Biology & Biotechnology*, **8(5)**, 48-58, doi: 10.7324/JABB.2020.80506 **(2020)**

22. Kumar A., Tripathi N. and Tripathi M., Fluoride-induced biochemical changes in fresh water catfish (*Clarias batrachus*, Linn.), *Fluoride*, **40(1)**, 37-41 **(2007)**

23. Laemmli U.K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, **227(5259)**, 680-685, https://doi.org/10.1038/227680a0 **(1970)**

24. Lowry O.H., Rosebrough N.J., Farr A.L. and Randall R.J., Protein measurement with the Folin phenol reagent, *J biol Chem.*, **193(1)**, 265-275 **(1951)**

25. Mann M., Hendrickson R.C. and Pandey A., Analysis of proteins and proteomes by mass spectrometry, *Annual Review of Biochemistry*, **70(1)**, 437-473, https://doi.org/10.1146/annurev.biochem.70.1.437 **(2001)**

26. Manna P., Sinha M. and Sil P.C., A 43 kD protein isolated from the herb *Cajanus indicus* L attenuates sodium fluoride-induced hepatic and renal disorders *in vivo*, *BMB Reports*, **40(3)**, 382-395, https://doi.org/10.5483/bmbrep.2007.40.3.382 **(2007)**

27. Meenakshi V.K., Garg K. and Renuka M.A., Ground water quality in some villages in Haryana, India: Focus on fluoride and fluorosis, *Journal Hazardous Material*, **106**, 85-97, doi:/10.1016/j.jhazmat.2003.09.007 **(2004)**

28. Narwaria Y.S. and Saksena D.N., Acute toxicity bioassay and behavioural responses induced by sodium fluoride in freshwater fish *Puntius sophore* (bloch), *Fluoride*, **45(1)**, 7-12 **(2012)**

29. Neelamathi E., Vasumathi E., Bagyalakshmi S. and Kannan R., Insilico prediction of structure and functional aspects of a hypothetical protein of *Neurospora crassa*, *Journal of Cell and Tissue Research*, **9(3)**, 1989 **(2009)**

30. Rogers S., Wells R. and Rechsteiner M., Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis, *Science*, **234(4774)**, doi: 10.1126/science.2876518, 364-368 **(1986)**

31. Sambrook J. and Russell D.W., Molecular Cloning: Ch. 8, *In Vitro* amplification of DNA by the polymerase chain reaction, Cold Spring Harbor Laboratory Press, 2 **(2001)**

32. Shelar A. and Bansal M., Sequence and conformational preferences at termini of α-helices in membrane proteins: Role of the helix environment, *Proteins: Structure, Function and Bioinformatics*, **82(12)**, https://doi.org/10.1002/prot.24696, 3420-3436 **(2014)**

33. Sippl M.J., Recognition of errors in three-dimensional structures of proteins, *Proteins: Structure, Function and Bioinformatics*, **17(4)**, https://doi.org/10.1002/prot.340170404, 355-362 **(1993)**

34. Szklarczyk D., Gable A.L., Lyon D., Junge A., Wyder S., Huerta-Cepas J., Simonovic M., Doncheva N.T., Morris J.H., Bork P. and Jensen L.J., STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Research*, **47(D1)**, https://doi.org/10.1093/nar/gky1131, D607-D613 **(2019)**

35. Szklarczyk D., Morris J.H., Cook H., Kuhn M., Wyder S., Simonovic M., Santos A., Doncheva N.T., Roth A., Bork P. and Jensen L.J., The STRING database in 2017: quality-controlled protein–protein association networks, made broadly

accessible, *Nucleic Acids Research*, gkw937, https://doi.org/10.1093/nar/gkw937 **(2016)**

36. Ulmschneider M.B. and Sansom M.S., Amino acid distributions in integral membrane protein structures, *Biochimica et Biophysica Acta (BBA)-Biomembranes*, **1512(1)**, 1-14, https://doi.org/10.1016/S0005-2736(01)00299-1 **(2001)**

37. Uslu B., Effect of fluoride on hemoglobin and hematocrit, *Fluoride*, **14(1)**, 38-41 **(1981)**

38. Vidhya V.G., Upgade A., Bhaskar A. and Deb D., *In silico* characterization of bovine (*Bos taurus*) antiapoptotic proteins, *J Proteins Proteom.*, **3(3)**, 187-196 **(2012)**

39. Wagh Pralhad S., Kahandal Sandeep S. and Bhumkar Ajit N., Thermophysical study for Binary mixtures of Dimethyl carbonate with Anisole, Butyl vinyl ether, Diisopropyl ether and Dibutyl ether at 303.15, 308.15 and 313.15 K, *Res. J. Chem. Environ.*, **28(11)**, 25-36 **(2024)**

40. WHO, International Standards for Drinking Water, 2nd ed., World Health Organization, Geneva **(1984)**

41. WHO, Inadequate or excess Fluoride: A major public health concern, Department of Public Health, Environmental and Social Determinants of Health, WHO/CED/PHE/EPE/19.4.5, World Health Organization, Geneva **(2019)**

42. Yan X., Feng C., Chen Q., Li W., Wang H., Lv L., Smith G.W. and Wang J., Effects of sodium fluoride treatment in vitro on cell proliferation, apoptosis and caspase-3 and caspase-9 mRNA expression by neonatal rat osteoblasts, *Archives of Toxicology*, **83**, 451-458, https://doi.org/10.1007/s00204-008-0365-z **(2009)**

43. Yilmaz M., Gul A. and Erbasli K., Influence of water hardness on cadmium toxicity to *Salmo gairdneri* (Rush), *Bull. Environ. Contam. Toxicol.*, **56(4)**, 575-582 **(2004)**

44. Zaccaria D., Greco R., Bozzaro S., Ceccarelli A. and MacWilliams H., UGUS, a reporter for use with destabilizing N-termini, *Nucleic Acids Research*, **26(4)**, 1128-1129, https://doi.org/10.1093/nar/26.4.1128 **(1998)**.

(Received 29th April 2025, accepted 16th June 2025)