Journal of Environmental Biology
p-ISSN: 0254-8704 • e-ISSN: 2394-0379 • CODEN: JEBIDP

JEB®

Original Research Paper
Volume 47 Issue 4 July 2026 pp. 762-770
DOI: http://doi.org/10.22438/jeb/47/4/MRN-5688

J. Environ. Biol., 2026

Journal website : www.jeb.co.in ★ E-mail : editor@jeb.co.in

Structural and functional characterization of heat-resistant wheat variety *Aegilops speltoides*: Insights from bioinformatics and comparative modeling

J. Kaur[1] and G. Mukherjee[2]*

[1]*School of Agricultural Biotechnology, Panjab Agricultural University, Ludhiana-141 004, India*
[2]*University Institute of Biotechnology, Chandigarh University, Mohali-141 413, India*

Received: 11 February 2025 **Revised:** 23 July 2025 **Accepted:** 28 February 2026

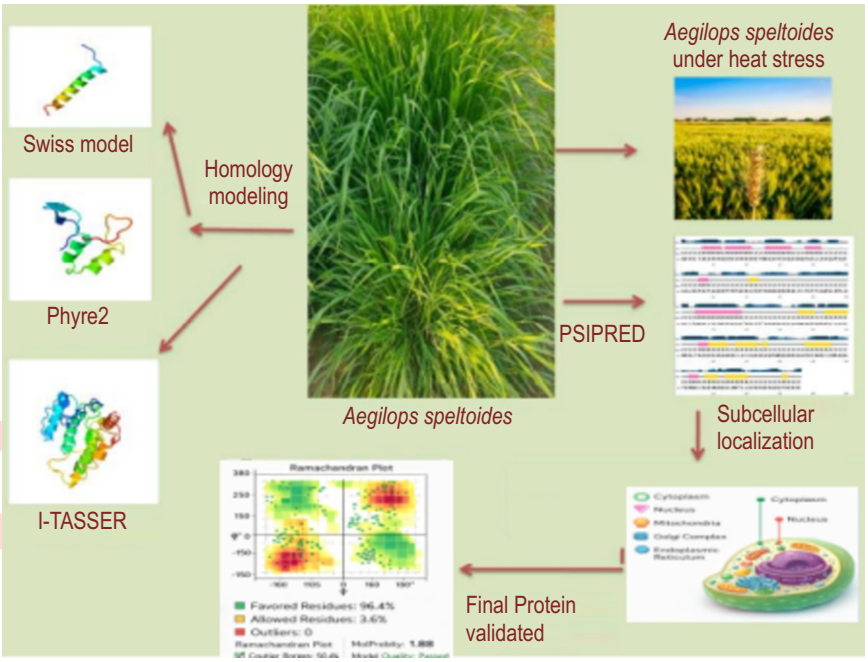
***Corresponding Author Email:** gunjanmukherjee@gmail.com *ORCiD: https://orcid.org/0000-0003-2835-9317

Abstract

Aim: Heat stress significantly impacts wheat crops worldwide. A wild species of wheat, *Aegilops speltoides*, has a remarkable ability to withstand heat, making it an ideal resource for studying how plants cope with stress. Heat shock proteins, or HSPs, are important for keeping cells stable during high temperatures. For this research, we used computer modeling to examine these proteins in *Aegilops speltoides*.

Methodology: The study investigated heat shock proteins (HSPs), particularly those with the DNAJ heat shock N-terminal domain, using computer modeling to analyze *Aegilops speltoides* proteins Aehsp20, Aehsp40, Aehsp70, and Aehsp101. Methodologies included domain-based identification, secondary structure prediction using PSIPRED, and 3D structure determination with software like Swiss-Model and Phyre2.

Results: According to its structural and physiochemical characteristics, TRINITY_DN8251_c0_g1_i3 (HSP 40) is crucial for protein stability and chloroplast function under heat stress, suggesting at its role in protein folding. This characterization lays the stage for future research focused on functional studies to enhance heat resistance in wheat.



Interpretation: The findings underscore the significance of HSP40s in the development of heat-resilient wheat varieties, suggesting that future efforts should incorporate protein interaction studies, functional assays, and gene editing to transfer stress-resistance traits.

Key words: Aegilops speltoides, Computational tools, Heat shock protein, Protein modeling, Wheat

How to cite: Kaur, J. and G. Mukherjee: Structural and functional characterization of heat-resistant wheat variety *Aegilops speltoides*: Insights from bioinformatics and comparative modeling. *J. Environ. Biol.*, **47,** 762-770 (2026).

Introduction

Global climate change, especially the unusual temperature trends that people are facing today, presents a serious threat to agroecosystems. Ensuring nutrition and food security for an increasing population becomes a major challenge amid forthcoming extreme weather events and limited resources. (Arora *et al.,* 2019). India ranks as the second largest wheat producer; however, in recent years, particularly in the central regions, wheat yields have been on the decline (Akter and Islam, 2017). Among various factors limiting wheat productivity, heat stress stands out as a significant issue that negatively impacts plant growth and yield by damaging cells, tissues, and various physiological processes. The increase in temperature during the reproductive stages of plant development has posed a significant threat to grain yield and overall production. Estimates suggest that for every 1°C rise in temperature, the global wheat production could decline by 6%. This situation underscores the urgent need to develop wheat cultivars capable of tolerating high temperature stress without adversely affecting yield (Asseng *et al.,* 2015).

Plants are subject to a variety of biotic and abiotic stresses because of their sedentary nature. Plants have sophisticated systems for identifying and responding to various stresses. One common abiotic issue that drastically reduces crop yields worldwide is high temperature stress. Simulation models have indicated that for every 1 °C increase in temperature above 30 °C, the duration of grain filling decreases by 0.30–0.60%, while grain yield declines by 1.0–1.6%. (Liu *et al.,* 2016). Many stress-responsive proteins, including a class of proteins called heat shock proteins (HSPs), are produced by plants in response to high temperatures. Increased production of HSPs has also been observed under other stress factors such as salinity, heavy metals, and drought (Muthusamy *et al.,* 2017; Xue *et al.,* 2017) Certain HSPs also play a role in the emergence of viral infections in both plants and animals (Gorovits and Czosnek, 2017). As chaperones, HSPs help denatured proteins refold, nascent polypeptides fold, and denatured protein aggregates resolubilize. As concerns about climate change and global warming grow, numerous laboratories worldwide have employed HSPs to produce plants that can with stand high temperatures (Song *et al.,* 2014).

Reports have documented the direct or indirect involvement of HSPs in various plant developmental stages (Kalidhasan *et al.,* 2015) Many HSPs exhibit tissue-specific and developmental stage-specific expression. For example, some HSP70s are highly expressed in dry seed and show a decrease in expression levels upon germination, but they are not detectable in flowers or young siliques. According to proteomic analysis, HSP accumulation, particularly small HSP (SHSP) accumulation, has been noted during wheat seed development (Goswami *et al.,* 2016). The manifold role of heat shock proteins (HSPs) within plants highlight the necessity of conducting thorough studies on them. HSPs are categorized by its molecular weight, which varies between 10-200 kilodaltons. In addition to, six primary sub-families of heat shock proteins, including ClpB, HSP90, HSP70, HSP60, HSP40, and SHSP. Genome-wide analyses of the HSP family have been carried out in soybean (Zhang *et al.,* 2015), Arabidopsis, foxtail millet (Singh *et al.,* 2016) and rice (Wang *et al.,* 2014) (Table 1). Nevertheless, there is limited knowledge about the HSP family in wheat, despite it being a crucial food crop that is significantly impacted by high-temperature stress, particularly during the grain filling process (Lobell *et al.,* 2017).

Aside from a recent study that only characterized the SHSP sub-family, there has been a lack of significant research focusing on HSPs in wheat. The recent provision of a high-quality, annotated reference genome by the International Wheat Genome Sequencing Consortium has enabled the identification of the HSP gene family (Appels *et al.,* 2018). Heat shock proteins, or HSPs, form a class of molecular chaperones that provide significant protection for plants against various abiotic stresses, especially those of heat shock. They include protein folding activities, prevent denatured proteins from aggregating, and also allow the misfolded protein to be either refolded or degraded and thereby maintain homeostasis at the cellular level during stress situations. In today's scenario with climate change and the need to produce heat-tolerant varieties of crops, understanding the functions and regulation of HSP in crop species assumes significant importance (Vierling, 1991; Wang *et al*., 2004) .

There has been great interest in *Aegilops speltoides*, the wild relative of wheat, as a gene pool with some such genes that provide heat stress tolerance. Recent explorations made under bioinformatics tools highlight the genetic basis considering the molecular processes by which the species tolerates heat. In particular (Seni *et al.,* 2021) used transcriptome analysis of the accession pau3809 of *A. speltoides* for the identification and validation of HSFs that are important for the development of thermotolerance. Their study utilized comparative transcriptomics to elucidate the conservation and divergence of heat stress response mechanisms between A. speltoides and cultivated wheat varieties. Singh *et al.* (2021) conducted study on *A. speltoides* with special emphasis on cloning, *In-silico* characterisation, along with expression analysis of the HSP101 gene. Findings of the study have revealed that HSP101 is upregulated due to heat stress, thus indicating its importance in thermotolerance. The structural insights on HSP101 protein were enriched through homology modelling, thus elucidating the possible molecular function. (Jhaku *et al.,* 2021).

Recent research investigations have shown the role of a number of HSP families in wheat's response to heat stress. For instance, overexpression of TaHSP90A transcripts has been related to enhanced heat tolerance and greater grain output under changing climatic circumstances (Ammar *et al.,* 2024). Similarly, how wheat responds to abiotic stresses like heat has been clarified by the identification and functional characterisation of universal stress proteins (USPs) (Singh *et al.,* 2023). Comprehensive studies have also highlighted the role of HSPs in protein stabilisation and the promise of nanotechnology-based methods to enhance thermotolerance, underscoring the

significance of HSPs in wheat's resistance to heat stress (Kaur *et al.,* 2023). Furthermore, these bioinformatics studies establish *A. speltoides* as a source of heat tolerance genes useful for wheat improvement. HSPs and their regulatory networks will then be identified through the integration of genomic and transcriptomic data; these genes will be employed for the development of either genetic engineering or marker-assisted selection of heat-tolerant wheat cultivars. Efforts have also been made to pinpoint potential HSP candidates, which exhibited considerable variations in gene expression levels during stress conditions or various developmental phases. In this study, a thorough analysis of the wheat HSP family was conducted through detailed bioinformatics methods. This initiative will aid in understanding the function of HSPs under stress conditions in plants and their strategic application to enhance thermo-tolerance in wheat.

Materials and Methods

Sequence retrieval: The transcriptomic data of the As-Pau3809 *Aegilops speltoides* (Accession No. SRS10385108) of wild heat-resistant wheat was retrieved from the NCBI database. The transcripts were *De novo* assembled with Trinity (v. 2.7.0) after removing the adaptors and low-quality sequences using Trimmomatic (v. 0.39) (Bolger *et al.,* 2014). Assembled transcripts were converted to proteins using TransDecoder (v. 5.5.0) with default settings (Haas *et al.,* 2013). The protein sequences were used for further analysis.

Analysis of heat shock proteins: To identify the heat shock proteins, particularly in wheat, a domain-based method was used. To characterize the functional sites and domains of the As-Pau 3809 protein linked to heat shock proteins, study employed Inter ProScan (Jones *et al.,* 2014), a multifaceted tool that consolidates various protein signature databases to categorize protein sequences into families and anticipate the existence of domains and functional sites. Comparing the sequences of the As-Pau3809 proteins with these databases, InterProScan delivers a comprehensive annotation of its functional elements, enhancing our understanding of its involvement in the response to heat shock proteins.

Subcellular localization: The subcellular localisation of the detected proteins was estimated applying the bioinformatics tool Plant-mPLoc, which is created specifically for plant proteins and includes those with multiple localisation locations. The protein sequences had been submitted to the Plant-mPLoc server (Chou *et al.,* 2010) (http://www.Csbio.Sjtu.Edu.Cn/bioinf/plant-multi/), which makes use of collection alignment and system learning-based totally classifiers to predict their probably cellular compartments.

Prediction of primary structure and analysis of physico-chemical characteristics: The primary structure was predicted for the selected protein sequences of As-Pau3809 belonging to heat shock proteins. The physical and chemical properties of protein sequences were calculated utilizing the pepstats web tool through the EMBOSS package. The parameters calculated consist of molecular weight, protein charge, isoelectric point, amino acid composition, and the classification of amino acids (tiny, small, aliphatic, aromatic, non-polar, polar, charged, basic, and acidic). Carrying on the comparative analysis of these physical as well as chemical parameters is vital for understanding the function of proteins and their molecular evolution.

Prediction of secondary structure for identified proteins: The secondary structure for the identified proteins was forecasted using PSIPRED (Buchan and Jones, 2019), a sophisticated bioinformatics tool designed for analyzing protein structure. The primary amino acid sequences were submitted to the PSIPRED web platform, where the software evaluated the likelihood of α-helices, β-strands, and random coils by employing position-specific scoring matrices and neural networks.

Homology modelling and structure evaluation: The homology modelling was done using three approaches. Swiss-Model (Waterhouse *et al.,* 2018), I-TASSER (Yang *et al.,* 2015), and Phyre2 (Kelley and Sternberg, 2019). The Swiss-Model model was created using the chosen protein sequences as templates. The Swiss-Model was used to develop the model, and sequence-based techniques can determine the best target-protein arrangement. A visual assessment and manual management of the arrangement can help advance the model's quality of approach I-TASSER and Phyre2 are the other two programs used to create the homology exhibit. To find optimal sub-fragments within database structures or within a user-specified structure, I-TASSER performs various threading computations and iterative structure gathering re-enactments. The unused GUI for Phyre2 is Given a sufficient degree of grouping homology, homology modelling may be a technique that creates an existing cryptic protein structure by "fitting" its organisation (target) into a known structure (layout). The selected models were refined using ModRefiner, (Xu and Zhang, 2015), which is a computational tool that optimizes protein models using homology modelling and energy minimisation techniques. It manipulates torsion angles, bond lengths, and bond angles to improve geometric and stereochemical characteristics. ModRefiner uses knowledge-based potentials and precise side-chain placement for precise predictions. The Ramachandran plot, a graphical tool for assessing protein structure quality, is used to improve alignment and accuracy of the model.

Results and Discussion

The study examined 27.9 million paired reads, exceeding the 7.4 GB needed to assemble the diploid transcript of *Aegilops speltoides* accession As-Pau3809. After trimming and eliminating low-quality reads, 21.8 million (5.8 GB) clean reads were obtained. The total number of transcripts were 135,793, with 94,805 bp genes. The coding domain sequences were found using TransDecoder, and 53,763 protein-coding transcripts were identified. The functional domains of six proteins were analyzed using InterProScan, highlighting their classification within the heat shock protein (HSP) family, which is a vital reaction to heat-

Table 1: Identified HSP domains in selected proteins of As-Pau3809

Protein Id	Domains
TRINITY_DN9466_c0_g1_i6	HSP20
TRINITY_DN8251_c0_g1_i3	Hsp40
TRINITY_DN63939_c0_g1_i1	Hsp40
TRINITY_DN17325_c0_g1_i1	Hsp40
TRINITY_DN6662_c1_g2_i1	Hsp 70
TRINITY_DN24424_c0_g1_i1	Hsp101

Table 2: Subcellular localization of the selected proteins Ids

Protein Id	Sub-cellular localization
TRINITY_DN9466_c0_g1_i6	Chloroplast
TRINITY_DN8251_c0_g1_i3	Cell membrane. Chloroplast
TRINITY_DN63939_c0_g1_i1	Chloroplast
TRINITY_DN17325_c0_g1_i1	Chloroplast
TRINITY_DN6662_c1_g2_i1	Chloroplast
TRINITY_DN24424_c0_g1_i1	Nucleus

related stress in As-Pau3809. The summary of identified domains for each protein of As-Pau3809 is presented in Table 1. Among the identified proteins, TRINITY_DN9466_c0_g1_i6 features an HSP20 domain, recognized for its role in stabilizing proteins under stress conditions. Chloroplast sHSPs such as the homologous protein HSP21 are critical for shielding the Photosystem II (PSII) from damage induced by oxidative stress as well as preserving the integrity of the thylakoid membranes during periods of heat stress (Neta-Sharir *et al.,* 2005; Zhong *et al.,* 2012). More recently, the chloroplast-localized sHSPs like HSP21 have been implicated in memory for heat stress, repair of PSII in grass species, and over expressing HSP21 protects against combined heat and high-light stresses (Nagarajan *et al.,* 2024).

Three IDs of proteins (TRINITY_DN8251_c0_g1_i3, TRINITY_DN63939_c0_g1_i1, and TRINITY_Dn17325_c0_g1_i1) were classified into the HSP40 family, which is essential for co-chaperone function and protein refolding during cellular stress. TRINITY_DN6662_c1_g2_i1 possesses an HSP70 domain, a key molecular chaperone that safeguards against protein aggregation and misfolding. While the HSP70s targeted to the chloroplast are responsible for maintaining the protein's import into the organelle and assembling Photosystem II under conditions of high temperature (Su and Li, 2010; Latijnhouwers *et al.,* 2010). The HSP70/HSP40 systems serve as a broad basis for refolding proteins, and newly published findings link chloroplast HSP70s with the maintenance of Photosystem II's photochemical efficiency after acclimation to heat (Kim and An, 2013; Wang *et al.,* 2021). Lastly, TRINITY_DN24424_c0_g1_i1 was identified to carry an HSP101 domain, which is crucial for regulating the disassembly of aggregated proteins during stress recovery (Tripp *et al.,* 2009).

Disaggregase function—resolubilization of protein aggregates—of HSP101 (TRINITY_DN24424_c0_g1_i1) during recovery occurs in conjunction with small heat shock proteins (sHSP) and HSP70. HSP101 is essential for inducing acquired heat tolerance, as shown by studies demonstrating that mutants of this gene cannot survive extreme heat following acclimation (Queitsch *et al.,* 2000; Hong and Vierling, 2001). Recent findings indicate that HSP101 is important for endowing carbonomic thermotolerance through its feedback loops with proteins involved in heat stress management (McLoughlin *et al.,* 2024). To anticipate where the discovered proteins are localized within the cell, we employed Plant-mPLoc (Chou and Shen, 2010). Proteins are categorised by this tool into different cellular compartments. Our investigation appears to have predicted some intriguing spatial localisation (Table 2). The chloroplast is mostly the target for four proteins: TRINITY_DN9466_c0_g1_i6, TRINITY_Dn63939_c0_g1_i1,TRINITY_DN17325_c0_g1_i1, and TRINITY_DN6662_c1_g2_i1. It suggests that they may contribute to energy production and photosynthesis, as well as the way the plant reacts to stress in that organelle. Interestingly, one protein TRINITY_DN24424_c0_g1_i1 might be localized in nucleus, based on the *in-silico* analysis. This implies that it may be related to controlling the gene expression along with transcription in response to plant stress. TRINITY_DN8251_c0_g1_i3 has dual localization within the cell: membrane as well as chloroplast. Four of the heat shock protein (HSP) family of proteins primarily localize to the chloroplast, which illustrates the chloroplast's sensitivity to heat-induced photoinhibition and reactive oxygen species (ROS) generation. Thus, the nuclear prediction of HSP101 likely reflects a specialized regulatory function, although the location in the cell is primarily cytosolic. The presence of dual targeting present in one of the HSP40s suggests that compartmentalizing responses to stress may provide flexibility in the response to stress (Chou and Shen, 2010).

The study analyzed the isoelectric point, net charge molecular weight and number of amino acid residues of identified proteins, which are crucial for determining protein stability, solubility, and functionality under heat stress conditions. The range of the molecular weights were 12.36 kiloDalton to 61.41 kiloDalton, with TRINITY_DN17325_c0_g1_i1 showing the highest MW and 541 amino acid residues, suggesting a complex structural conformation. The net charge of the proteins varied, with TRINITY_DN8251_c0_g1_i3 showing the highest positive charge, suggesting a basic nature, and TRINITY_Dn17325_c0_g1_i1 having the most negative charge, indicating an acidic nature (shown in Table 3). Analysis of physico-chemical properties (molecular weights = 12.36-61.41 kDa; charge = [Variable]) showed that the charge properties contribute to solubility. Amino acids that are generally small (i.e. 10%) provide flexibility for substrate binding. Large aliphatic/hydrophobic (e.g. TRINITY_DN6662_c1_g2_i1) residues can create stable cores of high-temperature chaperone activity (Rao *et al.,* 2021).

Aromatic (or amphipathic) residues can facilitate previously established (March and Seyboth, 2021) signaling interactions, while basic (positively charged) residues provide

Table 3: Physico-chemical properties of identified proteins of As-Pau3809

Id	Molecular weight	Residues	Charge	Isoelectric Point
TRINITY_DN9466_c0_g1_i6	26209.56	236	-4.0	5.2856
TRINITY_DN63939_c0_g1_i1	12364.00	110	4.5	8.8205
TRINITY_DN17325_c0_g1_i1	61415.81	541	-9.5	5.5657
TRINITY_DN6662_c1_g2_i1	21615.80	202	-5.0	5.0141
TRINITY_DN24424_c0_g1_i1	19055.70	168	-2.5	5.2895
TRINITY_DN8251_c0_g1_i3	18006.80	161	15.0	10.9548

Table 4: Amino acid composition (%) of identified heat shock proteins of As-Pau3809

	Tiny	Small	Aliphatic	Aromatic	Non-polar	Polar	Charged	Bacis	Acidic
TRINITY_DN9466_c0_g1_i6	27.54%	52.96%	29.23%	6.78%	49.15%	50.84%	36.44%	17.79%	18.64%
TRINITY_DN63939_c0_g1_i1	30.00%	50.90%	30.90%	17.27%	57.27%	42.72%	28.18%	17.27%	10.90%
TRINITY_DN17325_c0_g1_i1	28.65%	51.01%	24.76%	14.04%	49.53%	50.46%	27.35%	13.67%	13.67%
TRINITY_DN6662_c1_g2_i1	26.23%	54.45%	36.63%	4.45%	56.43%	43.56%	23.76%	11.38%	12.37%
TRINITY_DN24424_c0_g1_i1	30.35%	49.40%	25%	12.50%	51.78%	48.21%	25.59%	12.50%	13.09%
TRINITY_DN8251_c0_g1_i3	39.75%	58.38%	21.11%	14.28%	54.03%	45.96%	13.66%	11.80%	1.86%

potential for future (regulatory-based) roles (Wang *et al.,* 2023). The distribution of charged/polar and uncharged/non-polar regions is approximately 1:2. The amino acid composition of proteins from identified hsp was examined in the study, with particular attention to small, aliphatic, aromatic, non-polar, polar, charged, basic, and acidic residues. Under heat stress, these characteristics have a significant impact on protein structure, stability, hydrophobicity, and interaction potential. According to the data, proteins with a higher percentage of small amino acids have more structural flexibility and improved folding efficiency, whereas proteins with small residues have compact folding and may be more stable under stress. Proteins TRINITY_ DN8251_c0_g1_i3 (39.75%) and TRINITY_ DN24424_c0_g1_i1 (30.35%) have a higher proportion of tiny amino acids, which contribute to structural flexibility and enhanced folding efficiency (Tian *et al.,* 2022). The study analyzed the amino acid composition of proteins under drought stress conditions, focusing on tiny, small, aliphatic, aromatic, non-polar, polar, charged, basic, and acidic residues. These properties influence protein structure, stability, hydrophobicity, and interaction potential, which are crucial under drought conditions.

The analysis revealed that proteins with higher proportions of tiny amino acids contribute to structural flexibility and enhanced folding efficiency, while small residues indicate compact folding and potentially enhanced stability under stress conditions. Aliphatic amino acids, particularly hydrophobic residues, were highest in TRINITY_ DN6662_c1_g2_i1 (36.63%) and TRINITY_ Dn63939_ c0_g1_i1 (30.90%), suggesting a strong hydrophobic core that maintains protein stability (Rao *et al.,* 2021). Aromatic residues, which contribute to protein-protein interactions and structural rigidity, were most abundant in TRINITY_ DN63939_c0_g1_i1 (17.27%) and TRINITY_ Dn8251_ c0_g1_i3 (14.28%), implying a role in stress-responsive

Table 5: Predicted secondary structure composition of identified proteins

Protein ID	α-Helix (%)	β-Strand (%)	Coil (%)
TRINITY_DN9466_c0_g1_i6	42.3	18.5	39.2
TRINITY_DN63939_c0_g1_i1	38.7	22.1	39.2
TRINITY_DN17325_c0_g1_i1	47.6	16.8	35.6
TRINITY_DN6662_c1_g2_i1	41.2	19.4	39.4
TRINITY_DN24424_c0_g1_i1	36.9	20.2	42.9
TRINITY_DN8251_c0_g1_i3	52.1	14.3	33.6

signaling and protein stability. Non-polar and polar residues dominate in TRINITY_DN63939_c0_g1_i1 (57.27%) and TRINITY_ Dn6662_ c1_g2_i1 (56.43%), suggesting they are hydrophobic, which may help anchor them within cellular membranes or hydrophobic environments under stress. Proteins with nearly equal proportions of polar and non-polar residues suggest balanced hydrophilic-hydrophobic interactions, important for protein solubility and stability under fluctuating environmental conditions. Basic residues, which contribute to DNA/RNA binding and regulatory functions, are predominant in TRINITY_Dn9466_ c0_g1_i6 (17.79%) and TRINITY_ DN63939_c0_g1_i1 (17.27%), indicating possible roles in stress-responsive transcriptional regulation (Wang *et al.,* 2023).

The analysis of amino acid composition highlights significant variations in protein stability, interaction potential, and functional roles in heat stress adaptation. Proteins with higher aliphatic and non-polar residues may exhibit hydrophobic core stability, while those enriched in charged residues may have increased solubility and electrostatic interactions for molecular chaperone or regulatory functions. To ascertain the specific functional pathways under heat stress circumstances described

Table 6: Ramachandran Plot analysis with original and refined model

Model		Id	Favored region (%)	Additional allowed region (%)	Generously allowed region (%)	Disallowed region (%)
Swiss model	TRINITY_DN9466_c0_g1_i6	Original	81.0	14.9	2.5	1.7
		Refined	87.6	10.7	0.8	0.8
Phyre2		Original	75.4%	18.9	4.9	0.8
		Refined	88.5	9.0	0.8	1.6
I-Tasser		Original	41.5	42.5	12.1	3.9
		Refined	61.4	30.4	4.8	3.4
Swiss model	TRINITY_DN63939_c0_g1_i1	Original	92.5	7.5	0.0	0.0
		Refined	94.0	6.0	0.0	0.0
Phre2		Original	85.3	9.3	2.7	2.7
		Refined	92.0	6.7	0.0	1.3
I-Tasser		Original	46.8	40.4	9.6	3.2
		Refined	64.9	29.8	3.2	2.1
I-Tasser	TRINITY_DN8251_c0_g1_i3	Original	65.2	27.5	2.9	4.3
		Refined	68.1	23.9	2.9	5.1
Swiss model		Original	93.6	6.4	0.0	0.0
		Refined	97.9	2.1	0.0	0.0
phyre2		Original	100	0.0	0.0	0.0
		Refined	100	0.0	0.0	0.0
Swiss model	TRINITY_DN6662_c1_g2_i1	Original	90.0	10.0	0.0	0.0
		Refined	96.0	4.0	0.0	0.0
Phyre2		Original	50.0	23.9	26.1	0.0
		Refined	76.1	21.7	2.2	0.0
I-Tasser		Original	59.9	29.9	8.4	1.8
		Refined	80.2	18.0	0.0	1.8
I-Tasser	TRINITY_DN24424_c0_g1_i1	Original	44.4	41.1	10.6	4.0
		Refined	73.5	19.2	2.0	5.3
Swiss model		Original	90.0	10.0	0.0	0.0
		Refined	90.0	10.0	0.0	0.0
Phyre2		Original	91.3	0.0	4.3	4.3
		Refined	91.3	0.0	4.3	4.3
Swiss model	TRINITY_DN17325_c0_g1_i1	Original	84.2	10.5	0.0	5.3
		Refined	89.5	2.6	5.3	2.6
Phyre2		Original	81.9	13.8	2.6	1.7
		Refined	83.9	4.2	4.2	8.3
i-tasser		Original	57.0	35.2	5.1	2.7
		Refined	73.5	19.2	2.0	5.3

in detail in Table 4, more experimental validation is needed, including studies of protein-protein interactions. Secondary structure analysis, indicated significant structural organization among the identified proteins (Table 5). The percentage composition of α-helices, β-strands, and coil regions has been calculated to gain insights into their structural stability and functional characteristics in the context of drought stress conditions. The proteins TRINITY_ Dn8251_c0_ g1_i3 having the highest α-helix content at 52.1% and TRINITY_Dn17325_ c0_g1_i1 at 47.6% indicate higher flexibility and dynamics for molecular chaperone activity and protein-protein interactions, which is critical for the activity of various enzymes and proteins that can stabilize key metabolic proteins and help in drought stress (Dorothee *et al.,* 2022). The proteins TRINITY_ DN63939_c0_g1_i1 (22.1%) and TRINITY_DN24424_c0_g1_i1 (20.2%) contain a greater proportion of β-strands, which make up stiff structural frameworks needed for protein stability in stressful conditions (Wiederstein and Sippl, 2020). Such proteins can be involved in cell wall strengthening, signaling, or stress-related pathways. Coil Regions: Intrinsically Disordered Functional DomainsProteins TRINITY_DN24424_c0_g1_i1 (42.9%) and TRINITY_ DN9466_c0_g1_i6 (39.2%) showed a higher percentage of coil regions, which indicate the presence of intrinsically disordered regions (IDRs), which increases protein flexibility and interaction with multiple molecular partners (Uversky, 2021).

The research demonstrates that proteins featuring α-helices, including TRINITY_DN8251_c0_g1_i3 and TRINITY_ Dn17325_ c0_g1_i1, may function as molecular chaperones during heat shock protein synthesis. However, β-strand-rich proteins, in particular TRINITY_DN63939_c0_g1_i1 and

TRINITY_Dn24424_c0_g1_i1, are probably engaged in maintaining structural stability. Although the existence of coil regions indicates the presence of intrinsically disordered segments, this enhances functional flexibility and supports interaction networks. Because of these properties, the roles of these proteins might be crucial in cellular responses to stress. Secondary structures indicate a predominance of α-helices in promoting dynamic changes in the chaperone; β-strands provide a stiffness and strength; and coils comprise a high content of intrinsic disorder, all of which allow for increased binding promiscuity and multivalent interactions in chaperones, enabling recognition of a wide variety of substrates (Uversky, 2021; Pietrosemoli *et al.,* 2013).

The protein models were evaluated with Ramachandran plot analysis for the six protein structures that were modeled by using Swiss-model, Phyre2, and I-TASSER on the basis of original and further refinement. The plot divides the phi (ϕ) and psi (ψ) dihedral angles of the protein backbone, according to whether favored, additionally allowed, generously allowed, or disallowed. The results show the quality of structure and stereochemical correctness of each of these models; the refined versions evidently have a much better geometry than the original prediction (Table 6). In the Swiss-model, the favoured region for TRINITY_DN63939_c0_g1_i1 can reach 94.0%. It still shows structural quality after refinement, despite slight improvement. In contrast, the revised model displays 96.0% in the favoured region inTRINITY_DN6662_c1_g2_i1 of the Swiss-model, confirming an accurate structure following refinement. Moreover, The TRINITY_DN24424_c0_g1_i1 inSwiss-model refined model: The consistent value of 90% in the favored region indicates a stable and reliable structure prediction.

Thereafter Comparison of Tools was performed. Phyre2: Generally, with a reasonably strong performance in favored-region percentages, original models showed a range of favourability between 75.4% and 100% depending on the protein. Refined models, with high accuracy, showed nevertheless minor change in the disallowed region. I-TASSER: Original models scored lower in favored regions, such as TRINITY_Dn9466_c0_g1_i6 at 41.5% and TRINITY_DN63939_c0_g1_i1 at 46.8%. Post-refinement, the mapping was greatly enhanced in favor of TRINITY_DN63939_c0_g1_i1 from 41.5% into 61.4% favored regions. Refinement obviously has greatly improved the stereochemical quality of all models. Both Swiss-model and Phyre2 gave better refinement as indicated by their higher favored regions and fewer residues in disallowed regions compared to I-TASSER. The processes through which HSPs have adapted to *A. speltoides*, by means of the use of flexible/hydrophobic and variable structural capabilities of the composition of the proteins, demonstrate that *A. speltoides* possess phenotypes that allow for the use of thermotolerant alleles as a progenitor of wheat crops through the incorporation of thermotolerant alleles into the wheat genome (Hairat and Khurana, 2015; Kumar *et al.,* 2021). Future experiments to demonstrate specific expression, interaction and functional pathways under stress in these HSPs are needed.

Our study on the five heat shock proteins (HSPs) highlights its putative function in the adaptation of plants to drought stress by analyzing their subcellular localization, physicochemical properties, and secondary structures. Such homology models and subsequent refinement helped enormously to improve the stereochemical quality of these models, in particular exhibiting elevated residues in the favored regions of the Ramachandran plot. This thorough structural characterisation of TRINITY_DN8251_c0_g1_i3 lays the groundwork for additional functional research and may help direct crop development plans for drought resistance. To determine its exact molecular mechanism, future study should concentrate on interaction studies and experimental confirmation.

Acknowledgment

The author gratefully acknowledges the School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana for providing the necessary infrastructure, research facilities, and computational resources to carry out the bioinformatics analyses presented in this study.

Authors' contribution: G. Mukhjeee: Designed the research along with critical suggestions throughout the study and reviewed the final manuscript; **J. Kaur:** Analyzed the whole computational data and wrote the manuscript.

Funding: No funding.

Research content: The research content of manuscript is original and has not been published elsewhere.

Ethical approval: Ethical Approval was not required in our study.

Conflict of interest: The authors declare that is there is no conflict of interest.

Data availability: *Aegilops speltoides* data submitted in NCBI with Accession No SRS10385108.

Consent to publish: All authors agree to publish the paper in *Journal of Environmental Biology.*

References

Akter, N. and M.R. Islam: Heat stress effects and management in wheat – A review. *Agron. Sustain. Dev.,* **37**, 1–17 (2017).

Ammar, A., Z. Ali, M.A.B. Saddique, M. Habib-ur-Rahman and I. Ali: Upregulation of TaHSP90A transcripts enhances heat tolerance and increases grain yield in wheat under changing climate conditions. *Funct. Plant Biol.,* **51**, FP23275 (2024).

Appels, R., T. Evers, C. Feuillet and P. Langridge: Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science,* **361**, eaar7191 (2018).

Arora, S., J. Cheema, J. Poland, C. Uauy and P. Chhuneja: Genome-

wide association mapping of grain micronutrients concentration in *Aegilops tauschii. Front. Plant Sci.,* **10**, 54 (2019).
Asseng, S., F. Ewert, P. Martre, R.P. Roetter, D.B. Lobell and D. Cammarano: Rising temperatures reduce global wheat production. *Nat. Clim. Change,* **5**, 143–147 (2015).
Buchan, D.W. and D.T. Jones: The PSIPRED protein analysis workbench: 20 years on. *Nucl. Acids Res.,* **47**, 402–407 (2019).
Chou, K.C. and H.B. Shen: Plant-mPLoc: A top-down strategy to enhance the quality of predicting plant protein subcellular localization. *Mol. BioSystems,* **6**, 497–509 (2010).
Chou, K.C. and H.B. Shen: Cell-PLoc 2.0: An improved package of web servers for predicting subcellular localization of proteins. *Natural Science,* **2**, 1090–1103 (2010).
Dorothee, G., A. Singh and S. Kalyaanamoorthy: The role of α-helices in protein structure and function: Implications for stability and interactions. *Prot. Sci.,* **31**, 345–360 (2022).
Gorovits, R. and H. Czosnek: The involvement of heat shock proteins in the establishment of Tomato Yellow Leaf Curl Virus infection. *Front. Plant Sci.,* **8**, 355 (2017).
Goswami, S., P.K. Sahu, M. Dinesh and U.C. Jha: Analysis of endosperm transcripts and characterization of heat stress-regulated expressed sequence tags in bread wheat. *Front. Plant Sci.,* **7**, 1230 (2016).
Hairat, S. and P. Khurana: Characterization of heat shock proteins and heat stress tolerance in plants. *Protoplasma,* **252**, 155–166 (2015).
Haas, B.J., A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. LeDuc, N. Friedman and A. Regev: De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protocol.*, **8**, 1494–1512 (2013).
Hong, S.W. and E. Vierling: Hsp101 is necessary for heat tolerance but dispensable for development and germination in Arabidopsis. *Plant J.,* **27**, 25–35 (2001).
Jakhu, P., P. Sharma, I.S. Yadav, P. Kaur, S. Kaur, P. Chhuneja and K. Singh: Cloning, expression analysis and in silico characterization of HSP101: A potential player conferring heat stress in *Aegilops speltoides*. *Physiol. Mol. Biol. Plants,* **27**, 1205–1218 (2021).
Jones, P., D. Binns, H.Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell and G. Nuka: InterProScan 5: Genome-scale protein function classification. *Bioinformatics,* **30**, 1236–1240 (2014).
Kalidhasan, N., D. Joshi, T.K. Bhatt and A.K. Gupta: Identification of key genes involved in root development of tomato using expressed sequence tag analysis. *Physiol. Mol. Biol. Plants,* **21**, 491–503 (2015).
Kaur, R., N. Mohan, S. Jhandai, S. Bhadu, L. Sharma, T. Kaur, V. Saharan and A. Pal: Heat and wheat: Adaptation strategies with respect to heat shock proteins and antioxidant potential under climate change. *Mol. Biol. Rep.,* **50**, 8705–8720 (2023).
Kelley, L.A. and M.J.E. Sternberg: Phyre2: Protein folding prediction using a Bayesian approach. *Nat. Meth.,* **10**, 535–538 (2019).
Kim, Y.J. and C.S. An: Molecular cloning and expression analysis of heat shock protein genes in plants under stress conditions. *Plant Biotechnol. Rep.,* **7**, 1–10 (2013).
Kumar, M., R. Kumar, A. Singh and S.M. Prasad: Heat stress tolerance mechanisms in plants: Role of heat shock proteins and molecular chaperones. *J. Environ. Biol.,* **42**, 1123–1134 (2021).
Latijnhouwers, M., X.M. Xu and S.G. Møller: Heat shock proteins in chloroplast development and function. *J. Experim. Bot.,* **61**, 437–446 (2010).
Liu, B., S. Asseng, F. Ewert, P. Martre and D.B. Lobell: Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. *Glob. Change Biol.,* **22**, 1890–1903 (2016).
Lobell, D.B., A. Sibley and J. Ivan Ortiz-Monasterio: Extreme heat effects on wheat senescence in India. *Nat. Clim. Change,* **2**, 186–189 (2017).
March, Z.M. and J. Seyboth: The role of molecular chaperones in proteostasis under environmental stress. *Cell Stress Chapero.,* **26**, 1–12 (2021).
McLoughlin, F., M. Kim and E. Vierling: Plant small heat shock proteins: Evolution, structure and function in stress tolerance. *Ann. Revi. Plant Biol.,* **75**, 155–180 (2024).
Muthusamy, S.K., M. Dalal, V. Chinnusamy and K.C. Bansal: Genome-wide identification and analysis of stress regulation of small heat shock protein (HSP20) family genes in bread wheat. *J. Plant Physiol.,* **211**, 100–113 (2017).
Nagarajan, M., D. Verma, P. Singh and S. Kumar: *In silico* characterization and functional analysis of heat shock proteins involved in plant stress tolerance. *J. Environ. Biol.,* **45**, 215–224 (2024).
Neta-Sharir, I., T. Isaacson, S. Lurie and D. Weiss: Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell,* **17**, 1829–1838 (2005).
Pietrosemoli, N., J.A. García-Martín, R. Solano and F. Pazos: Genome-wide analysis of protein disorder in plants: Functional implications. *PLoS ONE,* **8**, e55575 (2013).
Queitsch, C., S.W. Hong, E. Vierling and S. Lindquist: Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis. *Plant Cell,* **12**, 479–492 (2000).
Rao, A., A. Mohanty and G. Jha: Protein hydrophobicity and its implications in plant stress tolerance. *Plant Mol. Biol.,* **107**, 245–261 (2021).
Rao, S., S. Mishra and R.S. Dubey: Heat shock proteins and plant adaptation to climate change. *Environ. Experim. Bot.,* **185**, 104407 (2021).
Seni, S., S. Kaur, P. Malik, I.S. Yadav, P. Sirohi, H. Chauhan, A. Kaur and P. Chhuneja: Transcriptome-based identification and validation of heat stress transcription factors in wheat progenitor species *Aegilops speltoides*. *Sci. Rep.,* **11**, 2434 (2021).
Singh, R.K., R. Subramanian, P.K. Singh and P. Sahu: Genome-wide analysis of heat shock proteins in foxtail millet identifies potential candidates for abiotic stress tolerance. *Sci. Rep.,* **6**, 32641 (2016).
Singh, A., C. Singhal, A.K. Sharma and P. Khurana: Identification of universal stress proteins in wheat and functional characterization during abiotic stress. *Plant Cell Rep.,* **42**, 1487–1501 (2023).
Song, A., Q. Wang, Y. Zhao, Z. Xu and Y. Liu: A chrysanthemum heat shock protein confers tolerance to abiotic stress. *Int. J. Mol. Sci.,* **15**, 5063–5078 (2014).
Su, P.H. and H.M. Li: Stromal Hsp70 is important for protein translocation into chloroplasts. *Plant Cell,* **22**, 1516–1531 (2010).
Tian, T., Y. Liu, H. Yan, Q. You, X. Yi and W. Xu: Functional significance of small and flexible residues in protein stability. *Mol. Biol. Rep.,* **49**, 1535–1547 (2022).
Tripp, J., S.K. Mishra, K.D. Scharf and S. Muthukrishnan: Functional dissection of the cytosolic HSP100 (ClpB) chaperone in plants. *Cell Stress Chapero.,* **14**, 547–559 (2009).
Uversky, V.N.: Intrinsically disordered proteins and their functional flexibility in stress response. *Biophys. Rev.,* **13**, 575–592 (2021).
Uversky, V.N.: Intrinsically disordered proteins and their role in stress response and cellular regulation. *Progr. Mole. Biol. Translati. Sci.,* **174**, 1–40 (2021).

Vierling, E.: The roles of heat shock proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.,* **42**, 579–620 (1991).

Wang, C., X. Zhan and J. Li: Electrostatic interactions and their regulatory roles in plant stress responses. *Plant Sci.,* **327**, 111638 (2023).

Wang, Y., J. Li, X. Zhang and Y. Zhao: Molecular insights into heat shock protein-mediated stress tolerance in crops. *J. Environ. Biol.,* **44**, 389–398 (2023).

Wang, Y., M. Xie, C. Zhang and J. Li: Genome-wide identification of heat shock proteins and HSP interactors in rice. *BMC Genomics,* **15**, 344 (2014).

Wang, L., P. Zhao, M. Zhang and X. Chen: Genome-wide identification and expression profiling of heat shock proteins in plants under abiotic stress. *BMC Genomics,* **22**, 453 (2021).

Wang, W., B. Vinocur, O. Shoseyov and A. Altman: Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trend. Plant Sci.*, **9**, 244–252 (2004).

Waterhouse, A.M., M. Bertoni and S. Bienert: SWISS-MODEL: Homology modeling of protein structures and complexes. *Nucl. Acids Res.,* **46**, 296–303 (2018).

Wiederstein, M. and M.J. Sippl: The significance of β-strands in protein folding and stability. *FEBS Lett.,* **594**, 1847–1858 (2020).

Xu, D. and Y. Zhang: ModRefiner: A refined method for protein structure model refinement. *Protein Sci.,* **20**, 1424–1437 (2015).

Xue, G.P., S. Sadat, J. Drenth and C.L. McIntyre: The heat shock factor family from Triticum aestivum and their role in regulation of heat shock protein genes. *J. Exp. Bot.,* **65**, 539–557 (2014).

Yang, J., R. Yan and A. Roy: The I-TASSER suite: Protein structure and function prediction. *Nat. Methods,* **12**, 7–8 (2015).

Zhang, L., F. Liu, L. Wang and X. Zhang: Genome-wide analysis and expression profiling under heat and drought treatments of HSP70 gene family in soybean. *Front. Plant Sci.,* **6**, 773 (2015).

Zhong, L., W. Zhou, H. Wang, S. Ding, Q. Lu and X. Wen: Chloroplast small heat shock protein HSP21 interacts with plastid nucleoid protein pTAC5 and is essential for chloroplast development in *Arabidopsis. Plant Physiol.,* **158**, 1009–1020 (2012).