Defence Life Science Journal, Vol. 9, No. 1, January 2024, pp. 44-54, DOI : 10.14429/dlsj.9.19455
© 2024, DESIDOC

SUMO Sites Prediction in Human Transcription Factors Involved in Hypoxia-induced Cardiac Illnesses

Dolly Sharma, Anu Kumari, Manisha Kumari and Ramesh C. Meena*

DRDO- Defence Institute of Physiology and Allied Sciences (DIPAS), Delhi–110 054, India
**E- mail: rcmeena.dipas@gov.in*

ABSTRACT

Protein SUMOylation is a reversible and well known post-translational modification process of the cells. It may change a protein's cellular location, interactions, and possible structural shape before it develops to carry out its basic functions. Also, it decides the binding of transcription factors and DNA binding proteins to chromatin in addition to various cis and trans regulatory factors. Alterations in protein SUMOylation have been linked with a variety of disorders and developmental anomalies. Tentative approaches to identify SUMO binding sites are challenging due to dynamic nature of the SUMOylation process and various critical lab experiments which are involved very high cost. Therefore, the computational methodologies may guide the experimental identification of SUMOylation sites and provide insights for improving comprehension of SUMOylation mechanism in the cells. In this study, we identify the SUMO binding sites in transcription factors that are actively involved and have crucial roles in cardiac development and pathophysiology of the heart. A list of important transcription factors was prepared from the human transcription factor database. The GPS-SUMO, SUMO plot, and JASSA web servers were used for the prediction of SUMO binding sites in cardiac transcription factors. We identified the SUMOylation of several novel, previously uncharacterized SUMO targets that are actively involved in the cardiovascular system. Thus, the present study may help to uncover the significance of SUMO modification in cardiac development and illnesses which creates a fresh avenue for future studies on target-specific SUMOylation for identification of novel therapeutic targets and management strategies for hypoxia-induced cardiovascular disorders.

Keywords: SUMOylation; Transcription factors; SUMO binding sites Cardiovascular; Hypoxia

1. INTRODUCTION

In response to extreme environmental conditions, imperative cellular processes such as regulation of cell cycle, apoptotic cell death, gene and protein expression, protein homeostasis, and sub-cellular transport through bio-molecular interactions are synchronised by the attachment of SUMO proteins. SUMO, a small ubiquitin-related modifiers, a family of low molecular weight proteins has five different types in humans (SUMO 1-5), which are structurally similar to ubiquitin and share 18-20 per cent sequence homology. The SUMO 1 protein is the foremost protein discovered from this family, initially as a covalent modification of Ran GTPase-activating protein and known as GMP 1, UBL 1, PIC 1, and Sentrin[1, 2]. The molecular weight of SUMO protein ranges from 10 to 15 kDa. SUMO 1 shares 55 percent similarity with SUMO 2/3 while SUMO 2 and SUMO 3 are more than 95 percent similar with each other. SUMO 1-3 shows a broad expression pattern in various tissues while SUMO 4 is expressed mainly in the kidney, spleen, and lymph nodes, and SUMO 5 protein is expressed in testes and peripheral blood leukocytes. SUMO 5 also showed its existence in the lungs, placenta, liver, spleen, and thymus. The attachment of SUMO proteins to their particular location is necessary for their optimal functioning in the cells and tissues. The exact location of proteins is mandatory for their attachment with SUMO proteins and best effective working in different cells and tissue types.

SUMOylation is a post-translational modifications (PTM) process of proteins in which attachment of SUMO proteins at canonical sequence "ψ K x E/D" in which ψ represents the hydrophobic amino acid, x can be any amino acid. The attachment of SUMO protein requires its activation where C-terminal diglycine residues are charged by a SAE1/SAE2 heterodimeric enzyme (E 1 activating enzyme) in ATP dependent process. The E 2 coupling enzyme UBC 9 accept the charged SUMO protein through glycyl thioester intermediate and catalyses its attachment with the E 3 ligase enzyme. There are many E 3 ligase enzymes discovered to date that catalyses the attachment of SUMO with lysine residue of consensus sequences present in the target proteins. The attachment of SUMO with the target protein leads to alteration in the downstream pathways, processes, and eventually overall cellular performance. It occurs in cardiomyocytes, fibroblasts, endothelial cells, and smooth muscle cells of the heart. It is vital for development of the heart and its finest functioning[3].

During hypoxia, SUMOylation of HIF1-α at its

Received : 31 March 2022, Revised : 06 September 2022
Accepted : 10 April 2023, Online published : 01 January 2024

different SUMO binding sites leads to enhanced or lowered stability. SUMO 1 conjugation to proteins expressed in the heart is increased during ischemia/reperfusion and protects it from oxidative damage. During myocardial infarction, Erk 5 protein gets SUMOylated by SUMO 1 which promotes inflammation and worsening the cardiac injury. SENP 2 deSUMOylate the Erk 5 protein and protect endothelial dysfunction partially[4]. SUMOylation promotes the degradation of HDACs and reduces the production of reactive oxygen species. SENP 5 is involved in the maintenance of mitochondrial dynamics by removing SUMO moieties from mitochondrial membrane protein DRP 1. During hypertrophied and heart failure, deSUMOylation of DRP 1 shows the way to mitochondrial fission followed by elevation of ROS levels, oxidative stress in cardiomyocytes, maladaptive ventricular remodelling, and a progressive decline in cardiovascular functions. Differentiation, contraction coupling, excitation, hypertrophy, and stress response functions are some of the redox-regulated pathways associated with cardiac hypertrophy[5]. SENP 5 recapitulate the process of dilated cardiomyopathy and heart failure. Recent observations suggest that the cardiac dysfunction can be rescued by increasing the expression of Bcl 2. Defective SUMO conjugation of lamin A protein alters its nuclear distribution and is directly implicated in the initiation of dilated cardiomyopathy.

In the past, molecular biologists conducted intricate experiments using expensive ingredients for identification of protein post-translational modifications such as SUMOylation. However, the advent of bioinformatic tools in recent years has enabled investigators to predict protein PTMs by integrating informatics, mathematics, and statistics. The GPS-SUMO, SUMO plot, and JASSA (Joint analyser of SUMOylation site and SIMs) are well-known tools for the prediction of SUMO binding sites. GPS-SUMO employs the advanced fourth-generation GPS algorithm to anticipate potential SUMOylation sites and SIMs[6]. While the SUMO-plot analysis program predicts and scores SUMOylation sites in the important proteins (https://www.abcepta.com/sumoplot). Similarly, JASSA offers an easily navigable interface that is available online at no cost, allowing users to identify SUMO proteins effortlessly[7]. SUMO proteins have been observed to influence the activity of crucial factors in cardiac development, including GATA-binding protein 4 (GATA 4), serum responsive factor, myocardin, myocyte enhancer factor 2 (MEF 2), as well as the T-box transcription factors 2 and 5 (TBX 2 and TBX 5)[8]. The involvement of SUMOylation in the cardiovascular system is becoming apparent, yet detailed information about the specific roles of SUMO isoforms in the heart and their precise cellular and molecular targets remains limited.

In the present investigation, we have revealed previously unknown SUMOylation sites of vital transcription factors (TFs) that play distinct roles in the cardiovascular system. The SUMO binding site prediction was done using GPS-SUMO, SUMO-Plot, and JASSA web servers. This research introduces a fresh avenue for future investigations into the specifics of SUMOylation at the individual target level. If specific SUMOylation sites on the particular proteins remain unidentified, various mutagenesis strategies can be implemented. Investigating the functional consequences of mutations in the identified protein SUMOylation sites along with exploring additional features of compromised SUMOylation within cardiac cells that can be further undertaken by *in-vitro* and *in-vivo* model systems.

2. METHODOLOGY

2.1. Data Collection of Human Transcription Factors

The transcription factors catalogue generated by lambert et al was used for the present study that comprises profound motif collections such as TRANSFAC[9], UniPROBE[10], JASPAR[11], HT-SELEX[12-14] along with previous catalogues of human TFs[15-17]. The evolutionary paths, expression profiles, and distinct roles of major groups of human TFs exhibit significant differences. The TF's used for the study are selected based on their assessment, classification, experimental derivation, and conformational binding sites and motifs. In addition to this, they are also sorted considering the human cardiovascular physiology, diseases, and variations to understand the TFs-mediated gene regulation of SUMOylated proteins[18].

2.2. Data Retrieval of Transcription Factors

Protein sequences of selected TFs that are involved in cardiac diseases and/or functions were obtained from the protein database of the National Centre for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/). Only one isoform of each transcription factor was used for the identification of SUMO binding sites (Table 2).

2.3 Prediction of SUMOylation Sites

Prediction tools for SUMOylation assess the chemical characteristics and consensus motifs within protein sequences or employ specialized algorithms to anticipate SUMOylation sites. The web servers GPS-SUMO, SUMO Plot, and JASSA are available at https://sumo.biocuckoo.cn/, https://www.abcepta.com/sumoplot, and https://www.jagga.fr respectively were used in the current report by a simple user friendly protocol for the identification of SUMOylation sites in addition to SUMO-interaction motifs (SIMs). In brief, the peptide sequences of TFs in FASTA format or Protein ID (SwissProt, PIR, GenPept, and RefSeq) were submitted in the portal and obtained all the SUMO binding sites and SIM's. The high and medium interaction parameters were used in GPS-SUMO. The default parameters were analysed on SUMO plot and JASSA web servers.

3. RESULTS AND DISCUSSION

3.1. Transcription Factors and Their Decisive Role in Hypoxia-induced Cardiac Illnesses

To meet the heightened energy requirement of the heart, a continuous and ample supply of oxygen is imperative. However, instances of hypoxia creates

an uneven balance between the delivery and demand of oxygen, thereby presenting the heart with various challenges of generating a comparable quantity of ATP using limited oxygen for its contracting efforts, all to uphold regular cardiac functionality. High-altitude (HA) populations reside within an environment characterised by both low barometric pressure and reduced oxygen levels, resulting in decreased oxygen saturation in the alveoli and subjected to long-term effect of persistent hypobaric hypoxia. From an epidemiological perspective, hypoxia holds significant importance as a key risk factor for a large number of populations. It does not only stand as a salient characteristic among high-altitude inhabitants but is also intricately involved in the underlying mechanisms of numerous cardiac ailments. These encompass ischemic heart disease, cardiac hypertrophy, hypertension, atherosclerosis, and heart failure (Fig 1).

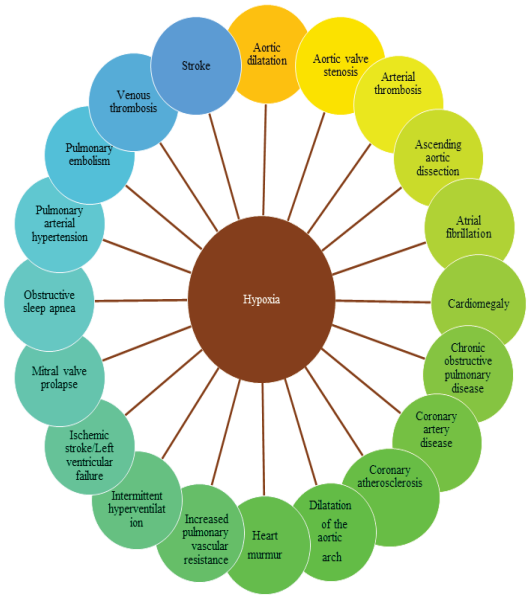

Figure 1. Schematic representation of hypoxia-induced cardiovascular diseases. The figure specifies the important hypoxia-induced cardiovascular abnormalities.

Transcription factors signify the point of conjunction of numerous signaling pathways within eukaryotic cells. TF's have major therapeutic potential since, dysregulation of transcription factors plays significant role in the development of various human diseases, including diabetes, inflammatory disorders, cardiovascular disease, and numerous malignancies[19]. The TF catalogue contains 2765 factors which are filtered for parameters that have the transcriptional activity, motif type, polymer subunit, purification, and cardiac disease specificity. 1639 factors have known transcriptional activity in 1211 with identified and inferred motifs. Out of which, 1205 are monomer or homomultimer and obligate heteromeric. 1063 are tested by HT-Selex while 526 are tested by PBM. A total 363 TFs are defining roles and 284 have been considered as TF class. The total of 37 TFs were chosen for cardiac-specific functions and have significant roles in cardiovascular diseases (Table 1). These TFs show a significant expression pattern in heart muscle

Table 1. Summary of hypoxia induced cardiovascular diseases and transcription factor genes

S. No.	Cardiac diseases	Transcription factor genes
1.	Aortic dilatation	SKI, SMAD4
2.	Aortic valve stenosis	FLI1, FOXF1, GATA5, NKX2-5, NR2F2, SMAD4, WT1
3.	Arterial thrombosis	STAT4, TET2, TP53
4.	Ascending aortic dissection	FOXE3, GATA5, NKX2-5, SMAD3
5.	Atrial fibrillation	NKX2-5, SMAD3, TBX5
6.	Cardiomegaly	FOXE3, GTF2I, GTF2IRD1, PLAGL1, SMAD3, ZIC3
7.	Chronic obstructive pulmonary disease	NFKB1, RREB1, TBX1
8.	Coronary artery disease	FOXE3, PPARG, SMAD3, ZNF687
9.	Coronary atherosclerosis	ESR1
10.	Dilatation of the aortic arch	GATA5, GTF2I, GTF2IRD1, NKX2-5, PRDM16, SKI
11.	Heart murmur	GATA5, NKX2-5
12.	Increased pulmonary vascular resistance	SMAD9
13.	Intermittent hyperventilation	MECP2, TCF4
14.	Ischemic stroke/Left ventricular failure	FOXE3, SMAD3
15.	Mitral valve prolapses	GTF2I, GTF2IRD1, MLXIPL, PRDM5, SKI, SMAD3, SMAD4, ZNF469
16.	Obstructive sleep apnea	AHDC1, NFIX, SKI
17.	Pulmonary arterial hypertension	FOS, FOXF1, GATA6, IRF5, MLX, NFIX, PPARG, RBPJ, SMAD4, SMAD9, STAT1, TBX2, TBX4
18.	Pulmonary embolism	SMAD4, STAT4, TET2
19.	Stroke	GTF2I, GTF2IRD1, MLXIPL, NR3C1, SMAD4, TET2
20.	Venous thrombosis	SMAD4, STAT4, TET2, TP53

Table 2. The expression values of transcription factor genes with the NCBI accession numbers and details of their family

S. No.	Transcription factors	Family	Expression values in heart muscle	Accession number
1.	AHDC1	Uncharacterized	7.56616	NP_001358857.1
2.	ESR1	ERF (ethylene response factor) subfamily B-1 of ERF/AP2 transcription factor family	0.56812	NP_000116.2
3.	FLI1	Fos family	7.95794	NP_002008.2
4.	FOS	Methyl-CPG-Binding	2.99425	CAG47063.1
5.	FOXF1	Basic helix-loop-helix leucine-zipper family	0.47495	NP_001442.2
6.	GATA5	Nuclear Factor I (NFI) family	5.62837	NP_536721.1
7.	GATA6	NF-κB family	25.5991	NP_005248.2
8.	GTF2I	Nuclear Receptor Subfamily 3 Group C Member 1	5.66745	NP_127492.1
9.	GTF2IRD1	Zinc-finger transcription factor	6.66505	NP_001397817.1
10.	IRF5	ligand-activated transcription factors of nuclear hormone receptor superfamily	1.55403	NP_001229381.1
11.	MECP2	(PR) domain subfamily of the Kruppel-like zinc finger protein family	7.77604	NP_004983.1
12.	MLX	Zinc finger transcription factor	13.2745	NP_937848.1
13.	MLXIPL	SMAD family of transcription factor	0.74514	NP_116569.1
14.	NFIX	STAT family	9.07706	NP_001257972.1
15.	NFKB1	Signal transducer and activator of transcription (STAT) family	4.17856	NP_003989.2
16.	NKX2-5	Tbx subfamily of T-box transcription factors	303.449	NP_004378.1
17.	NR2F2	p53 transcription factor family	4.77013	NP_066285.1
18.	NR3C1	E26 transforming-specific family	7.44016	NP_000167.1
19.	PLAGL1	Forkhead family	4.24471	NP_001074424.1
20.	PPARG	The GATA family	3.88175	NP_619725.3
21.	PRDM16	The GATA family	5.93261	NP_071397.3
22.	PRDM5	Transcription factor IIi	5.81658	NP_061169.2
23.	RBPJ	TFII-I family member	5.35937	NP_005340.2
24.	RREB1	Type I interferon system	4.6934	NP_001003699.1
25.	SKI	Basic helix-loop-helix leucine zipper transcription factor of the Myc/Max/Mad superfamily	7.96502	NP_003027.1
26.	SMAD3	NK2 family of homeobox genes	5.72309	NP_005893.1
27.	SMAD4	Steroid thyroid hormone superfamily of nuclear receptors	5.86328	NP_005350.1
28.	SMAD9	Zinc finger transcription factor	6.75405	NP_001120689.1

29.	STAT1	Transcriptional regulator	4.63734	NP_009330.1
30.	STAT4	Smad family	27.6431	NP_003142.1
31.	TBX1	Smad family	0.51771	NP_542377.1
32.	TBX2	Smad family	6.29704	NP_005985.3
33.	TBX5	T-box family	152.04	NP_852259.1
34.	TCF4	T-box family	9.67419	NP_001077431.1
35.	TET2	Helix–loop–helix (HLH) family	3.95656	NP_001120680.1
36.	TP53	Ten-eleven translocation (TET) family proteins	3.06442	NP_000537.3
37.	WT1	zinc finger transcription factor	1.68544	NP_000369.4

and cardiomyocytes (Table 2). The amino acid sequences of these transcription factors in the FASTA format were obtained from NCBI (Table 2).

Details of transcription factors involved in hypoxia-induced cardiovascular diseases[18].

Summary of expression values of transcription factor genes involved in cardiovascular diseases in heart muscle that was used for the analysis of SUMO binding sites[18] with the NCBI accession numbers and their family.

3.2. SUMO Sites Prediction of Key Cardiac Transcription Factors

Extreme environmental stress, such as hypoxia, can significantly alter protein function and their destiny through SUMO modification. The implications of this attachment are evident across various context, ranging from development to the progression of various disease conditions. The SUMOylation process has diverse role in a lot of dynamic biological functions/activities including signal transduction, epigenetic modulation, cell cycle progression, DNA replication, damage, and repair. It has also probable consequences in pathogenesis of human diseases e.g., neurodegenerative syndromes, cancer progression and metastasis, and craniofacial deformities. It is widely recognized that the activation of tissue-specific genes in the heart is controlled by a set of transcription factors. These factors collaborate with various signal transduction pathways and co-factors to regulate their functions. Therefore, to comprehend how the SUMO conjugation pathway regulates gene activity in the heart, it is crucial to investigate whether SUMO targets transcription factors essential for cardiomyocyte differentiation and/or proper heart development.

In the current study, GPS-SUMO, SUMO plot, and JASSA were used for prediction of SUMO binding sites in different TF's. Using GPS-SUMO, we detected SUMOylation sites in amino acid sequences of proteins, including both those conforming to the consensus pattern and deviated from it. The outcomes of the predictions are displayed in a table format, including details such as the FASTA title, the modified peptide and its position, the predicted score, the prediction cut-off, and the type of modification (Fig 2a, S file 1)[6]. Similarly, the SUMO plot scrutiny database forecasts the likelihood for the SUMO consensus sequence to be promised for the joining of SUMO moieties with the target proteins (Fig 2b, S file 2). The scoring system relies on two criteria: the direct matching of amino acids to SUMO-CS and the substitution of consensus amino acid residues with others that share similar hydrophobic properties (source: https://www.abcepta.com/sumoplot). Also, JASSA is examined using a scoring system derived from a position frequency matrix generated through the alignment of potential SUMOylation sites or its interacting motifs (Fig 2c, S file 3)[7]. The aforementioned online tools were employed to identify several transcription factors targeted by SUMOylation. These factors play a significant role in modifying cardiac gene activity and ensuring normal cardiovascular development.

The analysis of data confirmed that the SUMO binding sites in several novel and previously uncharacterized SUMO target proteins like transcription factors e.g. TET2, PRDM16, SRCAP, NR3C2, GTF2IRD1, GTF2I, GLI2, CTCF, NFKB2, ZNF148, SKI, GLIS3, MITF, TBX1, RBPJ, TBX5, MLXIPL, FLI1 (Fig 4d), TCF4, SMAD4, TBX20, WT1, SNAI2, IRF5, TBX19, FOXF1, PAX6, SMAD3, GATA6, NR2F2, HESX1, NKX2-5, GATA4, MEIS2, SPIB, GATA5, POU2AF1, SOX10, and LHX4. Out of these transcription factors, numerous binding sites and a tentative effect of their SUMO binding in cardiac diseases were analysed in Fig. 2 and discussed here.

Summary of cardiac transcription factors and their binding sites predicted by SUMO-Plot, GPS-SUMO and JASSA web servers.

3.2.1 Estrogen Receptor 1 (ESR 1):

Ubiquitin-conjugating enzyme (E2) primarily acts through nuclear receptors ER-α and ER-β (encoded by ESR 1 and ESR 2 genes) to carry out its biological functions. These receptors collectively contribute to maintaining cardiomyocyte homeostasis and providing cardio protection. Their effects involve diverse molecular mechanisms[20]. ER-α can be modified in several ways to control its

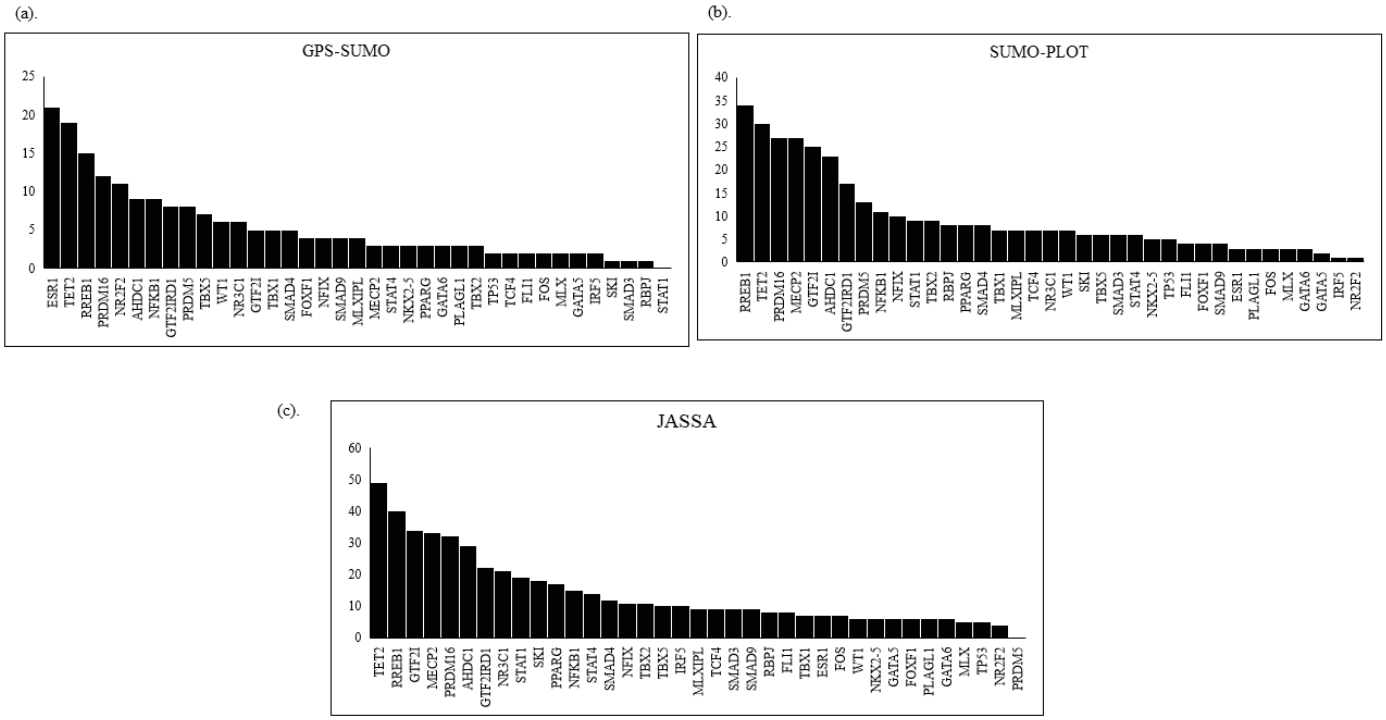

Figure 2. The SUMO binding site prediction of important transcription factors involved in hypoxia-induced cardiovascular illnesses which was analysed by a) SUMO-Plot, b) GPS-SUMO, and c) JASSA web servers.

Table 3. Transcription factors and web servers

S. No.	Transcription factors	SUMO-plot	GPS-SUMO	JASSA
1.	TET2	30	19	49
2.	PRDM16	27	12	32
3.	GTF2IRD1	17	8	22
4.	GTF2I	25	5	34
5.	SKI	6	1	18
6.	TBX1	7	5	7
7.	RBPJ	8	1	8
8.	ESR1	3	21	7
9.	TBX5	6	7	10
10.	MLXIPL	7	4	9
11.	FLI1	4	2	8
12.	TCF4	7	2	9
13.	SMAD4	8	5	12
14.	WT1	7	6	6
15.	IRF5	1	2	10
16.	RREB1	34	15	40
17.	FOXF1	4	4	6
18.	SMAD3	6	1	9
19.	GATA6	3	3	6
20.	NR2F2	1	11	4
21.	NKX2-5	5	3	6
22.	NFKB1	11	9	15
23.	AHDC1	23	9	29
24.	PRDM5	13	8	1
25.	GATA5	2	2	6
26.	NR3C1	7	6	21
27.	NFIX	10	4	11
28.	SMAD9	4	4	9
29.	PLAGL1	3	3	6
30.	PPARG	8	3	17
31.	TBX2	9	3	11
32.	MECP2	27	3	33
33.	STAT4	6	3	14
34.	FOS	3	2	7
35.	MLX	3	2	5
36.	TP53	5	2	5
37.	STAT1	9	-	19

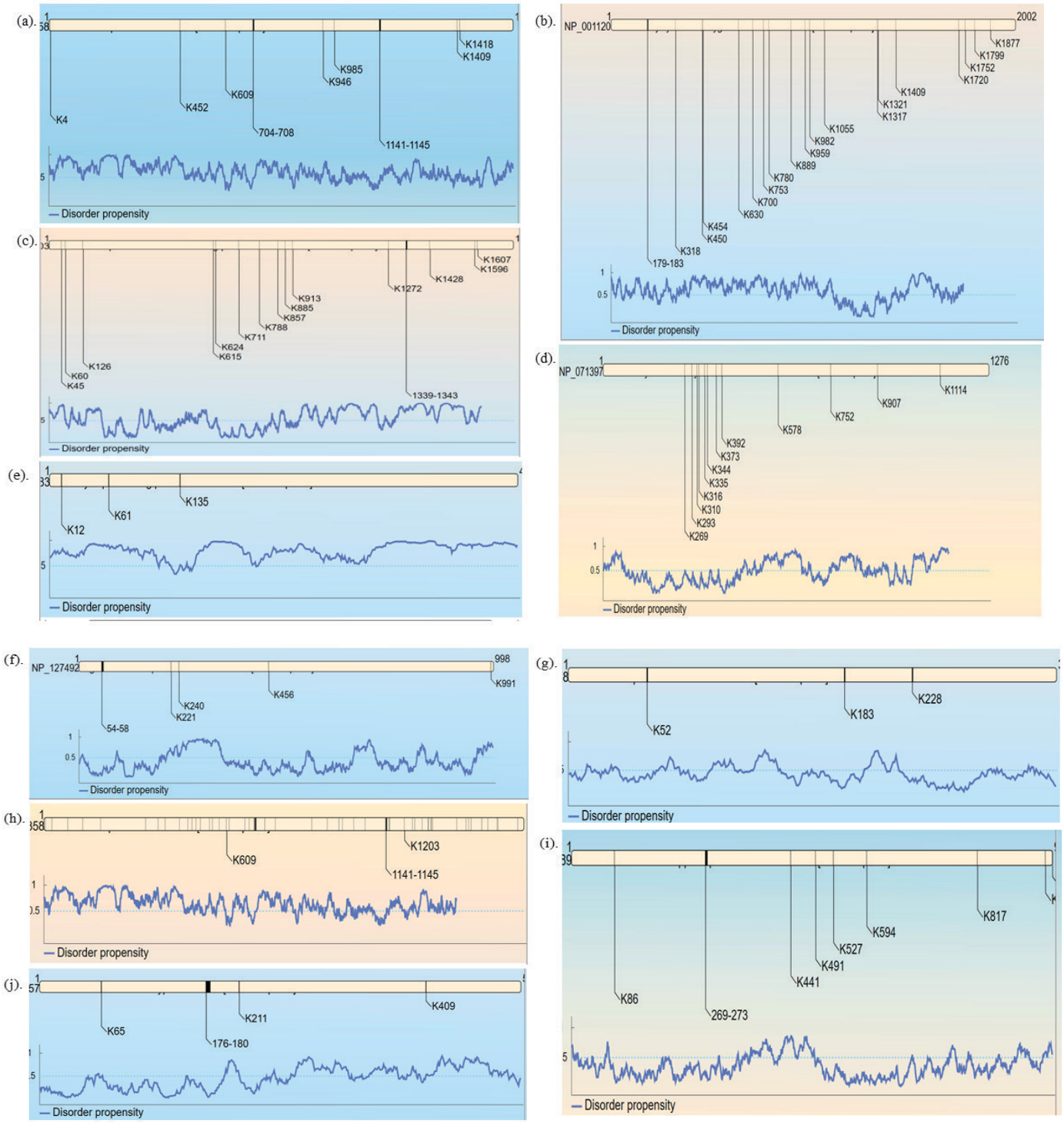

Figure 3. Representation of SUMO binding site prediction using SUMO-plot in a) ESR 1, b) TET 2, c) RREB 1, d) PRDM 16, e) MECP 2, f) NFKB 2, g) GTF2I, h) AHDC1, i) NF-kB 1, and j) NFIX.

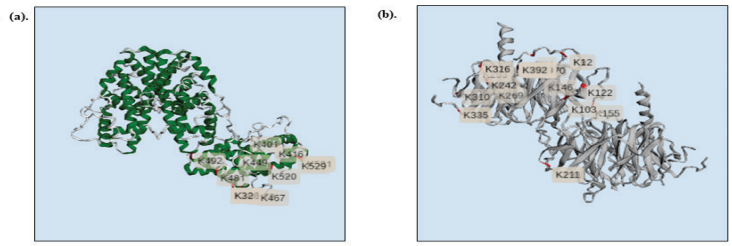

Figure 4. Structure representation in 3D with predicted SUMO binding sites by SUMO-plot, a) ESR 1, b) PRDM 16.

activity. TRIM3 encouraged SUMO alterations of estrogen receptor 1 (ESR 1) and stimulated the ER pathway[21]. A study by Sentis et al. revealed that ER-α is modified by SUMO 1, impacting its transcriptional activity[22]. Here, we also found that ESR 1 has numerous SUMO binding sites and a high probability of SUMOylation (Fig 2, Fig 3a, Fig 4a & Table 3). This can be further explored to understand the SUMOylation of estrogen receptors in cardiovascular disorders.

3.2.2 TET 2 (Ten-Eleven Translocation 2)

TET 2 is a well-recognized epigenetic regulator complex that facilitates transcriptional activation and repression depending on the state of disease condition. It was found to be associated with coronary heart disease, myocardial infarction, facilitates inflammatory responses, and cardiac apoptosis, and regulates the PI3K/AKT pathway[23, 24]. It also alleviates cardiac fibrosis, angiogenesis, action of matrix metalloproteinases (MMPs) and provokes the release of vascular endothelial growth factors[25]. Since, it is involved in oxygenation, TET2 regulates the expression of numerous downstream genes which decides the differentiation of cardiomyocytes and ultimately heart development. Here, we found that TET2 has several SUMO binding sites and a high probability of SUMOylation (Fig 2, Fig 3b & Table 3). This can be further studied using *in-vitro*/*in-vivo* experiments and that may lead to select novel therapeutic targets for the above abnormalities.

3.2.3 Positive Regulatory Domain 16 (PRDM 16)

The PRDM 16 is a vital transcriptional factor that affects cardiac development and regulates myocardial structure to maintain condensed myocardial cardiomyocyte identification by triggering essential solid myocardial genes and repressing trabecular myocardial genes. The loss of PRDM 16 results in a shift in the gene signature linked to the distinctive features of compact cardiomyocytes found in the myocardium of the left ventricle, resembling the characteristics of trabecular cardiomyocytes[26]. Defects in PRDM 16 lead to irregularities in cardiac conduction and trigger phenotypes associated with cardiomyopathies, such as cardiac fibrosis and hypertrophy in cardiomyocytes[27, 28]. The three web servers were used to identify the SUMO sites and found more than 25 SUMO binding sites in PRDM 16 (Fig 2, Fig 3d, Fig 4b & Table 3). To date, the role of SUMO binding to PRDM 16 has not assessed in the cardiovascular system. Although, it was known that SUMOylation of Prdm 16 at lysine 917 by Cbx4 blocks it's ubiquitination-mediated deprivation, thereby enhancing its stability and thermogenic role in adipose tissue[29]. The SUMO modification at lysine residue 568 and CtBP binding are essential in the repression of PRDM 16 mediated transcriptional and differentiation hindrance in the process of leukemogenesis[30].

3.2.4 General Transcription Factor 2I (GTF2I)

Transcription factor TFII-I is a universally expressed multifunctional protein that contains a nuclear localization sequence, basic DNA-binding, and multiple protein-protein interaction domains. TFII-I regulates the transcription of specific genes positively or negatively using initiator elements and cis-regulatory DNA-binding sequences[31]. Additionally, the network with CTCF regulates several gene expression in response to metabolic stress[32]. A study revealed that circ-GTF2I could intensify MIRI and myocardial infarction through induction of irregular expressions of IL-6, TNF-α, LDH, Bax, Bcl 2, and Cyt-c in MIRI[33]. Cardiac diseases associated with GTF2I include Williams-beuren syndrome and supra-valvular aortic stenosis. There is no SUMOylation studies were reported related to GTF2I to date. In our study, we identified more than 25 probable SUMO binding sites using all three web servers (Fig 2, Fig 3g & Table 3). It can be further explored using various study model systems which may lead to several novel therapeutic targets for hypoxia-induced cardiac illnesses.

3.2.5 NK2 Homeobox (Nkx2.5)

It is a cardiac-specific homeobox gene, associated with the *Nk*-2 class of homeodomain (HD) harboring factors, essential for standard cardiac development. Nkx2.5 genes are found to be expressed in early cardiac progenitor cells before cardiogenic differentiation. It is additionally subjected to SUMOylation at its conserved primary site, Lys51, which remains consistent across various species throughout the evolution. The attachment of Nkx2.5 with its associated proteins is improved by SUMO modification, which contributes to its interaction with the cofactor and improves Nkx2.5 function. NKX2.5 is majorly modified by SUMO 1 and results showed its enhanced transcriptional activity. The process of SUMOylation in NKX2.5 is crucial for typical heart development, while mutations in NKX2.5 are associated with the development of various congenital heart defects[3]. In our study, we identified more than 5 probable SUMO binding sites using all the above servers (Fig. 2 & Table 3). It can be further explored in mammalian models which may lead to therapeutic target for hypoxia-induced cardiac illnesses.

4. CONCLUSION

In summary, the transcription factors TET 2, PRDM 16, GTF2I, and NKX2.5 play crucial roles in various aspects of cardiovascular system development and functions. These transcription factors collectively orchestrated the intricate processes of cardiac development, maintenance, and response to different stressor including hypoxia. The identified SUMOylation binding sites within these TFs suggested the involvement of post-translational modifications in shaping their activities and proper functioning throughout the life including developmental stages of the heart. The present study, identified SUMO binding sites in above TFs which would be very crucial for the identification of biomarkers for the design and development of drugs/therapeutics for various pathological as well as hypoxia induced cardiovascular illnesses.

ACKNOWLEDGEMENTS

The authors are thankful to Ms. Chhavi Rai, Technical Officer 'A', Ms. Harshita Gupta, Technician 'B' and Ms. Reena, Project trainee, Department of Pathophysiology and Disruptive Technologies, DIPAS-DRDO for their constant technical and logistic support during the course of this study.

FUNDING INFORMATION

Authors are grateful to the funding agency, Director, Defence Institute of Physiology and Allied Sciences (DIPAS), Defence Research and Development Organization (DRDO), Ministry of Defence, Government of India. DS is supported by a fellowship from DIPAS-DRDO, Ministry of Defence, Government of India and MK is supported by a fellowship from Council of Scientific and Industrial Research-University Grants Commission (CSIR-UGC), Ministry of Human Resources, Government of India.

REFERENCES

1. Matunis, M.J.; Coutavas, E. & Blobel, G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J. Cell Biol.*, 1996, **135**(6 Pt 1), 1457-1470. doi: 10.1083/jcb.135.6.1457.
2. Mahajan, R.; Delphin, C.; Guan, T.; Gerace, L. & Melchior, F. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 1997, **88**(1), 97-107. doi: 10.1016/s0092-8674(00)81862-0.
3. Wang, J. & Schwartz, R.J. Sumoylation and regulation of cardiac gene expression. *Circ. Res.*, 2010, **107** (1), 19-29. doi: 10.1161/CIRCRESAHA.110.220491.
4. Shishido, T.; Woo, C-H.; Ding, B.; McClain, C.; Molina, C. A.; Yan, C.; Yang, J.; Abe, J-I. Effects of MEK5/ERK5 association on small ubiquitin-related modification of ERK5: Implications for diabetic ventricular dysfunction after myocardial infarction. *Circ. Res.*, 2008, **102**(11), 1416-1425. doi: 10.1161/CIRCRESAHA.107.168138.
5. Jin, J-Y.; Wei, X-X.; Zhi, X-L.; Wang, X-H. & Meng, D. Drp1-dependent mitochondrial fission in cardiovascular disease. *Acta Pharmacol. Sin.*, 2021, **42**(5), 655-664. doi: 10.1038/s41401-020-00518-y.
6. Zhao, Q.; Xie, Y.; Zheng, Y.; Jiang, S.; Liu, W.; Mu, W.; Liu, Z.; Zhao, Y.; Xue, Y. & Ren, J. GPS-SUMO: A tool for the prediction of sumoylation sites and SUMO-Interaction motifs. *Nucleic Acids Res.*, 2014, **42**(Web Server issue), W325-W330. doi: 10.1093/nar/gku383.
7. Beauclair, G.; Bridier-Nahmias, A.; Zagury, J.-F.; Saïb, A. & Zamborlini, A. JASSA: A comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics.,* 2015, **31**(21), 3483-3491. doi: 10.1093/bioinformatics/btv403.
8. Beketaev, I.; Kim, E.Y.; Zhang, Y.; Yu, W.; Qian, L. & Wang, J. Potentiation of tbx5-mediated transactivation by SUMO conjugation and protein inhibitor of activated STAT 1(PIAS1). *Int. J. Biochem. Cell. Biol.* 2014, **50**, 82-92. doi: 10.1016/j.biocel.2014.02.007.
9. Matys, V.; Kel-Margoulis, O.V.; Fricke, E.; Liebich, I.; Land, S.; Barre-Dirrie, A.; Reuter, I.; Chekmenev, D.; Krull, M.; Hornischer, K.; Voss, N.; Stegmaier, P.; Lewicki-Potapov, B.; Saxel, H.; Kel, A. E. & Wingender, E. TRANSFAC and its module TRANSCompel: Transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.*, 2006, **34**(Database issue), D108-110. doi: 10.1093/nar/gkj143.
10. Hume, M.A.; Barrera, L.A.; Gisselbrecht, S.S. & Bulyk, M.L. UniPROBE, Update 2015: New tools and content for the online database of protein-binding microarray data on protein-DNA interactions. *Nucleic Acids Res.*, 2015, **43**(D1), D117-D122. doi: 10.1093/nar/gku1045.
11. Mathelier, A.; Fornes, O.; Arenillas, D.J.; Chen, C.; Denay, G.; Lee, J.; Shi, W.; Shyr, C.; Tan, G.; Worsley-Hunt, R.; Zhang, A. W.; Parcy, F.; Lenhard, B.; Sandelin, A. & Wasserman, W. W. JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.***,** 2016, **44(**D1), D110-D115. doi: 10.1093/nar/gkv1176.
12. Jolma, A.; Yan, J.; Whitington, T.; Toivonen, J.; Nitta, K.R.; Rastas, P.; Morgunova, E.; Enge, M.; Taipale, M.; Wei, G.; Palin, K.; Vaquerizas, J.M.; Vincentelli, R.; Luscombe, N.M.; Hughes, T.R.; Lemaire, P.; Ukkonen, E.; Kivioja, T. & Taipale, J. DNA-binding specificities of human transcription factors. *Cell*, 2013, **152**(1-2), 327-339. doi: 10.1016/j.cell.2012.12.009.
13. Jolma, A.; Yin, Y.; Nitta, K.R.; Dave, K.; Popov, A.; Taipale, M.; Enge, M.; Kivioja, T.; Morgunova, E. & Taipale, J. DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature*, 2015, **527**(7578), 384-388. doi: 10.1038/nature15518.
14. Pantier, R.; Chhatbar, K.; Alston, G.; Lee, H.Y. & Bird, A. High-throughput sequencing SELEX for the determination of DNA-binding protein specificities in Vitro. *STAR Protoc.*, 2022, **3**(3), 101490. doi: 10.1016/j.xpro.2022.101490.
15. Fulton, D.L.; Sundararajan, S.; Badis, G.; Hughes, T.R.; Wasserman, W.W.; Roach, J.C. & Sladek, R. TFCat: the curated catalog of mouse and human transcription factors. *Genome Biol.*, 2009, **10** (3), R29. doi: 10.1186/gb-2009-10-3-r29.
16. Vaquerizas, J.M.; Kummerfeld, S.K.; Teichmann, S.A. & Luscombe, N.M. A Census of human transcription factors: Function, expression and evolution. *Nat. Rev. Genet.*, 2009, **10**(4), 252-263. doi: 10.1038/nrg2538.
17. Wingender, E.; Schoeps, T.; Haubrock, M.; Krull, M. & Dönitz, J. TFClass: Expanding the classification of human transcription factors to their mammalian orthologs. *Nucleic Acids Res.*, 2018, **46** (D1), D343-D347. doi: 10.1093/nar/gkx987.

18. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R. & Weirauch, M. T. The human transcription factors. *Cell*., 2018, **172**(4), 650-665. doi: 10.1016/j.cell.2018.01.029.
19. Papavassiliou, K.A. & Papavassiliou, A.G. Transcription factor drug targets. *J. Cell. Biochem*., 2016, **117**(12), 2693-2696. doi: 10.1002/jcb.25605.
20. Puzianowska-Kuźnicka, M. ESR1 in myocardial infarction. *Clin. Chim. Acta Int.*, 2012, **413**(1-2), 81-87. doi: 10.1016/j.cca.2011.10.028.
21. Ye, R.; AiErken, N.; Kuang, X.; Zeng, H.; Shao, N.; Lin, Y.; Liu, P. & Wang, S. Tripartite motif-containing 3 (TRIM3) enhances ER signaling and confers tamoxifen resistance in breast cancer. *Oncogenesis*., 2021, **10** (9), 60. doi: 10.1038/s41389-021-00350-x.
22. Nait Achour, T.; Sentis, S.; Teyssier, C.; Philippat, A.; Lucas, A.; Corbo, L.; Cavaillès, V. & Jalaguier, S. Transcriptional repression of estrogen receptor α signaling by SENP2 in breast cancer cells. *Mol. Endocrinol*., 2014, **28**(2), 183-196. doi: 10.1210/me.2013-1376.
23. Sano, S.; Oshima, K.; Wang, Y.; MacLauchlan, S.; Katanasaka, Y.; Sano, M.; Zuriaga, M.A.; Yoshiyama, M.; Goukassian, D.; Cooper, M.A.; Fuster, J.J. & Walsh, K. Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1β/NLRP3 inflammasome. *J. Am. Coll. Cardiol*., 2018, **71** (8), 875-886. doi: 10.1016/j.jacc.2017.12.037.
24. Xu, Y. TET2 expedites coronary heart disease by promoting microRNA-126 expression and inhibiting the E2F3-PI3K-AKT axis. *Biochem. Cell Biol.*, 2020, **98**(6), 698-708. doi: 10.1139/bcb-2020-0297.
25. Zadeh, F.J.; Akbari, T.; Zayeri, Z.D.; Samimi, A.; Davari, N. & Rezaeeyan, H. The role of molecular mechanism of ten-eleven translocation2 (TET2) family proteins in pathogenesis of cardiovascular diseases (CVDs). *Mol. Biol. Rep*., 2020, **47**(7), 5503-5509. doi: 10.1007/s11033-020-05602-4.
26. Wu, T.; Liang, Z.; Zhang, Z.; Liu, C.; Zhang, L.; Gu, Y.; Peterson, K.L.; Evans, S.M.; Fu, X.-D. & Chen, J. PRDM16 is a compact myocardium-enriched transcription factor required to maintain compact myocardial cardiomyocyte identity in left ventricle. *Circulation*., 2022, **145**(8), 586-602. doi: 10.1161/CIRCULATIONAHA.121.056666.
27. Arndt, A.-K.; Schafer, S.; Drenckhahn, J.-D.; Sabeh, M.K.; Plovie, E.R.; Caliebe, A.; Klopocki, E.; Musso, G.; Werdich, A.A.; Kalwa, H.; Heinig, M.; Padera, R. F.; Wassilew, K.; Bluhm, J.; Harnack, C.; Martitz, J.; Barton, P.J.; Greutmann, M.; Berger, F.; Hubner, N.; Siebert, R.; Kramer, H.-H.; Cook, S.A.; MacRae, C.A. & Klaassen, S. Fine mapping of the 1p36 deletion syndrome identifies mutation of PRDM16 as a cause of cardiomyopathy. *Am. J. Hum. Genet.,* 2013, **93**(1), 67-77. doi: 10.1016/j.ajhg.2013.05.015.
28. Nam, J.M.; Lim, J.E.; Ha, T.W.; Oh, B. & Kang, J.-O. Cardiac-specific inactivation of Prdm16 effects cardiac conduction abnormalities and cardiomyopathy-associated phenotypes. *Am. J. Physiol. Heart Circ. Physiol*., 2020, **318**(4), H764-H777. doi: 10.1152/ajpheart.00647.2019.
29. Chen, Q.; Huang, L.; Pan, D.; Zhu, L. J. & Wang, Y-X. Cbx4 sumoylates prdm16 to regulate adipose tissue thermogenesis. *Cell Rep.,* 2018, **22**(11), 2860-2872. doi: 10.1016/j.celrep.2018.02.057.
30. Nishikata, I.; Nakahata, S.; Saito, Y.; Kaneda, K.; Ichihara, E.; Yamakawa, N. & Morishita, K. Sumoylation of MEL1S at lysine 568 and its interaction with CtBP facilitates its repressor activity and the blockade of G-CSF-induced myeloid differentiation. *Oncogene*., 2011, **30**(40), 4194-4207. doi: 10.1038/onc.2011.132.
31. Roy, A.L. Biochemistry and biology of the inducible multifunctional transcription factor TFII-I. *Gene*., 2001, **274**(1-2), 1-13. doi: 10.1016/s0378-1119(01)00625-4.
32. Peña-Hernández, R.; Marques, M.; Hilmi, K.; Zhao, T.; Saad, A.; Alaoui-Jamali, M.A.; del Rincon, S.V.; Ashworth, T.; Roy, A.L.; Emerson, B.M. & Witcher, M. Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I. *Proc. Natl. Acad. Sci. U. S. A*., 2015, **112**(7), E677-686. doi: 10.1073/pnas.1416674112.
33. Yuan, C.; Lu, J.; Chen, Z. & Zhou, Q. Circ-GTF2I/MiR-590-5p axis aggravates myocardial ischemia-reperfusion injury by regulating kelch repeat and BTB domain-containing protein 7. *Evid. Based Complement. Alternat. Med*., 2022, **2022**, 2327669. doi: 10.1155/2022/2327669.

CONTRIBUTORS

Ms Dolly Sharma has received her MSc (Biotechnology) from Guru Jambheshwar University of Science and Technology, Hisar, Haryana. Currently, she is working as Senior Research Fellow at Defence Institute of Physiology and Allied Sciences, Delhi. Her research work mainly focused on SUMOylation of transcription factors in human cardiomyocytes exposed to hypoxic stress.
She has helped in conceptualisation, literature search and writing of the manuscript.

Ms Anu Kumari has received her MSc (Biochemistry) from UP Rajshri Tandon Open University, Prayagraj, UP. Currently, she is working as Senior Technical Assistant 'B' at Defence Institute of Physiology and Allied Sciences, Delhi. She helped in data collection, sorting, analysis and representation.

Ms Manisha Kumari has received her MSc (Biotechnology) from Central University of Rajasthan, Ajmer, Rajasthan. Currently, she is working as Senior Research Fellow at Defence Institute of Physiology and Allied Sciences, Delhi. Her research area includes study of gut microbiota composition at extreme environmental conditions including high altitude.
She helped in data collection and literature search.

Dr Rajeev Varshney has received his PhD from Jawaharlal Nehru University, New Delhi. Currently, he is working as Scientist 'G' and Director, Defence Institute of Physiology & Allied Sciences, Delhi. He has successfully contributed towards the development of mechanism-based drugs for countering radiation injuries.
He helped in guiding the research project and reviewing the present manuscript.

Dr Ramesh C. Meena has received his PhD in Life Sciences from Bharathiar University, Coimbatore. Presently, he is working as Scientist 'E' at Defence Institute of Physiology and Allied Sciences, Delhi. His major research area is to elucidate molecular mechanism of stress tolerance in yeast & mammalian model systems and exploration of human microbiome at extreme environmental conditions. He has various publications in national and international journals, book chapters and patents to his credit.
He has contributed in thorough conceptualisation, design, execution, data collection and writing of the manuscript.