

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ÉTUDE DE LA SUMOYLATION DES FACTEURS DE CHOC THERMIQUES IMPLIQUÉS
DANS LA RÉPONSE AU STRESS THERMIQUE CHEZ *ARABIDOPSIS THALIANA*

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN BIOCHIMIE

PAR

LAXIGA SINNATHURAI

MAI 2024

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.12-2023). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens à exprimer ma sincère gratitude à Pr. Laurent Cappadocia, pour son dévouement et son expertise en me guidant durant ma maîtrise. Ses connaissances approfondies, ses commentaires pertinents et ses précieuses suggestions ont joué un rôle central dans l'élaboration des résultats de ma mémoire. Son engagement envers mon succès a été vraiment inspirant, et je suis reconnaissante pour les innombrables heures que vous avez investies dans l'examen de mon travail, en me conseillant et en me poussant à atteindre de nouveaux sommets.

J'aimerais également remercier les membres du Laboratoire Cappadocia, Souleimen, Antoine, Ali, Julien, Mariama et Faustine ainsi que les membres du Laboratoire Joly-Lopez, Fabian et Aurore pour leur constant support.

À ma famille, mes parents, Suseela et Suthakaran, mon frère, Laxan, ma sœur, Laksana et mon fiancé, Thivishan, je ne saurais exprimer assez de gratitude pour leur amour, leurs encouragements et leur patience tout au long de ces deux ans et demi. Leur confiance en mes capacités, même dans les moments de doute, a été une source constante de force. Leur sacrifice, leur compréhension et leur volonté de me soutenir ont rendu ce voyage non seulement enrichissant sur le plan académique, mais aussi émotionnellement. Je suis éternellement reconnaissante de leur présence dans ma vie.

Enfin, je tiens à remercier le corps professoral et le personnel de l'Université du Québec à Montréal pour m'avoir fourni un environnement enrichissant et stimulant pour ma croissance académique. Le soutien, les ressources et les opportunités offertes par l'institution ont été inestimables pour façonner mon travail de mémoire. Je me sens privilégiée d'avoir fait partie de cette communauté universitaire aussi remarquable.

DÉDICACE

À ma mère Suseela Tharumalingam

INTRODUCTION

L'étude des mécanismes moléculaires régissant les réponses des plantes au stress thermique a connu un essor significatif, et au cœur de cette complexité émergent les facteurs de choc thermiques (HSFs) chez *Arabidopsis thaliana*. Ma mémoire se plonge dans l'exploration minutieuse de l'interconnexion fascinante entre les HSFs, le stress thermique, et un processus post-traductionnel clé, la SUMOylation. Le stress thermique représente l'un des défis environnementaux les plus critiques pour les plantes, influençant leur croissance, leur développement et, ultimement, leur survie. Les HSFs, en tant que régulateurs majeurs de la réponse au stress thermique, jouent un rôle crucial dans l'orchestration des réponses moléculaires permettant aux plantes de s'adapter à ces conditions adverses. Cependant, au-delà de cette première couche de complexité, la SUMOylation émerge comme un processus régulateur subtil et sophistiqué. Cette modification post-traductionnelle, impliquant l'ajout de petites protéines SUMO (Small Ubiquitin-like Modifier), est maintenant reconnue comme un acteur potentiellement influent dans la régulation des HSFs. Cette étude vise à démêler ces interactions complexes, à explorer comment la SUMOylation modifie l'activité des HSFs, et à discuter comment cette modification peut influencer la réponse globale de la plante au stress thermique.

L'importance de cette recherche réside dans sa capacité à apporter des éclairages novateurs sur les mécanismes moléculaires sous-jacents à la résistance des plantes au stress thermique. Comprendre comment la SUMOylation influe sur les HSFs peut non seulement élargir notre compréhension fondamentale de la biologie des plantes, mais aussi ouvrir des avenues passionnantes pour des applications pratiques. Ces connaissances pourraient éventuellement être exploitées pour moduler la résistance des cultures aux conditions climatiques changeantes, un enjeu pressant dans le contexte des défis actuels liés au changement climatique.

Enfin, je tiens à exprimer ma gratitude envers ceux qui ont contribué à l'élaboration de cette étude. Leur soutien indéfectible a été essentiel à chaque étape de cette recherche, et je suis optimiste quant aux perspectives nouvelles qu'elle pourrait offrir à la communauté scientifique et au-delà.

TABLE DES MATIÈRES

REMERCIEMENTS	ii
DÉDICACE.....	iii
INTRODUCTION.....	iv
LISTE DES FIGURES.....	vii
LISTE DES TABLEAUX.....	viii
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES	ix
LISTE DES SYMBOLES ET DES UNITÉS	x
RÉSUMÉ.....	xi
ABSTRACT	xii
CHAPITRE 1	1
REVUE DE LA LITTÉRATURE.....	1
1.1 Stress thermique : réponse physiologique et moléculaire.....	1
1.2 Les protéines de choc thermique et les facteurs de choc thermique.....	3
1.3 Modification post-traductionnelle : La SUMOylation	7
1.3.1 Les enzymes E3 ligases.....	9
1.3.2 Les protéines SUMO (Small Ubiquitin-like MOdifier)	14
1.4 La problématique.....	16
1.4 Le motif consensus ψ KxE/D	17
1.5 21 HSFs chez <i>Arabidopsis thaliana</i>	17
1.6 Hypothèse de l'étude	21
1.7 Objectif de l'étude	22
CHAPITRE 2 MATÉRIEL ET MÉTHODES	24
2.1 Bio-informatique.....	24
2.1.1 AlphaFold.....	24
2.1.2 Les logiciels JASSA et GPS-SUMO.....	26
2.2 Principaux réactifs	26

2.3	Clonage moléculaire	27
2.3.1	Amplification par PCR (Polymerase Chain Reaction).....	28
2.3.2	Clonage moléculaire par Gibson.....	31
2.3.3	Transformation bactérienne DH5 α	31
2.3.4	Isolation de l'ADN plasmidique (MiniPreps).....	32
2.4	Production et purification des protéines	34
2.4.1	Transformation bactérienne BL21 DE3 CodonPlus	34
2.4.2	Culture et expression des protéines.....	34
2.4.3	Chromatographie par affinité	35
2.4.4	Clivage à la TEV	36
2.4.5	Tamis moléculaire.....	36
2.4.6	Électrophorèse sur gel SDS-PAGE.....	36
2.5	Essai retard sur gel.....	37
2.6	Essais enzymatiques	39
CHAPITRE 3 RÉSULTATS.....		40
3.1	Analyse bio-informatique des sites de SUMOylation chez les HSFs par les logiciels JASSA et GPS-SUMO.....	40
3.2	Conservation des lysines chez les 21 protéines HSFs	53
3.3	Purification des protéines candidates HSFs.....	57
3.4	Purification de la protéine HSFB2b.....	60
3.5	HSFB2b se lie à la séquence spécifique HSE.....	65
CHAPITRE 4 DISCUSSION.....		71
4.1	Analyse biochimique	71
4.1.1	Purification des protéines HSFs.....	71
4.1.2	HSFB2b : Purification complète	73
4.1.3	HSFB2b : Liaison spécifique à HSE.....	74
4.2	Analyses bio-informatiques : Prédiction des sites de SUMOylation potentiels	80
CONCLUSION		91
BIBLIOGRAPHIE		93

LISTE DES FIGURES

Figure 1.1 Cascade enzymatique de la voie de SUMOylation.....	9
Figure 1.2 Enzyme de la SUMOylation.....	16
Figure 2.1 Visualisation des erreurs alignées prévues (PAE).....	25
Figure 2.1.4 Système de purification d'ADN par Minipreps.....	33
Figure 3.1 Analyse comparative des scores des serveurs web JASSA et GPS-SUMO : Une étude de corrélation pour la prédiction des sites de SUMOylation	443
Figure 3.2 Positionnement des sites de SUMOylation prédits par les logiciels JASSA et GPS-SUMO.	47
Figure 3.3 WebLogo de la conservation des sites de SUMOylation avec la lysine située au site prédit en vert et les autres résidus en gris des HSFs.	53
Figure 3.4 Conservation des résidus lysines dans les 21 HSFs d' <i>Arabidopsis thaliana</i>	55
Figure 3.5 Répartition des lysines SUMOylables chez les HSFs en fonction de leurs classes.....	56
Figure 3.6 Gel SDS-PAGE de la purification des protéines HSFs candidates	60
Figure 3.7 Séquence primaire de la protéine HSFB2b.....	61
Figure 3.8 Exemple de la purification de la protéine HSFB2b	63
Figure 3.9 Retard sur gel de la protéine HSFB2b présence d'ADN spécifique et non-spécifique	66
Figure 3.10 Essai enzymatique de la protéine HSFB2b en présence d'ADN spécifique à SIZ-1 .	67
Figure 3.11 Modélisation de la protéine HSFB2b avec la machinerie de la SUMOylation par AlphaFold	68
Figure 3.12 Conservation de la séquence protéique HSFB2b chez différentes espèces	69

LISTE DES TABLEAUX

Tableau 1.2 Cinq majeures familles de choc thermique et leur fonction sous stress thermique.....	5
Tableau 1.3.2 Différentes protéines SUMO présentent chez <i>Arabidopsis thaliana</i>	14
Tableau 2.1 Plasmides obtenus de chez ABRC contenant la séquence codante pour les différentes protéines HSFs utilisées dans le cadre de cette étude	27
Tableau 2.2 Caractéristiques des protéines utilisées dans cette étude.....	28
Tableau 2.3 Séquences des oligonucléotides employées pour amplifier les ADNc de facteurs de choc thermique.....	29
Tableau 2.4 Séquence nucléotidique de l'élément de réponse au stress thermique (HSE) impliqué dans la liaison et l'activation protéique.....	38
Tableau 3.1 Sites de SUMOylation chez les protéines HSFs prédites au moyen des logiciels GPS-SUMO (https://sumo.biocuckoo.cn/) et JASSA (http://www.jassa.fr).....	41
Tableau 3.2 Motifs d'interactions SUMO (SIM) chez les protéines HSFs prédits au moyen des logiciels GPS-SUMO et JASSA	44
Tableau 3.3 Résumé des clonages/expressions/purifications des protéines HSFs d' <i>Arabidopsis</i> .	59
Tableau 3.4 Exemple de la purification de la protéine HSFB2b.....	64

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

ADN	Acide désoxyribonucléique
HSF	Facteur de choc thermique
HSP	Protéine de choc thermique
MMS21	Methyl Methanesulfonate Sensitivity 21
MPT	Modifications post-traductionnelles
Ni-NTA	Nickel-Nitrilotriacetic Acid
PIAL 1/2	Protein Inhibitor of Activated like-STAT 1/2
SAE1	SUMO-activating enzyme 1 (E1)
SAE2	SUMO-activating enzyme 2 (E1)
SCE1	SUMO-conjugating enzyme 1 (E2)
SIM	SUMO-Interaction Motif
SIZ-1	SAP and Miz 1 E3 ligase
SUMO	Small Ubiquitine-like Modifier

LISTE DES SYMBOLES ET DES UNITÉS

μM	micromolaire
mL	millilitre
L	litre
μL	microlitre
K	lysine
E	acide glutamate
Ψ	psi – résidu hydrophobe

RÉSUMÉ

La SUMOylation est une modification post-traductionnelle des protéines qui permet aux plantes de répondre de façon rapide à des stress environnementaux comme la déshydratation ou l'exposition à des températures extrêmes. La SUMOylation est médiée par une cascade enzymatique impliquant une enzyme activatrice (E1), une enzyme conjugatrice (E2) et une protéine ligase (E3). Cette modification des protéines est typiquement dépendante de la présence des sites de SUMOylation et survient au niveau du motif Ψ -K-x-E où Ψ est un résidu hydrophobe, K la lysine ciblée par la SUMOylation, x un résidu quelconque et E un glutamate. Du fait de leur très petite taille et faible complexité, les sites de SUMOylation peuvent rapidement apparaître et disparaître au cours de l'évolution sous l'effet de simples mutations et de substitutions. Notre hypothèse est que des changements au niveau de l'abondance et l'emplacement des sites de SUMOylation, bien qu'ils puissent contribuer à court terme à des instabilités des réseaux protéiques, contribuent à plus long terme à une adaptation rapide des plantes à de nouveaux environnements. Notre objectif principal consiste à apporter une perspective évolutive à l'analyse comparative de la SUMOylation chez différentes protéines HSFs (Heat Shock Factors) impliquées dans la réponse aux stress thermiques chez *Arabidopsis thaliana*. L'objectif a été atteint en combinant des approches bio-informatiques et des analyses biochimiques. L'approche bio-informatique (JASSA et GPS-SUMO) a permis de mettre en évidence la présence de sites de SUMOylation dans les HSFs de la plante *Arabidopsis*. Parmi les 21 HSFs d'*Arabidopsis*, HSFA1, HSFA1b, HSFA1d, HSFA2, HSFA3, HSFA4c, HSFA5, HSFA6a, HSFA6b, HSFA7a, HSFA7b, HSFA8 et HSFA9 de la classe A, HSFB1, HSFB2b et HSFB3 de la classe B, ainsi que HSFC1 de la classe C, présentent des sites de SUMOylation prédits avec, des scores élevés. Dans une perspective biochimique, certaines des protéines candidates identifiées par les analyses bio-informatiques soit HSFA1a, HSFA2, HSFA1e, HSFA7a, HSFA7b, HSFA9, HSFB2 et HSFB3, ont été exprimées chez des bactéries. La protéine HSFB2b, un facteur de transcription répresseur, a été davantage étudiée et présente une bonne solubilité à une capacité de liaison à l'ADN et SUMOylée en présence d'ADN.

Mots clés : stress thermique, modification post-traductionnelle, SUMOylation, facteurs de choc thermique, motif de SUMOylation (Ψ -K-x-E)

ABSTRACT

SUMOylation is a post-translational modification of proteins that allows plants to respond quickly to environmental stresses such as dehydration or exposure to extreme temperatures. SUMOylation is mediated by an enzymatic cascade involving an activating enzyme (E1), a conjugating enzyme (E2) and a protein ligase (E3). This modification of proteins is typically dependent on the presence of SUMOylation sites and occurs at the level of Ψ -K-x-E motif, where Ψ is a hydrophobic residue, K the lysine targeted by SUMOylation, x any residue and E a glutamate. Due to their very small size and low complexity, SUMOylation sites can quickly appear and disappear during evolution under the effect of simple substitution mutations. Our hypothesis is that changes in the abundance and location of SUMOylation sites, although they may contribute in the short term to instabilities of protein networks, contribute in the longer term to rapid adaptation of plants to new environments. Our main objective is to bring an evolutionary perspective to the comparative analysis of SUMOylation in different HSFs (Heat Shock Factors) proteins involved in the response to heat shock in *Arabidopsis thaliana*. We achieved this goal by combining bioinformatics approaches and biochemical analyses. The bioinformatics approach (JASSA and GPS-SUMO) made it possible to highlight the presence of SUMOylation sites in the HSFs of the Arabidopsis plant. Among the 21 Arabidopsis HSFs, HSFA1, HSFA1b, HSFA1d, HSFA2, HSFA3, HSFA4c, HSFA5, HSFA6a, HSFA6b, HSFA7a, HSFA7b, HSFA8 and HSFA9 from class A, HSFB1, HSFB2b and HSFB3 from class B, as well that HSFC1 from class C, present SUMOylation sites predicted, with high scores. From a biochemical perspective, some of the candidate proteins identified by bioinformatics analyzes were expressed in bacteria, including HSFA1a, HSFA2, HSFA1e, HSFA7a, HSFA7b, HSFA9, HSFB2 and HSFB3. The HSFB2b protein, a repressor transcription factor, has been studied in detail, and it exhibits good solubility, DNA binding, and evidence of SUMOylation in the presence of DNA. In summary, our research shed light on some of the mechanisms underlying SUMOylation, thus providing interesting perspectives for understanding and improving plant responses to heat stress, with potential implications in agriculture.

Keywords: heat stress, post-translational modification, SUMOylation, heat shock factors, SUMOylation motif (Ψ -K-x-E)

CHAPITRE 1

REVUE DE LA LITTÉRATURE

1.1 Stress thermique : réponse physiologique et moléculaire

Le stress thermique, résultant de variations extrêmes de température, est un défi majeur pour les plantes, impactant à la fois leur croissance et leur productivité (Mishra et al., 2023). Les deux extrêmes, chaleur intense et gel, entraînent des conséquences significatives sur l'agriculture et le climat, et le Canada n'est pas épargné par ces problèmes. Les plantes sont des organismes sensibles aux changements environnementaux, et le stress thermique peut perturber leur physiologie, leur métabolisme et leur développement (Hasanuzzaman et al., 2013). Les températures élevées peuvent provoquer la dénaturation des protéines, la déshydratation cellulaire et la perturbation de la photosynthèse (Song et al., 2014). D'autre part, le gel peut entraîner la formation de cristaux de glace qui endommagent les membranes cellulaires, perturbent la circulation de l'eau et entraînent une instabilité osmotique (Pearce, 2001). D'un point de vue agricole, le stress thermique entraîne des conséquences directes sur les rendements des cultures. Les plantes vulnérables à la chaleur peuvent connaître une période de croissance plus courte, tandis que celles sensibles au gel risquent de subir des lésions permanentes. Les températures extrêmes peuvent également favoriser le développement de maladies et de ravageurs, exacerbant les pertes de récoltes (Ma et al., 2021). Les conséquences économiques de ces pertes peuvent être considérables pour les agriculteurs et les communautés dépendantes de l'agriculture. Selon des études récentes, le changement climatique mondial accentue les extrêmes thermiques, avec une augmentation des vagues de chaleur et des périodes de gel dans certaines régions (Orlowsky et al., 2011).

Au Canada, le réchauffement climatique est déjà perceptible, avec une hausse moyenne des températures annuelles. Par exemple, des régions comme les Prairies canadiennes connaissent une augmentation de la fréquence et de l'intensité des vagues de chaleur, tandis que d'autres régions, comme les provinces de l'est, peuvent être confrontées à des hivers plus rigoureux

(Smoyer-Tomic et al., 2003). En ce qui concerne les cultures spécifiques, les cultures céréalières telles que le blé, l'orge et le maïs sont particulièrement sensibles aux vagues de chaleur (Lamaoui et al., 2018). Les cultures fruitières, comme les pommes, les cerises et les raisins, peuvent également être affectées par le stress thermique (Fan et al., 2011). D'autre part, les cultures hivernales comme le colza et le blé d'hiver peuvent être vulnérables aux gelées printanières tardives (Fiebelkorn, 2016). De plus, le stress thermique peut affecter la production agricole en réduisant les rendements de cultures sensibles à la chaleur, comme le blé, le maïs, et d'autres cultures céréalières (Farhad et al., 2023). Des stratégies d'adaptation et d'atténuation sont essentielles pour minimiser les effets du stress thermique sur l'agriculture canadienne. Cela pourrait inclure le développement de variétés de cultures résistantes à la chaleur ou au froid, l'optimisation des pratiques agricoles, et l'investissement dans des technologies d'irrigation efficaces pour faire face aux périodes de sécheresse. En somme, le stress thermique représente un défi croissant pour l'agriculture au Canada, avec des répercussions potentiellement graves sur la sécurité alimentaire et l'économie (Agriculture et Agroalimentaire Canada). Le stress thermique chez les plantes peut avoir de nombreux impacts à la fois sur le plan biochimique, physiologique et au niveau de la régulation. Ces différentes réponses permettent aux plantes de survivre et de s'adapter aux changements climatiques.

Le stress thermique causé par des températures élevées constitue un défi significatif pour les plantes, influant profondément sur leur physiologie et leur métabolisme. Les mécanismes de régulation en réponse à la chaleur reflètent la complexité des réponses adaptatives des plantes (Guerra et al., 2015). L'ouverture et la fermeture des stomates sur les feuilles jouent un rôle central dans la régulation hydrique, une réponse cruciale face à l'augmentation de la transpiration induite par la chaleur (Nilson et al., 2007). Cependant, sous un stress thermique intense, cette régulation peut être compromise, entraînant une déshydratation des cellules végétales (Zhao et al., 2021). Le processus de photosynthèse, essentiel à la croissance des plantes, est également influencé par les températures élevées, réduisant l'efficacité de la fixation du carbone. Simultanément, la respiration cellulaire peut s'accélérer, intensifiant la consommation d'oxygène et impactant le métabolisme global (Tardieu, 2017). En réponse à ces perturbations, les plantes produisent des protéines de choc thermique qui agissent comme des chaperons moléculaires, facilitant la réparation ou l'élimination des protéines endommagées par la chaleur, préservant

ainsi l'intégrité des structures cellulaires et des fonctions métaboliques (Hu et al., 2022). De plus, les plantes mobilisent des métabolites protecteurs tels que les osmoprotecteurs, comprenant des sucres solubles et la proline, qui contribuent à maintenir l'homéostasie cellulaire et à atténuer les effets négatifs du stress thermique (Gosh et al., 2021). La signalisation hormonale, impliquant des composés tels que l'acide abscissique (ABA) et les cytokinines, joue un rôle crucial dans la réponse au stress thermique, régulant la transpiration, la croissance des racines et d'autres processus adaptatifs (Mandal et al., 2022).

Au Canada, où les changements climatiques entraînent une augmentation des températures, ces mécanismes de régulation revêtent une importance particulière. Les cultures telles que le blé et le maïs, sensibles aux vagues de chaleur, pourraient subir des rendements réduits sans l'application de stratégies d'adaptation (Couedel et al., 2021). Ainsi, la compréhension approfondie de ces mécanismes moléculaires pourra permettre de développer des variétés de cultures plus résistantes à la chaleur, une approche cruciale pour garantir la sécurité alimentaire dans un contexte de changement climatique croissant.

1.2 Les protéines de choc thermique et les facteurs de choc thermique

Les protéines de choc thermique (HSPs) jouent un rôle essentiel dans la capacité des plantes à faire face au stress thermique. Lorsque les plantes sont exposées à des températures élevées, les protéines cellulaires peuvent subir une dénaturation, compromettant ainsi leur structure et leur fonction (Hu et al., 2020). Les HSPs émergent comme des défenseurs essentiels contre ces dommages. Ces protéines agissent comme des chaperons moléculaires, favorisant le repliement correct des protéines dénaturées et facilitant leur stabilité fonctionnelle (Pommier, 2000). Elles sont impliquées dans la protection des protéines cellulaires, la prévention de l'agrégation protéique et la réparation des dommages induits par le stress thermique (Arrigo, 2005). Les HSP chez les plantes appartiennent à différentes familles (Tableau 1.1), dont les HSP70, HSP90 et les petites HSP, chacune ayant des fonctions spécifiques dans la réponse au stress thermique (Kang et al., 2021). Par exemple, les HSP70 interviennent dans le repliement des protéines nouvellement synthétisées, tandis que les HSP90 interagissent avec des protéines impliquées dans

la signalisation cellulaire (Kang et al., 2021). Les petites HSP, quant à elles, jouent souvent un rôle dans la protection des protéines contre le stress thermique (Kang et al., 2021). Les facteurs de choc thermique (HSF), des protéines régulatrices, jouent un rôle central dans la régulation transcriptionnelle des HSP chez les plantes (Guo et al., 2016). En conditions normales, les HSF existent sous une forme inactive (Morano et al., 1999). Cependant, lorsque les plantes sont exposées à des températures élevées, les HSFs subissent des modifications conformationnelles qui les activent (Fortunato et al., 2023). Une fois activés, ces HSF se transloquent dans le noyau cellulaire, où ils orchestrent l'expression des gènes des HSPs en se liant aux éléments de réponse au stress thermique présents dans les régions promotrices des gènes cibles (Morano et Thiele, 2018). Les HSFs chez les plantes sont également soumises à une régulation post-traductionnelle. Les modifications telles que la phosphorylation et l'oligomérisation des HSF contribuent à leur activation et à leur translocation nucléaire (Sarge et Morimoto, 1993). Ces mécanismes complexes de régulation assurent une réponse rapide et adaptative des plantes aux températures élevées, permettant la synthèse rapide de HSP pour atténuer les effets négatifs du stress thermique.

Arabidopsis thaliana, souvent utilisée comme modèle en biologie végétale, offre une perspective fascinante sur les HSFs et leur rôle crucial dans la réponse au stress thermique. Cette petite plante herbacée, appartenant à la famille des Brassicacées, est devenue une espèce modèle en raison de son cycle de vie court, de son génome séquencé et de sa facilité de manipulation génétique (Mukaimar, 2015). Chez *Arabidopsis*, il y a 21 HSFs et plusieurs d'entre elles ont été étudiées en détails pour comprendre leur fonction dans la régulation de la réponse au stress thermique (Wang et al., 2021). Les HSF jouent un rôle central dans la régulation transcriptionnelle des gènes de HSP chez *Arabidopsis*, ce qui conduit à une augmentation rapide de la synthèse de HSP (Swindell et al., 2007).

Tableau 1.2 Cinq majeures familles de protéines de choc thermique et leur fonction sous stress thermique (Kang et al., 2021)

Famille HSP	Poids moléculaire (kDa)	Emplacement subcellulaire	Fonction liée au stress thermique
HSP100	100-104	Cytosol Mitochondrie Chloroplaste	Désagrégation des protéines et implication dans la dégradation des protéines
HSP90	80-94	Cytosol Réticulum endoplasmique Noyau Mitochondrie Chloroplaste	Repliement des protéines, transduction des signaux (la plupart des substrats de HSP90s sont des kinases et des facteurs de transcription)
HSP70	68-75	Cytosol Réticulum endoplasmique Mitochondrie Chloroplaste	Assiste au repliement des protéines non-natives pour bloquer la dégradation des protéines dans le réticulum endoplasmique ainsi que l'importation et la translocation des protéines
HSP60	57-60	Mitochondrie Cytosol Noyau Chloroplaste	Assiste au repliement des polypeptides dépliés dans la matrice mitochondriale
HSPs	15-42	Cytosol Réticulum endoplasmique Mitochondrie Chloroplaste Membrane plasmique	Prévient l'agrégation et le repliement des polypeptides dépliés

Certaines HSFs peuvent être phosphorylées, acétylées ou subir d'autres modifications post-traductionnelles (Nitika et al., 2020). Ces altérations peuvent impacter leur activité, stabilité et localisation cellulaires, affectant ainsi leur capacité à protéger d'autres protéines contre les effets du stress thermique tel que la dénaturation, la perturbation de l'activité enzymatique, la réponse au choc thermique, la modification de la conformation protéique et l'expression génique (Miller, 1993). Les modifications post-traductionnelles peuvent donc jouer un rôle dans la régulation de la réponse au stress. En outre, les HSF eux-mêmes peuvent être modulés par des modifications post-traductionnelles. Par exemple, la phosphorylation des HSF peut moduler leur activité transcriptionnelle. La phosphorylation, qui est l'addition de groupes phosphate, est une modification post-traductionnelle commune qui peut influencer divers aspects des protéines, y compris leur activité fonctionnelle (Whittam et al., 1955). Les sites de phosphorylation sur les HSF peuvent être localisés sur des résidus spécifiques d'acides aminés de la protéine. Ces sites peuvent être ciblés par des kinases, des enzymes qui ajoutent des groupes phosphate (Evrard et al., 2013). La kinase MPK6 phosphoryle spécifiquement la protéine HSFA2 au niveau du résidu T249, entraînant une modification de sa localisation intracellulaire (Evrard et al., 2013). Dans l'ensemble, l'étude de Evrard et al. démontre que la stabilité de la protéine HSFA2 est étroitement régulée par des mécanismes dépendants de la phosphorylation (Evrard et al., 2013).

De plus, le ciblage de HSFA2 par MPK6 joue un rôle crucial dans la réponse au stress thermique, contribuant ainsi à un mécanisme de régulation complexe qui gouverne la réaction des plantes face à ce type de stress (Evrard et al., 2013). La phosphorylation des HSF peut se produire dans différentes parties de la protéine, telles que le domaine de liaison à l'ADN ou d'autres domaines régulateurs (Pommier, 2000). La localisation précise des sites de phosphorylation peut varier en fonction des HSF spécifiques et des conditions cellulaires. Quant à l'impact sur l'activité transcriptionnelle des HSF, la phosphorylation peut avoir des effets divergents. Dans certains cas, la phosphorylation peut activer les HSF, induisant ainsi leur capacité à se lier aux éléments de réponse au stress thermique dans les régions promotrices des gènes de HSP (Arrigo, 2005). Cela stimule la transcription des gènes des HSP renforçant ainsi la réponse cellulaire au stress thermique. Cependant, il est important de noter que la phosphorylation peut également inhiber l'activité transcriptionnelle des HSF dans d'autres contextes (Gomez-Pastor et al., 2018). Des

études ont montré que la phosphorylation peut réguler négativement les HSF, modifiant ainsi leur capacité à induire la transcription des gènes de HSP (Kim et al., 1998). Ces effets peuvent être dépendants du contexte cellulaire, du type de stress thermique et du niveau d'activation des HSF. Cette régulation entre les HSF et les modifications post-traductionnelles contribue à réguler la réponse au stress thermique.

1.3 Modification post-traductionnelle : La SUMOylation

La SUMOylation, une modification post-traductionnelle essentielle, consiste en l'ajout covalent de petites protéines semblables à l'ubiquitine, appelées Small Ubiquitin-like MOdifier (SUMO), à des protéines cibles spécifiques. (Braun et al., 2009). Bien que la SUMOylation soit un processus complexe impliquant différentes enzymes, son rôle général peut être exploré indépendamment des détails enzymatiques. Cette modification régule divers aspects de la biologie cellulaire en influençant la localisation, l'activité et les interactions des protéines cibles (Zhao et al., 2020). Contrairement à d'autres modifications similaires telles que l'ubiquitination, la SUMOylation n'entraîne généralement pas la dégradation des protéines cibles, mais modifie plutôt les propriétés fonctionnelles des protéines (Wilkinson et Henley, 2010). La SUMOylation est impliquée dans une variété de processus cellulaires, dont la signalisation cellulaire, la régulation transcriptionnelle, la réparation de l'ADN et la réponse au stress (Enserink, 2015). Dans le contexte du stress thermique, la SUMOylation joue un rôle crucial en permettant aux cellules de s'adapter et de survivre à des conditions environnementales difficiles (Filippopoulou et al., 2020). Les protéines impliquées dans la réponse au stress thermique sont souvent sumoylées, modifiant ainsi leur fonction et leur interaction avec d'autres composants cellulaires (Robert, 2019). La modification par SUMO est également impliquée dans la régulation de la conformation des protéines, ce qui peut influencer leur stabilité et leur capacité à remplir leurs fonctions biologiques. En réponse au stress thermique, la SUMOylation peut agir comme un mécanisme de régulation fine pour maintenir l'homéostasie cellulaire, favoriser la survie cellulaire et prévenir les dommages cellulaires irréversibles (Robert, 2019). La SUMOylation représente ainsi un mécanisme sophistiqué et hautement régulé de modification post-traductionnelle des protéines,

jouant un rôle crucial dans la régulation des processus cellulaires, particulièrement importants dans la réponse adaptative des cellules au stress thermique.

La modification post-traductionnelle qui implique les enzymes E1 (SUMO-activating enzyme 1/2), E2 (SUMO-conjugating enzyme 1), E3 (ligase) et SUMO (Figure 1.1 et 1.2) commence par le rôle essentiel de l'enzyme E1 dans l'activation de la protéine SUMO. Cette activation se produit lorsque l'enzyme E1 utilise l'hydrolyse de l'ATP pour former un pont thioester avec SUMO. L'enzyme E1 se présente sous la forme d'un dimère, composé d'une sous-unité activating enzyme subunit 1A (SAE1A) ou activating enzyme subunit 1B (SAE1B) et d'une sous-unité activating enzyme subunit 2 (SAE2) (Brackett et al., 2022). La protéine SAE1A d'Arabidopsis (Uniprot Q8VY78) est constituée de 322 acides aminés. Parallèlement, la protéine SAE2 d'Arabidopsis (Uniprot Q9SJT1) est composée de 700 acides aminés. La structure d'une enzyme E1, dont le poids moléculaire varie généralement entre 95 et 115 kDa, a été élucidée dans les années 1980 grâce à la radiocristallographie. Son site actif renferme une cystéine en position C168 de la sous-unité 2, jouant ainsi un rôle central dans le déroulement de l'activation du SUMO (Szczepanowski et al., 2005). Cette étape initiale de la SUMOylation sert de point de départ à la cascade de réactions post-traductionnelles, permettant la liaison ultérieure de SUMO à des protéines cibles spécifiques. Concernant l'enzyme E2, il s'agit d'une enzyme de conjugaison qui assume la responsabilité du transfert de l'adhésif SUMO activé par l'enzyme E1 vers les protéines cibles. Comparativement à l'enzyme E1, l'enzyme E2, impliquée dans la SUMOylation, est généralement de plus petite taille, affichant un poids moléculaire d'environ 15 à 25 kDa. La structure d'enzyme E2 a été élucidée au cours des années 1980 et 1990 grâce à la radiocristallographie (Gareau et al., 2010). L'enzyme E2 abrite un site actif qui permet le transfert de SUMO depuis la cystéine 94 vers des résidus lysines de protéines cibles (Gareau et al., 2010). La protéine SCE1 d'Arabidopsis (Uniprot Q42551) est composée de 160 acides aminés. Les enzymes E3, quant à elle, jouent un rôle crucial en facilitant la conjugaison spécifique de SUMO à des protéines cibles particulières, en favorisant leur rapprochement mutuel (Xiong et al., 2013). Les enzymes E3 présentent une diversité marquée, avec une variété de types particulièrement notable pour les E3 de l'ubiquitine. En règle générale, elles présentent une taille plus importante que l'enzyme E2, avec des poids moléculaires variables (Toma-Fukai et al., 2021). Le rôle de médiateur des enzymes E3 dans l'interaction spécifique entre SUMO et les

protéines cibles contribue de manière significative à la précision et à la spécificité de la SUMOylation dans la régulation des processus cellulaires (Wilkinson & Hensley).

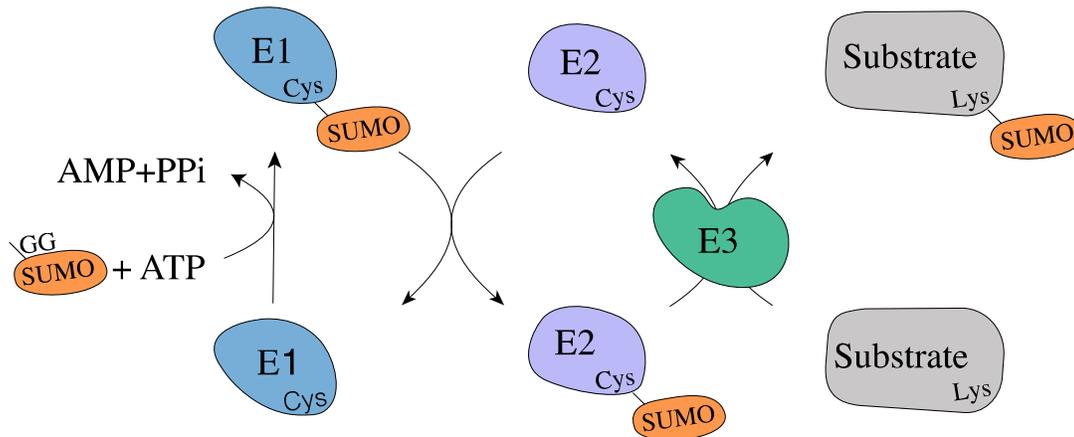


Figure 1.1 Cascade enzymatique de la voie de SUMOylation. L'enzyme E1 utilise une molécule d'ATP pour activer la molécule SUMO. La SUMO activée est par la suite transférée sur l'enzyme de conjugaison E2. L'enzyme E2 et l'enzyme E3 ligase permettent la fixation de la protéine SUMO sur un résidu lysine des protéines substrats. Figure prise de (Jmii & Cappadocia, 2021).

1.3.1 Les enzymes E3 ligases

Les SUMO E3 ligases (Figure 1.2) sont des acteurs clés dans la régulation fine de la SUMOylation des protéines, apportant une contribution significative à une variété de processus cellulaires tels que la régulation de l'expression génique, la réparation de l'ADN et la signalisation cellulaire (Yang et al., 2021). Leurs rôles essentiels dans ces mécanismes sont observés de manière universelle chez les plantes, les levures et les humains, bien que des nuances spécifiques puissent exister entre ces groupes biologiques. Ces protéines sont universellement impliquées dans l'ajout de la protéine SUMO à des protéines cibles spécifiques, participant ainsi à la stabilité et à la fonction des protéines engagées dans des processus clés tels que la régulation de l'expression génique, la réparation de l'ADN et la signalisation cellulaire (Shi et al., 2022). Malgré ces similitudes fondamentales, des différences notables existent entre les espèces. Les plantes, en tant qu'organismes photosynthétiques, peuvent présenter des adaptations spécifiques à

leur environnement, influençant ainsi la régulation de la SUMOylation par les SUMO E3 ligases. Les levures, en tant que modèles unicellulaires, ont grandement contribué à la compréhension initiale des mécanismes moléculaires de la SUMOylation (Di Gregorio et al., 2018). Ces organismes unicellulaires ont des mécanismes spécifiques d'homéostasie cellulaire et de régulation des fonctions cellulaires qui impactent le rôle des E3 ligases dans la SUMOylation (Enserink, 2015). Les humains, en raison de leur complexité biologique, présentent des réseaux de signalisation plus élaborés, où les E3 ligases sont impliquées dans des voies complexes régulant finement les processus cellulaires. De plus, les implications pathologiques liées à des dysfonctionnements dans la SUMOylation, régulée par les E3 ligases, sont spécifiques aux organismes humains, associés à diverses maladies telles que certains cancers et maladies neurodégénératives (Lescouzères et Bomont, 2020). Ainsi, bien que les principes fondamentaux soient conservés, la spécificité moléculaire, les adaptations biologiques et les implications pathologiques varient entre les plantes, les levures et les humains.

Chaque SUMO E3 ligase a une spécificité de criblage pour les protéines substrats, et elles jouent un rôle clé dans la modulation de l'activité des protéines sumoylées. La protéine SIZ-1 (E3 SUMO-protein ligase, Uniprot Q680Q4) Arabidopsis fait 884 acides aminés. Une partie de sa structure soit de l'acide aminé 104 à 168 a été élucidée par RMN. La protéine SIZ-1 (SAP and MIZ1) comporte quatre différents domaines conservés; un domaine SAP (SAF-A/B, Acinus) qui est conservé dans différentes protéines. Le domaine SAP, avec une séquence consensus de 35 résidus, a été initialement proposé comme un nouveau domaine de liaison à l'ADN trouvé dans diverses protéines nucléaires impliquées dans l'organisation chromosomique. Les domaines SAP ont la capacité de se lier à un ADN riche en A/T (Suzuki et al., 2008). Chez SIZ-1, le domaine SAP occupe les résidus 11 à 45 acides aminés. Le domaine PHD (Plant Homeodomain) de 112 à 168 acides aminés s'associe plutôt à d'autres protéines impliquées dans la régulation de l'expression génique et dans la modification épigénétique notamment la méthylation de l'ADN et l'acétylation des histones (Rytz, 2018). Le domaine PINIT de certaines protéines PIAS (Protein Inhibitor of Activated STAT) mammifères régule leur localisation subcellulaire et influence la sélectivité du substrat (Duval et al., 2003; Wong et al., 2004). PIIT (pour Pro-Ile-Ile-Thr) est la forme modifiée du motif PINIT qui se trouve dans SIZ-1. Le domaine s'associe à d'autres protéines (Rytz, 2018, Cheong et al., 2009, Jmii et al., 2021). Finalement, le domaine SP-RING

(SIZ/PIAS RING) de 346 à 429 acides aminés est un domaine responsable de l'activité de conjugaison de SUMO aux protéines cibles (Rytz, 2018).

La protéine SIZ-1 est impliquée dans la régulation de la réponse au stress de haute température chez les plantes via la SUMOylation de protéines cibles. Par exemple, en réponse à une élévation de la température, la plante active la synthèse de protéines de choc thermique pour protéger ses protéines et maintenir l'homéostasie cellulaire (Jardin et Lejour, 1991). SIZ-1 modifie ces HSP par SUMOylation, ce qui modifie leur activité et leur localisation, renforçant ainsi leur efficacité dans la réponse au stress thermique (Takahashi et Kikuchi, 2008). De plus, SIZ-1 régule la stabilité et l'activité d'autres protéines clés impliquées dans la réponse au stress thermique, créant un réseau complexe de régulation moléculaire (Takahashi et Kikuchi, 2008). Des études ont montré que les mutants SIZ-1 présentent une sensibilité accrue au stress thermique, soulignant l'importance de SIZ-1 dans l'adaptation des plantes aux températures élevées (Rytz, 2018, Yoo et al., 2006). SIZ-1 agit également comme un régulateur négatif de l'expression génique de l'HSFA2, un facteur de transcription crucial dans la réponse au stress thermique, suggérant une régulation fine du réseau de signalisation (Zhang et al., 2017). L'implication de SIZ-1 dans la réponse au stress de haute température offre des perspectives prometteuses pour améliorer la résilience des plantes face aux changements climatiques, en exploitant cette voie de signalisation pour renforcer leur capacité à survivre et à prospérer dans des conditions environnementales défavorables.

La protéine MMS21 (E3 SUMO-protein ligase, Uniprot Q8GYH7) est aussi reconnue sous le nom Nse2 (Non-Smc Element 2) chez la levure *Saccharomyces cerevisiae* (Jmii et Cappadocia, 2021). C'est une protéine de 249 acides aminés et qui possède un domaine SP-RING de 141 à 228 acides aminés. Cette SUMO E3 ligase est importante dans la régulation de la SUMOylation des protéines associées à la chromatine, notamment dans des processus liés à la réparation de l'ADN et à la régulation de la structure chromatinienne (Zhao et Blobel, 2005).

La protéine PIAL1 (PIAS-like protein 1; E4 SUMO-protein ligase, Uniprot A0A0A7EPL0) chez *Arabidopsis* fait 847 acides aminés et possède un domaine SP-RING de 268 à 340 acides aminés, ainsi qu'un domaine IND (domaine d'interaction) nécessaire pour l'interaction avec MOM1 et PIAL2. La protéine PIAL2 (PIAS-like protein 2; E4 SUMO-protein ligase PIAL2, Uniprot F4JYG0) chez *Arabidopsis* fait quant à elle 760 acides aminés, a un domaine SP-RING de l'acide

aminé 298 à l'acide aminé 379, ainsi qu'un domaine d'interaction entre 143 et 301 acides aminés afin de pouvoir interagir avec MOM1 et PIAL1. Ces protéines sont impliquées dans la régulation de stabilité des protéines et la réponse aux stress environnementaux. Les protéines PIAL1 et PIAL2 ainsi que la protéine SIZ-1 sont toutes des membres de la famille des E3 ligases SUMO, mais elles se distinguent par leurs caractéristiques structurales et leurs rôles spécifiques dans la régulation de la SUMOylation (Jmii et Cappadocia, 2021). SIZ-1 est impliquée dans la réponse au stress thermique en régulant la SUMOylation de protéines clés, y compris les protéines de choc thermique, ce qui contribue à l'adaptation des plantes à des températures élevées. SIZ-1 agit également comme un régulateur de la transcription, influençant l'expression de gènes impliqués dans la réponse au stress thermique. Les protéines PIAL1 et PIAL2, en revanche, appartiennent à la famille des protéines PIAS (Protein Inhibitor of Activated STAT), qui est également une famille d'E3 ligases SUMO (Tomanov et al., 2014). Contrairement à SIZ-1, les protéines PIAS sont souvent associées à la régulation de la signalisation des cytokines et des facteurs de transcription dans les organismes animaux (Niu et al., 2018). Elles sont bien étudiées dans le contexte de la régulation immunitaire et de la réponse inflammatoire chez les mammifères. Ces distinctions reflètent la diversité fonctionnelle au sein de la famille des E3 ligases SUMO en réponse aux besoins spécifiques des organismes et des voies biologiques qu'elles régulent.

Les SUMO E3 ligases représentent des acteurs cruciaux du processus de SUMOylation au sein d'*Arabidopsis thaliana*, particulièrement dans le contexte de la régulation de protéines liées à la structure de l'ADN, la réponse au stress, ainsi que d'autres processus associés à l'ADN (Siu et al., 2020). Bien que SIZ-1 ne lie pas l'ADN de manière séquence-spécifique, des études ont mis en lumière des interactions protéine-protéine spécifiques entre la SIZ-1 et des protéines étroitement associées à l'ADN (Kim et al., 2017). Ces protéines comprennent des facteurs de transcription et des régulateurs de la chromatine (Kim et al., 2017). La présence de SUMO sur la chromatine est corrélée aux marqueurs de la chromatine active, en accord avec des expériences de microscopie à fluorescence montrant que SUMO1 de maïs est davantage associé à l'euchromatine qu'à l'hétérochromatine (Chen et al., 2019; Jmii et al., 2021). Lors de l'induction du stress thermique, SUMO se réarrange pour réguler à la hausse les gènes de stress thermique tout en régulant à la baisse les gènes de croissance (Han et al., 2020). Il est important de noter que peu d'associations de SUMO à l'ADN se produit en l'absence de SIZ-1, soulignant ainsi l'importance de SIZ-1 pour

cibler SUMO sur l'ADN. Lorsque la protéine SIZ-1, impliquée dans la SUMOylation, est absente, cela entraîne une perturbation de la SUMOylation des protéines associées à l'ADN, ce qui se traduit par une activation des gènes de réponse au stress et une inhibition des gènes de croissance. Ainsi, la SUMOylation exerce un contrôle sur l'expression des gènes via des modifications post-traductionnelles des protéines associées à l'ADN, ce qui a un impact sur les réponses de la plante au stress et à la croissance. En accord avec un rôle de SUMO en tant qu'amplificateur de la réponse au stress, l'activation des gènes responsables du stress et l'inhibition des gènes liés à la croissance étaient toujours présentes chez les plantes dépourvues de SIZ-1. (Han et al., 2020; Jmii et al., 2021). Ces recherches laissent entendre qu'une fraction de la SUMOylation, régulée par les ligases E3 chez les végétaux, pourrait avoir lieu directement sur l'ADN. En outre, les protéines SIZ des plantes comportent un domaine SAP qui a été confirmé comme étant en liaison avec l'ADN (Suzuki et al., 2009). Ces interactions complexes peuvent faciliter l'ancrage indirect de la SIZ-1 à l'ADN, révélant une dimension nouvelle et indirecte de son implication dans la régulation des processus liés à l'ADN (Kong et al., 2020). Par ailleurs, une analyse protéomique a révélé que de nombreuses cibles de SIZ-1 sont constituées de facteurs de transcription ou de protéines impliquées dans le remodelage de la chromatine, lesquelles possèdent des domaines de liaison à l'ADN (Rytz et al., 2018). La nature non-séquence-spécifique de la liaison de SIZ-1 à l'ADN suggère qu'elle pourrait éventuellement se lier à proximité des facteurs de transcription associés à l'ADN, influençant potentiellement la sélection de la lysine ou complétant les interactions protéine-protéine par des interactions protéine-ADN. Au-delà d'un simple contact avec l'ADN, il est envisageable que SIZ-1 puisse interagir avec la chromatine ouverte, grâce à son domaine PHD qui a été démontré comme étant en liaison avec l'histone H3K4 tri-méthylée (Shindo et al., 2012 ; Miura et al., 2020). Cette influence sur les protéines régulatrices de la chromatine, en particulier les facteurs de transcription, constitue un aspect essentiel de son rôle fonctionnel (Miura et al., 2020). La SUMOylation de ces éléments régulateurs peut altérer leur activité et leur localisation, modulant ainsi la compaction de la chromatine et influençant l'accessibilité des sites de liaison à l'ADN. La SUMOylation par la SIZ-1 ne se contente pas de réguler la structure de la chromatine, mais elle entraîne des répercussions significatives sur la transcription génique, notamment des gènes impliqués dans la réponse au stress thermique (Sharma et al., 2023). Ainsi, le domaine SAP de la SIZ-1 lie l'ADN, son impact sur la régulation de la structure de la chromatine et sa capacité à moduler la réponse au stress thermique soulignent son rôle central et multifacette dans

la coordination de réponses cellulaires complexes au sein d'*Arabidopsis thaliana*. Ces découvertes suggèrent un réseau sophistiqué de régulation moléculaire dans lequel la SUMOylation orchestrée par la SIZ-1 émerge comme un mécanisme crucial dans la modulation des réponses adaptatives aux changements environnementaux, en particulier lors d'épisodes de stress thermique.

1.3.2 Les protéines SUMO (Small Ubiquitin-like MOdifier)

Il existe huit variants distincts de SUMO, à savoir SUMO-1, SUMO-2, SUMO-3, SUMO-4, SUMO-5, SUMO-6, SUMO-7 et SUMO-8 (Tableau 1.2 et Figure 1.2), qui peuvent établir des liaisons réversibles avec les protéines cibles. Chacun de ces types de SUMO présente des particularités légèrement différentes, notamment des poids moléculaires et des points isoélectriques (pI) spécifiques. Ces différentes formes de SUMO reflètent possiblement une diversité fonctionnelle qui peut influencer leurs interactions spécifiques avec les protéines cibles et, par conséquent, leur rôle dans la régulation des processus cellulaires. Les protéines SUMO-1 et SUMO-2 d'*Arabidopsis* constituent les principales protéines conjuguées sur les protéines, particulièrement en condition de stress (Kurepa et al., 2003).

Tableau 1.3.2 Différentes protéines SUMO présentent chez *Arabidopsis thaliana*

SUMO	Numéro d'accèsion (Uniprot)	Taille (a. a.)	pI
SUMO-1	P55852	100	4.91
SUMO-2	Q9FLP6	103	5.35
SUMO-3	Q9FLP5	111	5.09
SUMO-4	Q6EEV6	95	6.57
SUMO-5	Q8VZI7	108	8.91
SUMO-6	Q9FKC6	117	6.85
SUMO-7	Q3E8A8	95	6.03
SUMO-8	B3H5R8	97	6.03

La conjugaison de SUMO influence les fonctions des facteurs de transcription de plusieurs manières, notamment : (a) en influençant l'activité de liaison à l'ADN, (b) en favorisant l'élimination du facteur de transcription de la chromatine, (c) en modifiant l'étendue de l'association du facteur de transcription à ses sites cibles associés sur la chromatine, (d) en régulant l'abondance et/ou la localisation du facteur de transcription, (e) en altérant la capacité du facteur de transcription à interagir avec des régulateurs transcriptionnels tels que des co-répresseurs ou des coactivateurs et, (f) en interagissant avec d'autres modifications post-traductionnelles qui peuvent influencer l'activité du facteur de transcription (Roy et al., 2021). Chacun des types de SUMO présente des caractéristiques structurales et fonctionnelles qui les rendent uniques dans leur contribution à la régulation cellulaire. Parmi eux, chez l'humain, SUMO-1 est souvent associé à des modifications monoSUMOylation et exprimé de manière ubiquitaire, régulant divers processus cellulaires tels que la signalisation, la transcription et la stabilisation des protéines (Wang et Dasso, 2009). En revanche, SUMO-2 et SUMO-3 sont principalement impliqués dans des modifications polySUMOylation, formant des chaînes de SUMO, et jouent un rôle clé dans la régulation de la réponse au stress, notamment le stress thermique et la perturbation génomique chez l'humain (Park et al., 2011). Au sein du règne végétal, les SUMO-1, SUMO-2 et SUMO-3 partagent des mécanismes fondamentaux tout en assumant des rôles spécifiques dans la régulation des protéines cibles et dans la réponse au stress. Quant à SUMO-4, SUMO-5, SUMO-6, SUMO-7 et SUMO-8, leur rôle précis dans la conjugaison et leurs fonctions spécifiques restent moins caractérisés par rapport aux autres membres de la famille SUMO. Leur expression peut être plus spécifique à certains tissus ou conditions, et davantage de recherches sont nécessaires pour élucider pleinement leurs rôles distincts dans la régulation cellulaire. La recherche constante sur les SUMO continue à dévoiler de nouvelles perspectives sur la complexité de ces modifications post-traductionnelles et leur impact sur la régulation précise des processus cellulaires chez les mammifères, influencée par une séquence d'acides aminés limitée (Robert, 2019).

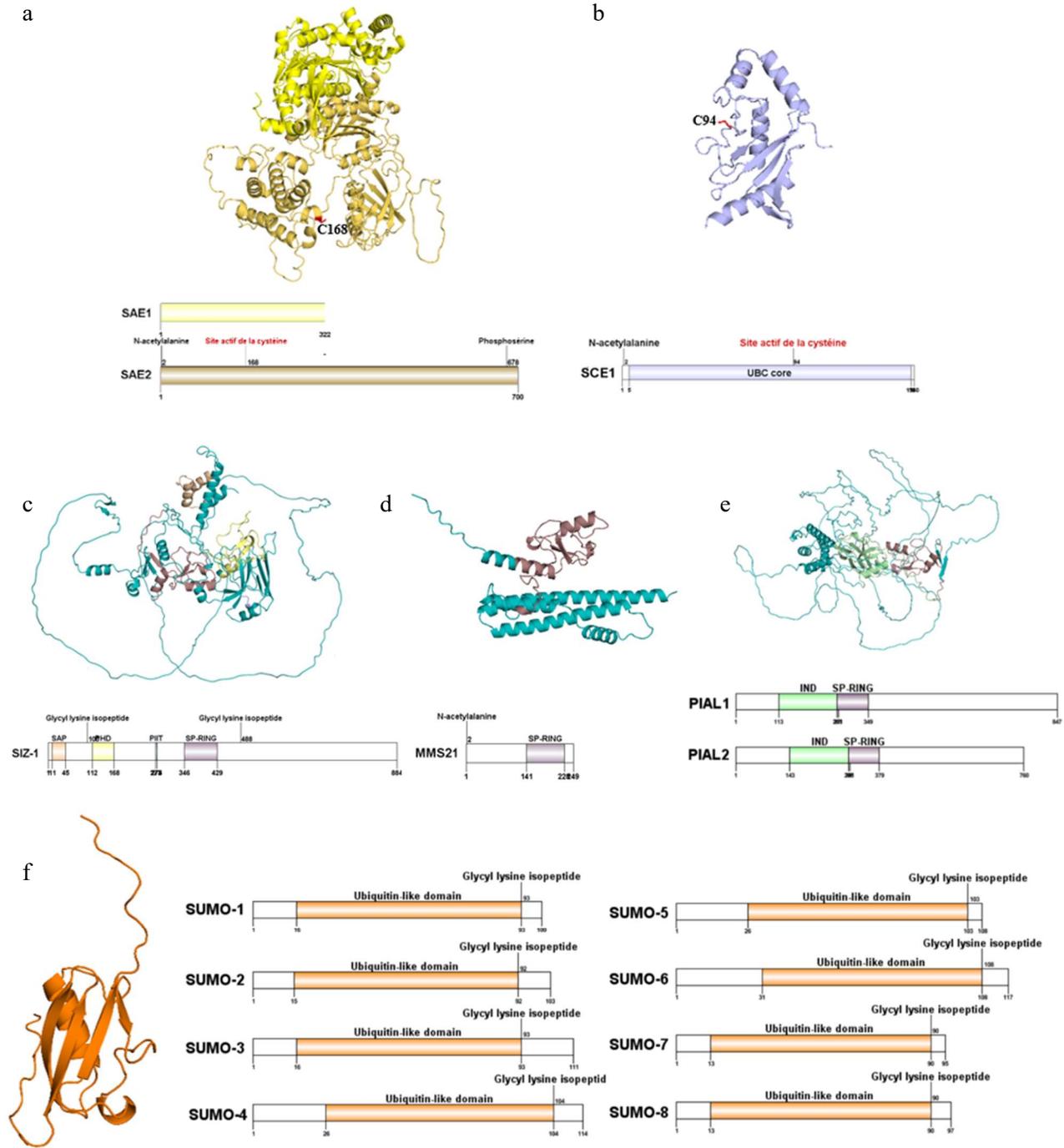


Figure 1.2 Enzymes de la SUMOylation. a) Structure des deux sous-unités de l'enzyme E1 ainsi que le site actif de la cystéine. b) Structure de l'enzyme E2 avec le site actif de la cystéine. c) Structure de la SIZ-1 (E3 ligase) avec les domaines SAP, PHD, PIIT et SP-RING. d) Structure de MMS21 (E3 ligase) avec son domaine SP-RING. e) Structure représente la protéine PIAL1/2 qui a un domaine d'IND pour interagir avec MOM-1 et PIAL). f) Structure de la protéine SUMO ainsi que les différents domaines d'Ubiquitin-like chez le SUMO-1 à SUMO-8.

1.4 Le motif consensus ψ KxE/D

Le motif consensus pour le site d'acceptation de la SUMOylation est défini comme ψ KxE/D, une séquence spécifique retrouvée chez plusieurs substrats du processus de SUMOylation chez la plante modèle *Arabidopsis thaliana*. Ce motif se caractérise par la présence d'un acide aminé hydrophobe (ψ), (K) étant la lysine cible pour la SUMOylation, (x) peut être n'importe quel acide aminé, et (E/D) représente un acide aminé acide (glutamate ou aspartate). Ce motif est basé sur des analyses bio-informatiques des sites de SUMOylation connus et a été déduit par l'observation de motifs récurrents dans les protéines cibles de la SUMOylation (Yau et al., 2021). Pour les plantes, le motif de SUMOylation accepteur est similaire à celui des autres organismes, mais il peut y avoir des variations spécifiques à chaque espèce. Les études expérimentales et l'analyse bio-informatique des protéines SUMOylées chez les plantes, telles que notre organisme modèle, *Arabidopsis thaliana*, ont contribué à la détermination de ce motif. Les protéines cibles de la SUMOylation sont souvent impliquées dans la régulation de la signalisation cellulaire, la réponse au stress, la stabilité des protéines, et d'autres processus cellulaires essentiels. La détermination précise du motif de SUMOylation aide à prédire et à identifier les sites potentiels de SUMOylation dans les protéines, ce qui facilite la compréhension de la régulation post-traductionnelle des fonctions cellulaires. Ce motif est reconnu par l'enzyme E2 pour l'addition de la protéine SUMO. Le motif de SUMOylation sur les HSF peut être spécifique au contexte de stress thermique (Wilkinson et Henley, 2010). De plus, du fait de leur très petite taille, les sites de SUMOylation peuvent rapidement apparaître ou disparaître au cours de l'évolution sous l'effet de simples mutations de substitutions (Wilkinson et Henley, 2010).

1.5 21 HSFs chez *Arabidopsis thaliana*

Le génome complet de l'*Arabidopsis thaliana* a été séquencé, et il est constitué de cinq chromosomes, cumulant environ 135 mégabases. Sa taille en fait l'un des plus petits génomes parmi les plantes à fleurs, avec chaque chromosome portant des gènes et des régions régulatrices distincts. De manière approfondie, entre 27 000 et 35 000 gènes ont été identifiés dans le génome de cette plante. Ces gènes englobent divers processus biologiques, tels que la croissance, le

développement, la réponse au stress et la régulation métabolique. La plante modèle contient 21 HSFs. Ces 21 HSFs sont subdivisés en 3 grandes classes, soit la classe A (HSFA1, HSFA1b, HSFA1d, HSFA1e, HSFA2, HSFA3, HSFA4a, HSFA4c, HSFA5, HSFA6a, HSFA6b, HSFA7a, HSFA7b, HSFA8 et HSFA9), la classe B (HSFB1, HSFB2a, HSFB2b, HSFB3 et HSFB4) et finalement la classe C (HSFC1).

Tel que mentionné précédemment, la famille des facteurs de transcription HSF joue un rôle central dans la régulation des réponses au stress thermique chez les plantes. Ces HSFs partagent plusieurs caractéristiques fondamentales qui sous-tendent leur rôle crucial dans la survie des plantes face aux variations de température. Un trait commun à tous les HSFs est la présence de domaines conservés, tels que le domaine de liaison à l'ADN (DBD) et le domaine de liaison aux co-chaperons (HR-A/B), qui sont essentiels pour leur activité transcriptionnelle (Jiang et al., 2023). Le domaine de liaison à l'ADN (DBD) est responsable de la reconnaissance des éléments promoteurs de gènes cibles, spécifiquement ceux impliqués dans la réponse au stress thermique. La conservation de ce domaine parmi les HSFs suggère une similarité fonctionnelle dans la régulation des gènes de réponse au stress thermique.

Les HSFs interagissent avec des séquences spécifiques appelées éléments de réponse au stress thermique (HSEs), généralement caractérisées par la séquence de nucléotides 5'-nGAAn-3' (Zhao et al., 2020). Outre le DBD, le domaine de liaison aux co-chaperons (HR-A/B) est crucial pour la régulation de l'activité transcriptionnelle des HSFs (Gomez-Pastor et al., 2017). Ce domaine permet l'interaction avec des co-chaperons, tels que les protéines HSP70, facilitant ainsi la formation de complexes protéiques fonctionnels. Ces complexes assurent le dépliage des protéines dénaturées et leur translocation vers le noyau pour l'activation de la réponse au stress thermique (Chen et Liu, 2000). La diversité dans les domaines HR-A/B peut refléter des spécificités fonctionnelles, permettant aux différentes HSFs de s'associer à des co-chaperons spécifiques et de moduler ainsi leur activité.

Malgré ces similitudes structurales, les HSFs présentent également des divergences significatives qui confèrent une diversité fonctionnelle à cette famille de facteurs de transcription. Par exemple, la distribution tissulaire et l'expression régulée des HSFs peuvent varier, suggérant une spécialisation fonctionnelle au sein de la famille. Certains HSFs peuvent être spécifiquement

induits en réponse à des conditions de stress thermique, tandis que d'autres peuvent être exprimés de manière constitutive ou régulée par d'autres stimuli environnementaux (Hantelys, 2017). Les HSFs peuvent également former des dimères ou trimères, ajoutant une couche de complexité à leur régulation (Latchman, 2008). Cette capacité à former des complexes multimériques suggère une régulation fine et flexible de la réponse au stress thermique, permettant une adaptation aux conditions environnementales spécifiques rencontrées par la plante (Latchman, 2008). En résumé, les 21 protéines HSFs partagent des caractéristiques structurales essentielles, telles que les domaines de liaison à l'ADN et aux co-chaperons. Ces traits communs définissent leur rôle dans la régulation de la réponse au stress thermique. Cependant, des variations dans la distribution tissulaire, l'expression régulée et la capacité à former des complexes multimériques ajoutent des nuances fonctionnelles à cette famille de facteurs de transcription, permettant aux plantes de mettre en œuvre des stratégies adaptatives complexes face aux défis thermiques (Latchman, 2008).

Tant chez les humains que chez les plantes, la SUMOylation peut réguler l'activité des HSFs par des mécanismes similaires, tels que l'influence sur la liaison à l'ADN, les interactions protéine-protéine et la localisation subcellulaire. Cependant, les détails spécifiques et les résultats de la SUMOylation sur les HSFs peuvent différer entre ces deux règnes. Chez l'humain, la SUMOylation peut altérer l'activité de liaison à l'ADN des HSFs humains, influant sur leur capacité à réguler les gènes associés au stress thermique. Elle exerce également une influence sur les interactions entre les HSFs humains et des protéines co-régulatrices telles que les co-répresseurs ou les coactivateurs, ce qui entraîne des répercussions sur la régulation transcriptionnelle (Gomez-Pastor et al, 2017). En outre, la SUMOylation participe à la localisation subcellulaire des HSFs humains, modifiant leur déplacement entre le cytoplasme et le noyau (Gomez-Pastor et al, 2017). De manière analogue, chez les plantes, la SUMOylation peut moduler l'activité de liaison à l'ADN des HSFs, influençant ainsi leur implication dans l'activation des gènes associés au stress thermique. La SUMOylation joue également un rôle dans la modulation des interactions entre les HSFs végétaux et d'autres protéines impliquées dans la réponse au stress thermique. En outre, la SUMOylation intervient dans la détermination de la localisation subcellulaire des HSFs végétaux, impactant ainsi leur activité dans le noyau (Park et al., 2011). Bien que les principes généraux de la SUMOylation influençant la fonction des HSFs

soient conservés, les protéines spécifiques impliquées, les sites précis de SUMOylation et l'impact global sur la réponse au stress thermique peuvent différer entre les humains et les plantes en raison de la divergence évolutive de ces deux systèmes biologiques.

La question centrale de la localisation des sites de SUMOylation des facteurs de transcription HSFs lors de situations de stress environnemental, notamment le stress thermique, demeure une interrogation cruciale pour élucider les mécanismes régulant la réponse au stress chez les plantes. Les HSFs sont des régulateurs clés impliqués dans la transcription des gènes de choc thermique, orchestrant la réponse cellulaire face aux fluctuations de température. La SUMOylation, une modification post-traductionnelle qui implique l'ajout de SUMO est connue pour moduler l'activité des HSFs et influencer leur capacité à réguler les gènes de réponse au stress thermique. La liaison de SUMO à ces facteurs peut moduler leur activité en altérant leur interaction avec d'autres protéines ou leur liaison à l'ADN, et cette modification peut également servir de signal pour le ciblage des protéines vers des voies cellulaires spécifiques, y compris la dégradation par le protéasome, et dans certains cas, elle peut même stabiliser les protéines en les protégeant de la dégradation. Les ligases ajoutent ou enlèvent les groupements SUMO de leurs cibles, et une fois modifiés, ces protéines peuvent être soumises à la dégradation par le protéasome si elles sont ciblées pour cette voie. La complexité de cette problématique réside dans la diversité des HSFs et des conditions de stress environnementaux. Les différents membres de la famille des HSFs peuvent présenter des spécificités fonctionnelles, et les conditions de stress, telles que le stress thermique, peuvent induire des réponses distinctes selon l'intensité et la durée du stress. La question clé consiste à déterminer comment la SUMOylation influence la fonction des HSFs dans des contextes de stress environnementaux variables. Comprendre les sites de SUMOylation spécifiques des HSFs sous stress thermique est essentiel pour déchiffrer les mécanismes précis par lesquels cette modification post-traductionnelle régule la réponse au stress. Ces sites de SUMOylation peuvent affecter la stabilité des HSFs, leur capacité à former des complexes protéiques avec d'autres co-régulateurs, ainsi que leur recrutement sur les promoteurs des gènes de choc thermique. Les résultats de telles études pourraient révéler des motifs de SUMOylation caractéristiques associés à des fonctions spécifiques des HSFs, offrant ainsi des informations cruciales pour la compréhension des réseaux de régulation sous-jacents. Enfin, cette problématique a des implications directes pour la biologie des plantes et la recherche en

agronomie. Comprendre la localisation des sites de SUMOylation et la capacité des différentes HSFs à être SUMOylées en présence ou en absence de SIZ-1 et d'ADN dans des conditions de stress thermique pourrait fournir des pistes pour améliorer la résistance des cultures aux conditions environnementales adverses. En somme, cette recherche contribue à élucider les mécanismes moléculaires complexes qui sous-tendent la capacité des plantes à s'adapter et à survivre dans des environnements changeants, ce qui est crucial pour relever les défis liés aux changements climatiques.

1.6 Hypothèse de l'étude

Notre hypothèse est que des altérations dans l'abondance et la localisation des sites de SUMOylation, bien qu'elles puissent initialement induire des perturbations dans les réseaux protéiques, contribuent à une adaptation à long terme des plantes à de nouveaux environnements. La SUMOylation, en tant que modification post-traductionnelle cruciale, serait ainsi impliquée dans la régulation fine des protéines et des voies cellulaires. Les changements dynamiques dans les motifs de SUMOylation pourraient jouer un rôle essentiel dans la réponse adaptative des plantes aux fluctuations environnementales. À court terme, des modifications dans l'abondance des sites de SUMOylation pourraient entraîner des instabilités dans les réseaux protéiques. En modifiant la SUMOylation de protéines clés impliquées dans la réponse au stress ou la signalisation, les plantes pourraient rapidement ajuster leur physiologie pour faire face aux défis environnementaux immédiats. Ces ajustements pourraient se traduire par des changements dans la stabilité des protéines, leur localisation subcellulaire, ou leur activité fonctionnelle. Cependant, à plus long terme, cette variabilité dans les motifs de SUMOylation pourrait fournir une souplesse adaptative aux plantes. En diversifiant les sites de SUMOylation, les plantes pourraient avoir la capacité d'explorer un large éventail de réponses adaptatives sans subir des changements génétiques permanents. Ainsi, les plantes peuvent rapidement adapter leur physiologie et leur métabolisme pour faire face aux stress environnementaux fluctuants tout en préservant leur intégrité génomique. L'adaptation à long terme pourrait également impliquer des ajustements dans la localisation des sites de SUMOylation. Par exemple, la redistribution des sites de SUMOylation vers des protéines clés régulant la réponse au stress pourrait offrir une mémoire

cellulaire, permettant aux plantes de s'adapter plus efficacement à des stress récurrents. Cette redistribution pourrait également faciliter la régulation de nouvelles voies métaboliques ou la modulation de l'interaction entre différentes protéines en réponse à des changements environnementaux prolongés. En conclusion, notre hypothèse suggère que la variabilité dans les motifs de SUMOylation, bien qu'elle puisse initialement perturber les réseaux protéiques, pourrait être un mécanisme évolutif permettant aux plantes de s'adapter rapidement et de manière flexible à des conditions environnementales changeantes. Des études approfondies sur les motifs de SUMOylation et leur rôle dans l'adaptation des plantes pourraient ainsi fournir des perspectives cruciales sur la plasticité des réponses cellulaires et métaboliques dans un contexte évolutif.

1.7 Objectif de l'étude

L'objectif principal de cette étude sur la SUMOylation des facteurs de choc thermique est de dévoiler les mécanismes moléculaires sous-jacents à la régulation fine de la réponse au stress chez les plantes, en se concentrant sur l'influence spécifique de la SUMOylation. Face aux défis climatiques et environnementaux, le stress thermique est l'un des facteurs les plus préoccupants pour la survie des plantes, et la compréhension détaillée de la régulation des facteurs de choc thermique à travers la SUMOylation pourrait révéler des pistes importantes pour améliorer la résilience des cultures. L'approche bio-informatique joue un rôle essentiel dans cette étude, en exploitant des logiciels spécialisés pour prédire et identifier les sites potentiels de SUMOylation sur les facteurs de choc thermique. Ces outils bio-informatiques analysent les séquences protéiques à la recherche de motifs caractéristiques impliqués dans la SUMOylation, tels que le consensus ψ KxE/D. Cette approche permet une identification rapide des sites de SUMOylation potentiels, offrant ainsi une base solide pour les étapes expérimentales ultérieures. L'approche biochimique complète la démarche bio-informatique en se concentrant sur la purification des facteurs de choc thermique pour une analyse approfondie de leur modification post-traductionnelle par SUMO. Les techniques de chromatographie sont couramment utilisées pour isoler et purifier les protéines d'intérêt à partir d'échantillons cellulaires. La purification des facteurs de choc thermique permet d'obtenir des échantillons hautement spécifiques pour des

analyses ultérieures, facilitant ainsi l'identification des sites de SUMOylation. Cette combinaison d'approches bio-informatique et biochimique offre une perspective holistique pour comprendre la SUMOylation des facteurs de choc thermique. En reliant les prédictions bio-informatiques aux données expérimentales obtenues par des méthodes biochimiques, cette étude aspire à identifier de manière fiable les sites de SUMOylation spécifiques, à élucider les motifs préférentiels de SUMOylation et à dévoiler les conséquences fonctionnelles de ces modifications pendant le stress thermique. En résumé, l'objectif de cette étude est d'intégrer des approches complémentaires, allant de la prédiction bio-informatique à l'analyse biochimique, pour déchiffrer les mécanismes de la SUMOylation sur les facteurs de choc thermique. Ces connaissances approfondies pourraient non seulement contribuer à notre compréhension fondamentale de la réponse au stress thermique chez les plantes, mais également ouvrir des perspectives pour des applications biotechnologiques visant à renforcer la résistance des cultures aux conditions environnementales adverses.

CHAPITRE 2

MATÉRIEL ET MÉTHODES

2.1 Bio-informatique

2.1.1 AlphaFold

L'utilisation d'AlphaFold (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) via la plateforme *Colaboratory* de Google (AlphaFold Colab version 1.2 et AlphaFold version 2.0) a permis la prédiction de la structure des protéines. AlphaFold effectue ses prédictions à partir de la séquence primaire des protéines. Ces séquences sont obtenues à partir de la base de données Uniprot (<https://www.uniprot.org/>) sous le format FASTA. AlphaFold (version 2.0) se concentre sur la prédiction des structures monomériques, toutefois, elle peut aussi prédire des structures oligomériques dont des structures trimériques. Une version améliorée, AlphaFold-Multimer, est spécifiquement conçue pour prédire les structures des oligomères et d'autres complexes protéiques (Bryant et al., 2022). Pour chaque séquence de protéine présente dans le fichier FASTA, un alignement de séquences est fait par l'outil 2 avec les paramètres de défauts et celui-ci sera utilisé comme fichier d'entrée par AlphaFold. AlphaFold génère de 1 à 5 modèles tridimensionnels avec des variations mineures de structure sous forme pdb qui vont être lus par le logiciel PyMOL pour visualiser et analyser les structures prédites des protéines. Chacun de ces modèles est classé en fonction de sa probabilité d'être une structure précise et correcte pour la protéine. Une fois qu'AlphaFold génère le modèle initial de la structure de la protéine, il est soumis à une étape de relaxation afin d'améliorer la précision de la structure prédite ainsi que de réduire les tensions et les erreurs de prédictions.

AlphaFold génère également des données PAE (Predicted Alignment Error) pour chacune des structures correspondant à la protéine d'intérêt afin d'évaluer la qualité des prédictions de structures protéiques. Les données de PAE sont visualisées sous forme de graphiques. L'axe

horizontal et l'axe vertical du graphique PAE correspondent aux numéros de résidus de la séquence de la protéine. Chaque point sur cet axe correspond à un résidu de la protéine et est généralement numéroté de manière séquentielle. La valeur attribuée à chaque position X, Y représente l'évaluation de l'erreur de positionnement du résidu X, sous l'hypothèse que la structure réelle et la structure prédite seraient alignées au niveau du résidu Y. Plus la valeur PAE est élevée, plus la prédiction est éloignée de la structure de référence. Les valeurs PAE sont mesurées en angströms (Å). On examine les graphiques PAE pour identifier des zones d'interaction entre différentes régions de la même protéine ou entre une protéine donnée et d'autres composants du complexe protéique.

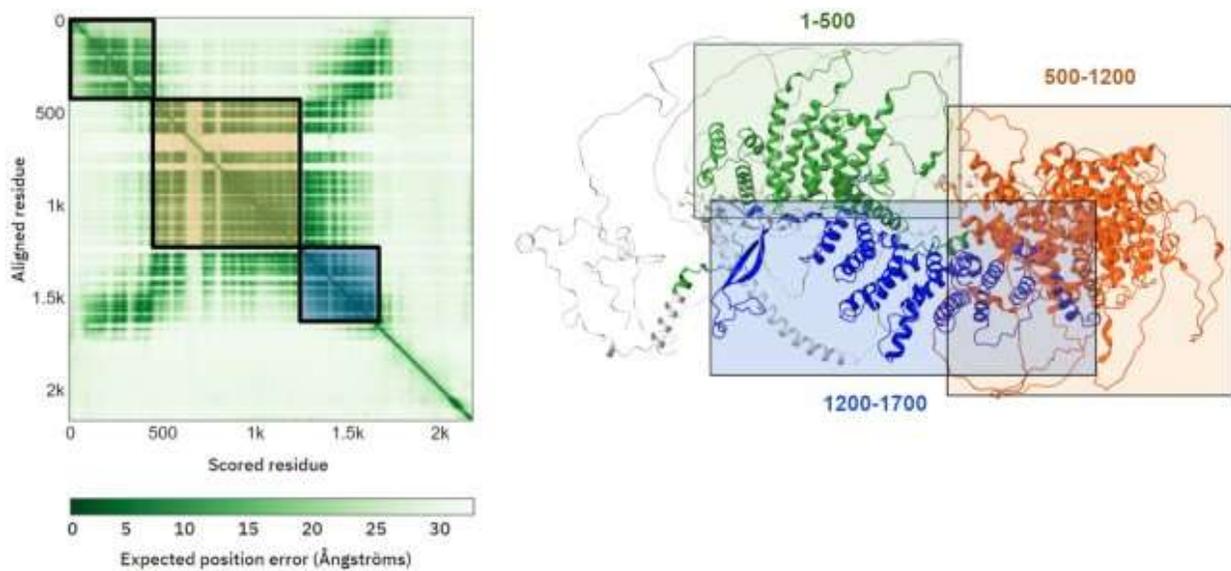


Figure 2.1 Visualisation des erreurs alignées prévues (PAE). La figure de gauche représente les interactions entre les domaines des protéines en 2D. La figure à droite est la modélisation de la protéine en 3D avec l'identification de trois régions distinctes. AlphaFold prédit que les résidus 1 à 500 (vert) adoptent une structure rigide et interagissent avec les résidus 1200 à 1700 (bleu) qui forment également une structure rigide. En revanche, les résidus 500 et 1200 (orange) n'ont que peu de contact avec les domaines formés par les résidus 1 à 500 et 1200 à 1700. (Varadi et al., 2021)

AlphaFold génère aussi des données de pLDDT (Predicted Local Distances Difference Test) qui estiment la fiabilité des prédictions à l'échelle locale. Enfin, l'analyse des fichiers PDB générée par AlphaFold a été faite à l'aide de PyMOL, un logiciel de visualisation de structures 3D développé par Schrodinger (version 2.5).

2.1.2 Les logiciels JASSA et GPS-SUMO

Les logiciels JASSA (<http://www.jassa.fr>) et GPS-SUMO (<https://sumo.biocuckoo.cn/>) permettent de prédire les sites de SUMOylation à partir de la séquence primaire des protéines. Les paramètres d'entrée pour le logiciel JASSA peuvent consister en la séquence de la protéine au format FASTA ou en son numéro d'accèsion Uniprot (disponible sur <https://www.uniprot.org/>). Par exemple le numéro d'accèsion pour la protéine HSFB2b est Q9T0D3. Les paramètres de sortie pour ce logiciel peuvent être ajustés en fonction des résultats recherchés. Dans notre cas, les seuls paramètres modifiés sont les résultats de SUMO et SIM qui vont être mis avec des résultats de seuil (*cut-off*) élevés, utilisés pour filtrer des interactions SUMO-SIMs ou les sites de SUMOylation les plus probables. En ce qui concerne le logiciel GPS-SUMO, il suffisait d'entrer la séquence de la protéine. Dans la section de limite (threshold), il a été nécessaire de définir la valeur haute pour la SUMOylation et pour l'interaction SUMO afin d'obtenir les résultats souhaités. L'utilisation d'un seuil de prédiction élevé dans la prédiction de site de SUMOylation, est généralement pour s'assurer une plus grande fiabilité des résultats. Ces différents logiciels utilisent des données expérimentales de sites de SUMOylation pour entraîner un modèle de prédiction basé sur des caractéristiques pertinentes extraites de séquences de protéines (Palacios et al., 2024). Après l'entraînement du modèle, un seuil est déterminé en optimisant la précision, la sensibilité et la spécificité du modèle, généralement à l'aide de techniques de validation croisée. Ce seuil est ensuite utilisé pour prédire avec précision les sites de SUMOylation dans de nouvelles séquences de protéines.

2.2 Principaux réactifs

Les plasmides codants pour les différentes protéines HSFs ont été obtenus auprès de l'Arabidopsis Biological Resource Center (ABRC). Le tableau 2.1 indique les codes de plasmides pour chacun des facteurs de choc thermique chez la banque de ABRC. Les bactéries compétentes DH5 α et les BL21 DE3 CodonPlus proviennent à l'origine de la compagnie Novagen.

Tableau 2.1 Plasmides obtenus de chez ABRC contenant la séquence codante pour les différentes protéines HSFs utilisées dans le cadre de cette étude.

Protéines	Numéro de stock ABRC
HSFA1a	DKLAT4G17750
HSFA2	DKLAT2G26150
HSFA1e	U85727
HSFA7a	U23517
HSFA7b	PENTR221
HSFA9	U18576
HSFB1	U16604
HSFB2b	U24153
HSFB3	U86661
HSFC1	U14335

2.3 Clonage moléculaire

Le clonage moléculaire est une technique fondamentale en biologie moléculaire qui permet de copier et d'amplifier des fragments d'ADN spécifiques. Il implique l'insertion d'un fragment d'ADN dans un vecteur, tel qu'un plasmide bactérien, pour produire une molécule d'ADN recombinant. Cette molécule recombinante peut ensuite être introduite dans des cellules hôtes, généralement des bactéries, où elle est répliquée pour produire de multiples copies du fragment d'ADN. Dans le clonage moléculaire, les oligonucléotides jouent un rôle essentiel en tant qu'amorces pour l'amplification et l'insertion de fragments d'ADN. Leur sélection est basée sur la spécificité de la séquence cible, la température de fusion, et l'évitement des régions répétitives ou instables. De plus, l'ajout de séquences d'amorce supplémentaires et l'inclusion de sites de restriction peuvent faciliter le processus de clonage.

Les tableaux 2.2 indique certaines des propriétés des 21 HSFs d'Arabidopsis avec, en gris les protéines HSFs utilisées dans le cadre de cette étude.

Tableau 2.2 Caractéristiques des protéines utilisées dans cette étude

Protéine	Numéro d'accèsion	Locus	Taille (a.a)	MW (Da)	Point isoélectrique
HSFA1a	P41151	At4g17750	495	55744	5,12
HSFA1b	O81821	At5g16820	481	53576	4,80
HSFA1d	Q9LQM7	At1g32330	485	54524	4,87
HSFA1e	Q9SCW5	At3g02990	468	51962	4,82
HSFA2	O80982	At2g26150	345	39114	5,13
HSFA3	Q8GYY1	At5g03720	412	46458	5,32
HSFA4a	O49403	At4g18880	401	46244	5,33
HSFA4c	Q9FK72	At5g45710	345	39647	5,61
HSFA5	Q94BZ5	At4g13980	466	52355	6,16
HSFA6a	Q1PDN3	At5g43840	282	33155	5,51
HSFA6b	Q9LUH8	At3g22830	406	46735	4,94
HSFA7a	Q9SV12	At3g51910	272	31776	6,03
HSFA7b	Q9M1V5	At3g63350	282	32623	5,13
HSFA8	Q9S7U5	At1g67970	374	42629	4,84
HSFA9	Q9LVW2	At5g54070	331	38146	5,49
HSFB1	Q96320	At4g36990	284	31327	6,27
HSFB2a	Q9SCW4	At5g62020	299	34050	5,84
HSFB2b	Q9T0D3	At4g11660	377	39705	4,92
HSFB3	O22230	At2g41690	244	28307	5,26
HSFB4	Q9C635	At1g46264	348	39615	7,30
HSFC1	Q9LV52	At3g24520	330	37714	5,82

2.3.1 Amplification par PCR (Polymerase Chain Reaction)

Le tableau 2.3 correspond à la liste des séquences d'oligonucléotides utilisées pour réaliser leur clonage moléculaire par la méthode GIBSON des différents HSFs.

Tableau 2.3 Séquences des oligonucléotides employées pour amplifier les ADNc de facteurs de choc thermique

Protéine	Séquences
HSFA1a	Sens : GAAGATCCGGACGGCGAAGAG <u>GGATCC</u> ATGGACGGTGTACCGGCGGAG Anti-sens : CGGAGCTCGAATTCGGATCTCA <u>GGATCC</u> CTAGTGTCTGTTTCTGATGTGAGAGAC
HSFA2	Sens : GAAGATCCGGACGGCGAAGAG <u>GGATCC</u> ATGGAAGAAGTAAAGTGGAAATGGAGGAAG Anti-sens : CGGAGCTCGAATTCGGATCTCA <u>GGATCC</u> TTAAGGTTCCGAACCAAGAAAACCCATTG
HSFA1e	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGGGAACGGTTTTCGAATCTGTAGCG Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> TCATTTTCTGAGAGCATCTGATGTGAGAAG
HSFA7a	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGATGAACCCGTTTCTCCCGGAAGG Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> TTAGGAGGTGGAAGCCAAACTCTCATCACTAAG
HSFA7b	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGGACCCGTCGTCAAGCTCCAG Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> CTAATCTTGCTTACATTCGCCTCTTCTTCGTC
HSFA9	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGACGGCAATTCCAAACGTCGTCG Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> CTACTCTATCTATCCTCCCATTACATCC
HSFB1	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGACGGCTGTGACGGCGGC Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> TTAGTTGCAGACTTTGCTGCTTTTCCACAAC
HSFB2b	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGCCGGGGGAACAAACCGGAG Anti-sens : GCTTGTCGACGGAGCTCGAATTC <u>GGATCC</u> TCATTTTCCGAGTTCAAGCCACGACCC
HSFB3	Sens :

	GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGGAAGATGCTGGTGAACATTTACGGTG Anti-sens : GCTTGTTCGACGGAGCTCGAATTC <u>GGATCC</u> TCATTCAAGTTTTACTCCGAACAACCTTAAGCC
HSFC1	Sens : GGTACCGAAAACCTTGTATTTCCAG <u>GGATCC</u> ATGGAGGACGACAATAGTAACAACAACAAC Anti-sens : GCTTGTTCGACGGAGCTCGAATTC <u>GGATCC</u> CTAAAAGCCACCTCGAAACAGAGAAAATGG

BamHI est une enzyme de restriction utilisée pour cliver l'ADN. Les séquences d'oligonucléotides sont représentées au tableau 2.3 en caractère gras et soulignées contiennent toutes un site de reconnaissance de BamHI, 5'-G[^]GATCC-3'. Lorsque BamHI reconnaît ce site, il coupe l'ADN entre les deux guanines (G).

Chaque réaction d'amplification par PCR a été réalisée en mélangeant 1 ng de l'ADN, 4 µM de chaque amorce, 200 µM dNTP, 10 µg/mL polymérase Pfu et du tampon de réaction comprenant 50 mM Tris-HCl pH 8.0, 20 mM de KCl, 15 mM (NH₄)₂SO₄ et 0.14mg/mL BSA (concentrations finales). L'amplification a été faite selon le cycle suivant en utilisant le thermocycleur T100 (BioRAD); la dénaturation à 95 °C pendant 2 minutes, 29 cycles d'hybridation à 50 °C pour une durée de 20 secondes et l'élongation à 72 °C à environ 3 minutes. Enfin l'élongation finale à 72 °C pendant 20 minutes. Au début de chaque cycle, une étape de dénaturation de 20 secondes à 95 °C a été faite. Les échantillons sont ensuite chargés sur un gel d'agarose 1 % afin de visualiser si l'amplification a fonctionné. La migration est faite à 120 V pour une durée de 22 minutes. Le standard est un mélange 1 : 1 (v/v) de EZ Load 1 kb Molecular Ruler et EZ Load 100 bp Molecular Ruler pour obtenir une échelle moléculaire de 100 pb à 15 000 pb. Après migration, le gel a été analysé sous un transilluminateur à l'aide de l'appareil ChemiDoc (BioRad) et le logiciel ImageLab par BioRad, version 5.0 built 18.

2.3.2 Clonage moléculaire par Gibson

La réaction de clonage par Gibson a été faite en mélangeant ensemble le plasmide préalablement digéré à l'aide de l'enzyme BamHI et l'insert en présence du tampon Gibson comprenant 1.00 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 5% PEG-8000, 10 mM DTT, 1 mM NAD, 0.5 mM dGTP, 0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dTTP ainsi que de 0.0008 U/μL de T5 exonucléase, 0.05 U/μL Phusion polymérase et 2 U/μL TAQ ligase (concentrations finales). Le calcul pour déterminer la quantité d'insert à utiliser pour une quantité de vecteur dépend de plusieurs facteurs, notamment la taille du vecteur (dans notre cas, il s'agit du vecteur pRSF-Trx), de sa concentration, de la taille de l'insert et du rapport molaire souhaité entre l'insert et le vecteur (3 :1 dans notre cas).

$$ng\ d'insert = \frac{((ng\ de\ vecteur\ x\ taille\ en\ kb\ de\ l'insert)\ x\ 3)}{taille\ en\ kb\ du\ vecteur}$$

Une incubation d'une heure à 50 °C permet à l'exonucléase de digérer des nucléotides dans l'extrémité 5' du vecteur et des inserts. L'exonucléase, qui est thermosensible s'inactive toutefois rapidement sous l'effet de la température élevée. La ligase et la polymérase restent fonctionnelles à cette température, et permettent ensemble d'insérer la séquence codante des protéines HSFs dans le vecteur.

2.3.3 Transformation bactérienne DH5α

Les bactéries compétentes *Escherichia coli* (*E. coli*) de type DH5α ont été utilisées afin de pouvoir obtenir des colonies bactériennes contenant les plasmides d'intérêt (vecteurs pRSF-Trx-HSF). 20 μL de bactéries chimio-compétentes ont été incubés sur glace avec 2 ng d'ADN plasmidique (produit de la réaction Gibson) pendant 30 minutes suivi d'un choc thermique de 45 secondes à 42 °C et finalement d'une incubation d'une minute sur glace. Un volume de 300 μL de milieu Luria-Bertani (LB) (10 g/L NaCl, 5 g/L extrait de levure et 10 g/L peptone) a été ensuite ajouté aux bactéries en condition stérile et les bactéries sont incubées à 37 °C avec

agitation à 350 rpm pendant 1 heure. Cette étape de récupération permet d'exprimer le gène de résistance à l'antibiotique avant l'étalement sur gélose. Après cette incubation, la culture bactérienne a été étalée sur une gélose LB-Agar-Kanamycine en condition stérile suivie d'une incubation d'environ 18 heures à 37 °C. Les colonies ont été validées par PCR sur colonies et sélectionnées en condition stérile et incubées dans 2 mL de milieu Super Broth (SB) (5 g/L NaCl, 20 g/L extrait de levure, 32 g/L peptone) avec 50 µg/mL de kanamycine pour extraire le vecteur par Mini Prep (voir section 1.2.4).

La validation par PCR de colonies est une technique pour identifier des colonies de bactéries contenant des séquences d'ADN d'intérêt. Dans notre cas, cette technique nous a permis d'identifier les colonies qui possèdent un plasmide pRSF-Trx-HSF. Dans des tubes à PCR, il a fallu homogénéiser une colonie bactérienne dans une solution de lyse (10 mM Tris-HCl pH 8.0, 1 mM EDTA et Triton X-100 0.1 %). Ces échantillons ont été incubés sur le Eppendorf Thermomix C à 50 °C pendant 10 minutes. Par la suite, 2 µL de l'échantillon a été prélevé et 4 µM de chaque amorce soit ACYC (5'-GGATCTCGACGCTCTCCCT-3') et T7- (5'-TAATACGACTCACTATAGGG-3') a été rajouté. La présence de l'insert entraîne un changement de la taille de l'amplicon par rapport au vecteur dépourvu d'insert et au tampon de réaction comprenant 50 mM Tris-HCl pH 8.0, 20 mM de KCl, 15 mM (NH₄)₂SO₄, 1.2 mM MgSO₄ et 0.14 mg/mL BSA (concentrations finales) et a été mise sur le thermocycleur avec les mêmes paramètres que pour l'amplification par PCR (section 2.3.1) afin d'amplifier la séquence précise du plasmide d'ADN recombinant. Les échantillons sont ensuite chargés sur un gel d'agarose 1%. La migration est faite à 120 V pour une durée de 22 minutes. Le gel a été analysé à l'aide de l'appareil ChemiDoc par BioRad et le logiciel ImageLab par BioRad, version 5.0 built 18.

2.3.4 Isolation de l'ADN plasmidique (MiniPreps)

Une trousse de MiniPreps (Proméga; Numéro de catalogue A1460) a été utilisé pour l'extraction et la purification des plasmides (Figure 2.1) selon les instructions du fabricant. 1.5 mL de la culture bactérienne a été mis dans un Eppendorf et centrifugé à 6000 rpm pendant 5 minutes. Le culot a été resuspendu avec 250 µL de solution de suspension cellulaire (Cell Suspension

Solution) et 250 μL de solution de lyse cellulaire (Cell Lysis Solution) ont été ajoutés et la solution été mélangée 4x par inversion. 10 μL de solution de protéase alcaline (Alkaline Protease Solution) ont été ajoutés et le tube a été mélangé 4x par inversion. Ce mélange est incubé pendant 5 minutes à température pièce. 350 μL de solution de neutralisation (Neutralization Solution) ont été ajoutés et le tube a été mélangé 4x par inversion. Ce mélange est ensuite centrifugé à une vitesse maximale de 8000 g pendant 10 minutes. Le lysat a été versé dans une colonne de silice qui a été préalablement placée dans un tube de 2 mL. Ceci a été centrifugé à 8400 rpm pendant 1 minute. L'éluat écoulé a été jeté. 750 μL de solution de lavage (Wash Solution) ont été ajoutés suivis d'une centrifugation à 8400 rpm pendant 1 minute et, de nouveau, le liquide écoulé a été jeté. 250 μL de solution de lavage sont ajoutés suivis d'une centrifugation à 8400 rpm pendant 2 minutes et le liquide écoulé est jeté. La colonne a finalement été transférée dans un nouveau tube Eppendorf. 60 μL d'eau dépourvue de nucléases a été ajouté avant une centrifugation à 8400 rpm pendant 1 minute. La quantification des plasmides a été quantifiée à l'aide du NanoDrop (Thermo Scientific).

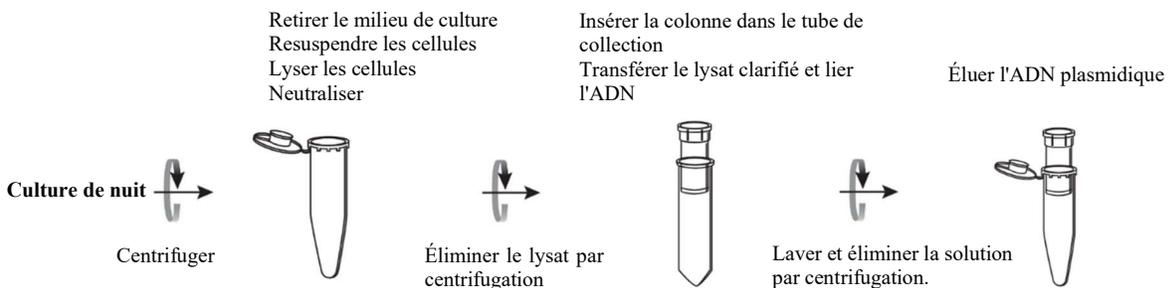


Figure 2.1.4 Système de purification d'ADN par Minipreps

Une approche de digestion enzymatique a ensuite été utilisée pour vérifier la présence de l'insert d'intérêt dans les plasmides isolés. Le mélange réactionnel consiste 2 μL du tampon 10X CutSmart, 0.4 μL BamHI, 10 μL du vecteur (produit de la Mini Prep) et 7.6 μL d'eau déionisée de haute pureté ont été ajoutés. Cette réaction a été incubée 37 °C pendant 1 heure. Par la suite, cette réaction a été mise sur un gel d'agarose 1% de la même manière qu'à la section 2.3.1 afin de pouvoir voir l'insert et le vecteur. La digestion par BamHI induit des coupures spécifiques

dans le plasmide, suivies de la ligation de notre fragment d'ADN d'intérêt. Parallèlement, il a aussi été possible d'envoyer les échantillons pour le séquençage Sanger à Génome Québec afin de pouvoir déterminer que la séquence correspond à notre protéine d'intérêt. Les amorces ACYC et T7m ont été utilisés pour le séquençage.

2.4 Production et purification des protéines

2.4.1 Transformation bactérienne BL21 DE3 CodonPlus

Les bactéries compétentes *E. coli* de type BL21 DE3 CodonPlus ont été utilisées 20 µL de bactéries compétentes ont été incubés sur glace avec 2 ng du vecteur digéré (produit de la Mini Prep) pendant 30 minutes suivi d'un choc thermique de 45 secondes à 42 °C et finalement d'une incubation d'une minute sur glace. Un volume de 300 µL de milieu LB (Luria-Bertani) (10 g/L NaCl, 5 g/L extrait de levure et 10 g/L peptone) a été ajouté en condition stérile et les bactéries ont été incubées à 37 °C avec agitation à 350 rpm pendant 1 heure de temps. Le volume de la culture bactérienne a été mis sur pétri LB-Agar-Kanamycine-Chloramphénicol en condition stérile suivi d'une incubation d'environ 18 heures.

2.4.2 Culture et expression des protéines

Des bactéries provenant d'une colonie bactérienne BL21 DE3 CodonPlus ont étéensemencées dans 100 mL de milieu de culture LB-Kanamycine-Chloramphénicol contenant 2% glucose stérile puis incubé à 32 °C sous agitation à 200 rpm pendant 16 heures. Cette pré-culture a ensuite été permis d'ensemencer à 3 L de milieu de culture SB- Kanamycine-Chloramphénicol qui ont été incubée à 37 °C sous agitation à 200 rpm. Lorsque la densité optique (DO) a atteint entre 0.8 et 1.0 à 600 nm, l'expression protéique a été induite avec 0.3 mM d'isopropyl β-D-1-thiogalactopyranoside (IPTG). La culture a été ensuite incubée à 18 °C à 200 rpm pendant environ 16 à 18 heures. Les bactéries ont été récupérées par centrifugation à 6000 rpm pendant 40 minutes à 4 °C. Le culot bactérien a été suspendu dans 20% de sucrose 50 mM Tris-HCl pH 8.0

avec une concentration d'environ 0.1mg/mL de lysozymes et congelé dans le -20°C à l'aide de l'azote liquide. Le culot bactérien a été décongelé et complété à 50 mL avec du tampon A (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM Imidazole et 5 mM β -mercaptoéthanol). Les bactéries ont été lysées sur glace au moyen d'un sonicateur Branson Sonifier 450 avec une puissance de 5 et un cycle de travail à 50 %. La sonication a été effectuée en 3 cycles comportant chacun 3 minutes de sonication et 3 minutes de pause. Les bactéries ont été ensuite culottées en centrifugeant à 12 298 g pendant 40 minutes à 4 °C. Le lysat total des protéines a été récupéré pour la purification des protéines.

2.4.3 Chromatographie par affinité

La chromatographie par affinité est une méthode de purification des protéines basées sur leurs interactions spécifiques avec des ligands d'affinité. Le principe de cette chromatographie repose sur l'utilisation de résine spéciale qui fonctionne avec des ligands d'affinité. La résine utilisée est composée de groupement chimique avec les ions de nickel (Ni^{2+}) (*Cytiva*) liés pour la purification des protéines étiquetées 6 histidines (His-tag). La première étape a été de préparer et conditionner la résine dans une colonne chromatographique avec de l'eau et du tampon A (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole et 5 mM β -mercaptoéthanol) environ 6 fois le volume de colonne. Ensuite, l'échantillon contenant la protéine d'intérêt a été chargé sur la colonne. Les impuretés passent à travers la colonne tandis que la protéine d'intérêt se lie spécifiquement aux ligands d'affinité. Suite au chargement de l'échantillon contenant la protéine, la colonne a été lavée avec environ 75 mL de tampon A afin de pouvoir éliminer les impuretés non liées tout en maintenant la protéine d'intérêt sur la colonne. La protéine d'intérêt a ensuite été libérée de la colonne en utilisant un tampon d'élution spécifique, le tampon B (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 0.5 M Imidazole et 5 mM β -mercaptoéthanol) qui perturbe l'interaction entre la protéine et les ligands d'affinité. Cela a permis de récupérer la protéine purifiée dans des Eppendorf à fraction de 1 mL pour un volume total de 5 mL. Ces échantillons sont ensuite quantifiés par NanoDrop et analysés par SDS-PAGE. Les différentes fractions contenant le plus de la protéine d'intérêt ont été combinées par la suite.

2.4.4 Clivage à la TEV

Le clivage à la protéase Tobacco Etch Virus (TEV) est une méthode de clivage des protéines purifiées de manière spécifique au niveau d'un site de reconnaissance. Plus précisément, la protéase TEV clive au niveau de la séquence d'acide aminé : ENLYFQ. Ce clivage permet de séparer la protéine d'intérêt de l'étiquette d'affinité, soit la 6x histidine. Après purification de la protéine par chromatographie d'affinité, un volume de 100 μ L de protéase TEV à une concentration de 150 μ M a été incorporé à 5 mL de la protéine d'intérêt combinée. Cet échantillon sera incubé dans le réfrigérateur à 4 °C pour une durée d'environ 18 heures.

2.4.5 Tamis moléculaire

La chromatographie par exclusion stérique est une méthode de purification des protéines basée sur leur poids moléculaire. Elle permet notamment d'éliminer les petites molécules contaminantes ainsi que les agrégats protéiques. La première étape a été de préparer et conditionner la colonne S200 ainsi d'équilibrer avec le tampon approprié, *GF Buffer* (150 mM NaCl, 20 mM Tris-HCl pH 8.0 et 5 mM β -mercaptoéthanol) en utilisant un débit de 1 mL par minute. Ensuite, l'échantillon de 5 mL contenant la protéine d'intérêt a été chargé sur la colonne à l'aide d'une seringue tout en évitant d'introduire de l'air dans la colonne, les petites molécules pénètrent dans les pores et sont retardées tandis que les protéines les plus grosses circulent plus rapidement. Une analyse par gel SDS-PAGE a permis d'identifier les fractions contenant la protéine d'intérêt. Celles-ci ont été combinées, concentrées et congelées à l'aide d'un bain d'azote liquide, et conservées à -80 °C jusqu'à leur utilisation.

2.4.6 Électrophorèse sur gel SDS-PAGE

L'électrophorèse sur gel de polyacrylamide en présence de dodécyl sulfate de sodium (SDS-PAGE) est une méthode utilisée pour séparer les protéines en fonction de leur poids moléculaire.

Des gels de 12 % ont été utilisés afin de pouvoir visualiser les échantillons issus de la purification des protéines par chromatographie par affinité ainsi que le tamis moléculaire. La migration des protéines est faite à 180 V pendant 1 heure. La visualisation des protéines sur gel a été faite à l'aide de l'appareil ChemiDoc par BioRad et le logiciel ImageLab par BioRad version 5.0 build 18. Les gels ont été visualisés de deux façons; la première étant avec du TCE (2,2,2-trichloroethanol) qui réagit avec les résidus tryptophane des protéines. La deuxième étant la coloration avec du bleu de Coomassie de permettant de visualiser la présence des petites quantités de protéines sur le gel. Le gel est immergé dans une solution de coloration de Coomassie composée de 40% de méthanol, 10% d'acide acétique et de 0.1% de Coomassie R-250 pendant environ 15 minutes et ensuite décoloré avec une solution composée de 10% de méthanol et de 10% d'acide acétique afin de pouvoir visualiser les bandes de protéines sur le gel.

2.5 Essai retard sur gel

Les essais de retard sur gel sont utilisés pour étudier la mobilité et la liaison de différents ADN à la protéine d'intérêt. Des concentrations de 0 μM , 0.3 μM , 0.7 μM , 1.5 μM , 3.1 μM , 6.2 μM , 12.5 μM et 25 μM de protéines HSF ont été préparées dans du tampon *GF Buffer* et 1 μM d'ADN non-spécifique et d'ADN spécifique est rajouté dans chacun des tubes de concentrations avec du colorant de charge, Orange G 1X (0.2 mg/mL de Orange G et 3 % (v/v) de glycérol) (voir tableau 2.4). Ces échantillons ont été chargés sur un gel d'agarose 1 %. La migration a été faite à 120 V pour une durée de 22 minutes. Le gel a par la suite été analysé à l'aide de l'appareil ChemiDoc par BioRad et le logiciel ImageLab par BioRad, version 5.0 built 18.

Les protéines se lient à des séquences spécifiques d'ADN, appelées éléments de réponse à la chaleur (HSE – Heat Shock Elements) pour activer l'expression génique. La séquence consensus de HSE présente trois occurrences de la séquence « AGAAT/C » ou « A/GTTCT » dans son complément inverse et celle-ci est démontrée en gras et souligné dans le tableau ci-dessous.

Tableau 2.4 Séquence nucléotidique de l'élément de réponse au stress thermique (HSE) impliqué dans la liaison et l'activation protéique

HSE	Séquence
HSF20nt (non-spécifique)	Sens : CGC <u>AGAATATTCTAGAACGC</u> Anti-sens : GCG <u>TTCCTAGAATATTCT</u> GCG
SIZ20nt (non-spécifique)	Sens : AAAATAAAACTAGCTATTTT Anti-sens : AAAATAGCTAGTTTTATTTT
SIZHSF40nt (spécifique)	Sens : AAAATAAAACTAGCTATTTTTCGC <u>AGAATATTCTAGAACGC</u> Anti-sens : GCG <u>TTCCTAGAATATTCT</u> GCGAAAATAGCTAGTTTTATTTT
SIZHSF40nt2 (spécifique)	Sens : AAAATAAAACTAGCTATTTTTCG <u>TTCCTAGAATATTCT</u> GCG Anti-sens : CGC <u>AGAATATTCTAGAACGC</u> AAAATAGCTAGTTTTATTTT
HSFSIZ40nt (spécifique)	Sens : CGC <u>AGAATATTCTAGAACGC</u> AAAATAAAACTAGCTATTTT Anti-sens : AAAATAGCTAGTTTTATTTTTCG <u>TTCCTAGAATATTCT</u> GCG
HSFSIZ40nt2 (spécifique)	Sens : GCG <u>TTCCTAGAATATTCT</u> GCGAAAATAAAACTAGCTATTTT Anti-sens : AAAATAGCTAGTTTTATTTTTCGC <u>AGAATATTCTAGAACGC</u>

2.6 Essais enzymatiques

L'essai enzymatique de la SUMOylation est fait afin de pouvoir étudier la conjugaison de SUMO à des protéines d'intérêt. Pour commencer, il est important d'avoir le mélange réactionnel qui comprend la protéine d'étude soit la HSF (1 μM), la protéine SUMO (50 μM), les enzymes E1 (1 μM), E2 (1 μM), E3 (1 μM), le tampon de réaction (20 mM HEPES pH 8.0, 5 mM MgCl_2 , 125 mM NaCl et Tween 0.1 %), le DTT (3 mM) et l'ATP (2 mM) afin de pouvoir initier la réaction enzymatique. Le DTT (dithiothréitol) est utilisé dans les réactions enzymatiques pour réduire les ponts sulfure, protéger les protéines de l'oxydation, maintenir leur activité et améliorer leur solubilité. Les plasmides de la cascade enzymatiques (SUMO1-pRSFDuet1, SAE1-pRSF-Trx, SAE2-pRSF-Trx, SCE1-pRSF-Trx et SIZ-1-MBP) ont été purifiés par ma collègue Faustine Hutinet de la même façon que les facteurs de choc thermique. Le mélange de réactions a été incubé à 37 °C sur le Eppendorf ThermoMixer C. À 0, 5, 10, 20, 40 et 80 minutes, 60 μL d'échantillon ont été prélevés et mis dans 20 μL de Laemmli 1X. La composition de la solution Laemmli 1X est 62.5 mM Tris-HCl pH 6.8, 2% SDS (sodium dodecyl sulfate), 10% glycérol, 0.01% bromophénol bleu et 5% β -mercaptoéthanol. Cette solution a été utilisée pour diluer les échantillons de protéine afin de pouvoir visualiser ces derniers sur gel comme expliqués à la section 2.4.6.

CHAPITRE 3

RÉSULTATS

Les résultats de l'analyse bio-informatique des sites de SUMOylation, conjointement avec les études *in vitro* visant à caractériser la SUMOylation des facteurs de choc thermique (HSFs), fournissent des données cruciales pour comprendre la régulation de ces protéines essentielles dans la réponse au stress thermique chez *Arabidopsis thaliana*. Cette approche intégrée permet d'identifier des cibles potentielles pour la modulation de la réponse au stress thermique chez les plantes, ouvrant ainsi de nouvelles perspectives pour l'ingénierie génétique visant à améliorer la tolérance au stress thermique des cultures.

3.1 Analyse bio-informatique des sites de SUMOylation chez les HSFs par les logiciels JASSA et GPS-SUMO

Dans le but d'étudier la SUMOylation chez les différentes HSFs d'*Arabidopsis thaliana*, il est important d'identifier les lysines (K) pouvant être SUMOylées ou ayant une plus grande probabilité d'être SUMOylées chez ces protéines. L'étude des lysines conservées sera une des démarches importantes afin de pouvoir comprendre la régulation de ces protéines essentielles dans la réponse au stress thermique. Le tableau 3.1 représente les lysines candidates qui sont potentiellement SUMOylables avec leurs scores obtenus avec les logiciels JASSA (<http://www.jassa.fr>) et GPS-SUMO (<https://sumo.biocuckoo.cn/>). Ce tableau présente une classification des protéines HSFs avec une subdivision en fonction de leur classe, incluant les positions variées de lysines pouvant subir une SUMOylation.

Tableau 3.1 Sites de SUMOylation chez les protéines HSFs prédites au moyen des logiciels GPS-SUMO (<https://sumo.biocuckoo.cn/>) et JASSA (<http://www.jassa.fr>)

	HSF	Position de la lysine	Séquence	Score GPS-SUMO	Score JASSA	Publications
Classe A	*HSFA1a	K189	FGLEEEVEQLKRDKNVLMQEL	0.3340 (Faible)	Faible	Hirata et al., 2010
		K212	LRQQQQT ^T DNKLQVLVKHLQV	0.7167 (Moyenne)	Faible	
		K308	RNDSMMWNMMKTDDKY ^T PFLDG	0.5964 (Moyenne)	Faible	
	*HSFA1b	K155	FGIEEEVERL ^K RDKNVLMQEL	0.3320 (Faible)	Faible	Hirata et al., 2010
		K402	GDAISPEG ^E GKMNELLEGV ^P K	0.7004 (Moyenne)	Faible	
	*HSFA1d	K19	SDGGGDSM ^E TK ^S SPQPQ ^A AI	0.7594 (Fort)	Fort	Mascle et al., 2007
		K169	FGLEEEVERL ^K RDKNVLMQEL	0.2817 (Faible)	Faible	
		K251	ISDTSKRR ^F KRDGIVRN ^N DS	0.3694 (Faible)	Faible	Hirata et al., 2010
		K288	QAKAMFKQLMKME ^P YK ^T GDDG	0.8539 (Fort)	Faible	Chun et al., 2003
	*HSFA2	K5	-----ME ^E L ^K VEMEE ^E TVTF	0.9276 (Fort)	Fort	Hirata et al., 2010
		K167	YGF ^D GEVERL ^K R ^D HGVLVAEV	0.3751 (Faible)	Faible	
		K269	LLHDQ ^E FDRMKDDMEMLFAAA	0.3948 (Faible)	Faible	Cohen-Peer et al., 2010
		K315	DGNLEAALD ^V KVE ^D LVGSPLD	0.9187 (Fort)	Fort	Cohen-Peer et al., 2010
	HSFA3	K4	-----MS ^P KKDAVSK ^P TPI	0.4888 (Moyenne)	Faible	
		K273	QDSPTGGEV ^V KYEADD ^W ERLL	0.8130 (Fort)	Fort	Eloranta et al., 2002
		K333	MSF ^S PEGL ^I KQE ^E TTWSMG ^F	0.8063 (Fort)	Fort	
HSFA4c	K136	RSMEDQIERL ^K NEKEGL ^L LAEL	0.5864 (Moyenne)	Faible		
*HSFA5	K151	LSREKAAI ^E AKL ^L KFKQ ^K QV ^V	0.4440 (Moyenne)	Faible		
	K203	NPTFVK ^N FGK ^K VE ^Q LDISAY ^N	-	Faible		
HSFA6a	K97	RWEFANEH ^F LKGE ^R HLL ^L KN ^I K	0.7906 (Fort)	Faible		
	K158	VKTYLHL ^M E ^E KL ^K VTEV ^K QEM	0.3689 (Faible)	Faible		
	K165	MEEKL ^K VTEV ^K QEMMN ^F LL ^K	0.7794 (Moyenne)	Fort		
	K217	HGVEDNG ^K FV ^K AE ^P EEY ^G DD ^I	0.8608 (Fort)	Fort		
HSFA6b	K9	--MD ^P SFR ^F I ^K EE ^F PAG ^F SD ^S	0.9853 (Fort)	Fort		
	K221	TKMYLTL ^I E ^E KL ^K KTES ^K Q ^K Q	0.4497 (Moyenne)	Faible		
	K224	YLT ^L IEE ^K L ^K KTES ^K Q ^K Q ^M S	0.5153 (Moyenne)	Faible		
HSFA7a	K202	LHQLL ^K QR ^D K ^K I ^K ELED ^N ESA	0.5362 (Moyenne)	Faible		
HSFA7b	K225	AAMIDQAGL ^I KME ^E VEHL ^S SEL	0.8259 (Fort)	Fort		
	K280	ENSDEEEAN ^V KQ ^D -----	0.8104 (Fort)	Fort		
HSFA8	K143	SGLWKEV ^D L ^L KGD ^K QVLA ^Q E ^L	-	Faible		
	K166	VRQYQEV ^T D ^T KMLHLE ^D RVQ ^G	0.5624 (Fort)	Faible		
	K303	LLSPSRK ^K T ^K QEN ^I VK ^K G ^K D	0.4161 (Moyenne)	Faible		
HSFA9	K47	SKPDDV ^V LL ^I KEE ^E DDAV ^N LS	0.8754 (Fort)	Fort		
	K83	LRKTFEI ^V DD ^K V ^T DPV ^V SW ^S P	0.4861 (Moyenne)	Fort		
	K185	TTETE ^V ES ^L KEE ^Q SPMR ^L EM	0.6526 (Moyenne)	Faible		
Classe B	HSFB1	K164	ADLSGENE ^K L ^K RENN ^N LSSEL	0.4749 (Moyenne)	Faible	
		K244	EGLK ^L FGV ^W LKGER ^K K ^R DR ^D E	0.7044 (Moyenne)	Faible	
		K248	LFGV ^W LKGER ^K K ^R DR ^D EKN ^Y V	0.2572 (Faible)	Faible	
		K268	VVSGSR ^M TE ^I KN ^V DFHAP ^L W ^K	0.8012 (Fort)	-	
	*HSFB2b	K279	EDCAHLL ^P E ^G KPLDLL ^P ER ^Q E	0.2947 (Faible)	Fort	
		K356	AQEGE ^Q SSD ^V KAE ^P ME ^E NN ^S G	0.9315 (Fort)	Fort	
HSFB3	K242	EDEGL ^K L ^F GV ^K LE-----	0.8886 (Fort)	Fort		

Classe C	*HSFC1	K131	GEIVREIERLKEEQRELEAEI	0.7703 (Fort)	Faible	
		K202	KRRRVMTSTVKSEEEVEEDE	0.7963 (Fort)	Fort	

Ces données fournissent des informations sur les sites de SUMOylation de chaque HSF, identifiés par spectrométrie de masse et annotés d'un astérisque. Les lysines situées dans un site de SUMOylation consensus direct (Ψ -K-x-E) sont représentées en vert, celles dans un consensus inverse (E-x-K- Ψ) en orange, et celles présentes dans un consensus à la fois direct et inverse sont marquées en rose.

La figure 3.1 présente une représentation graphique détaillée de la corrélation entre les scores attribués par les serveurs web GPS-SUMO et JASSA pour les sites de SUMOylation sur les protéines étudiées. Cette courbe met en évidence des nuances importantes dans la relation entre les deux ensembles de scores, soulignant ainsi l'absence d'une relation linéaire claire entre eux. Un aspect notable de cette faible corrélation est que des niveaux de confiance jugés élevés par GPS-SUMO peuvent correspondre à des niveaux de confiance relativement faibles selon JASSA. Cette divergence entre les deux outils est particulièrement remarquable et nécessite une attention particulière lors de l'interprétation des résultats. L'analyse approfondie révèle que la corrélation entre les scores des deux serveurs varie considérablement en fonction des sites spécifiques de SUMOylation sur les protéines examinées. En examinant de près le graphique, il devient évident que les fluctuations dans cette corrélation sont plus prononcées pour certaines positions spécifiques de SUMOylation. En comparant les scores individuels attribués par chaque logiciel, il est alors possible d'identifier les sites de SUMOylation présentant les écarts les plus importants entre les deux logiciels, ce qui permet une meilleure compréhension de la variabilité dans la prédiction des sites de SUMOylation. Ces différences notables dans l'évaluation de la confiance entre GPS-SUMO et JASSA soulignent l'importance cruciale de prendre en compte la spécificité des positions individuelles lors de l'analyse des résultats. Il est essentiel de reconnaître que la faible valeur du coefficient de corrélation suggère que les scores attribués par GPS-SUMO et JASSA ne suivent pas une tendance uniforme sur l'ensemble des sites de SUMOylation.

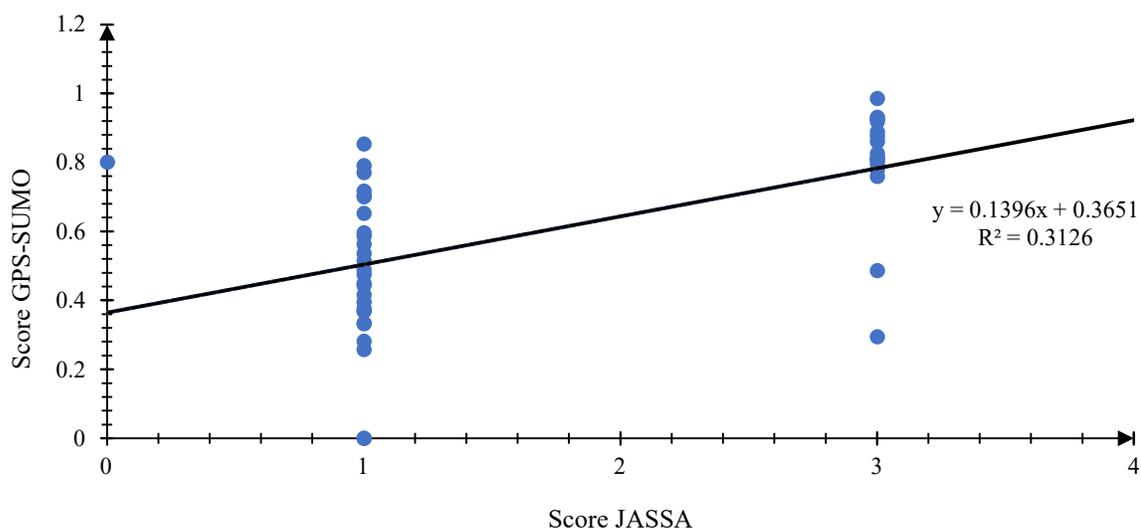


Figure 3.1 Analyse comparative des scores des serveurs web JASSA et GPS-SUMO : Une étude de corrélation pour la prédiction des sites de SUMOylation

Le tableau 3.2 présente les motifs d'interaction avec SUMO (SIM) chez différents facteurs de choc thermique. Les SIMs jouent un rôle dans le recrutement de l'enzyme E2 indirectement, favorisant ainsi la SUMOylation de la protéine. On peut remarquer que tous les HSFs contiennent des motifs SIM, à l'exception de la protéine HSFA7b. L'absence de motif SIM dans une protéine indique qu'elle ne comporte pas la séquence spécifique qui facilite son interaction avec des protéines SUMOylées. Ces motifs SIM agissent comme des points d'ancrage, permettant à la protéine de se lier spécifiquement aux protéines SUMOylées. L'absence de ce motif peut suggérer que la protéine en question ne participe pas directement à des interactions spécifiques avec des protéines SUMOylées, ou du moins, qu'elle n'utilise pas les motifs SIM conventionnels à cet effet. Par ailleurs, les scores fournis par des outils bio-informatiques tels que GPS-SUMO et JASSA pour les motifs SIM représentent des mesures de prédiction quant à la capacité d'une séquence donnée à interagir avec des protéines SUMOylées. Ces scores sont généralement calculés en évaluant la similarité de la séquence avec des motifs SIM connus, en tenant compte de la présence de résidus clés et d'autres caractéristiques spécifiques aux motifs SIM. Il est à noter que les scores peuvent varier entre les outils en raison des différences dans les algorithmes utilisés et les bases de données de motifs SIM prises en compte, ce qui est également le cas avec les deux serveurs web que nous avons utilisés pour extraire les scores des motifs.

Tableau 3.2 Motifs d'interactions SUMO (SIM) chez les protéines HSFs prédits au moyen des logiciels GPS-SUMO et JASSA

	HSF	Position du SIM	Séquence	Score GPS-SUMO	Score JASSA
Classe A	HSFA1a	135-139	FLRGQKHLLKKISRKRKSVQ	0.1447	-
		199-203	KNVLMQELVKLRQQQQTDT	0.1955	0.001
		482-486	NGRHMDKLEELGLLTSET	0.9960	-
	HSFA1b	165-169	KNVLMQELVRLRQQQATE	0.5226	0.074
		210-214	AVQSPGFLNQLVQQNNNDG	0.5276	-
		292-296	SNNPDSFLLGDVPSSTSVD	0.3108	-
		383-387	SGECFDPIMAVLDESEGDA	0.7520	-
		448-452	GSVENNDLVLEQEPNEWTR	0.2491	0.059
	HSFA1d	120-124	FLRGQKHLQSI TRRKPAP	0.2028	-
		179-183	KNVLMQELVRLRQQQSTD	0.2945	0.074
	HSFA2	127-131	FLAGQKHLLKNIKRRRNMG	0.2343	-
		172-176	RLKRDHGVVAEVVRLRQQ	0.6831	0.000
		177-181	HGVLVAE VVRLRQQQHSK	0.7587	0.573
	HSFA3	217-221	AEQRQKQLLSFLAKLFQNR	0.2103	-
	HSFA4c	142-146	LKNEKEGLLAELQEQEER	0.2017	-
		177-181	MEQHQSIVAYVSQVLGKP	0.2471	-
HSFA5	322-326	TGTCPRRLLLDNTRVETL	0.3480	0.000	
HSFA6a	56-60	WEPETFALICLPCRFKHN	0.4537	0.155	
HSFA6b	98-102	WDPQAFSVTLPLPRFFKHN	0.4303	0.413	
	144-148	-FLRGQKHLLKNI RRRKTSN	0.2187	-	
	246-250	AMQNPDFIQQLVEQKEKRR	0.2373	-	
	338-342	NSSAREEVLVNVEKGNDEEE	0.9818	0.081	
	389-393	FEGDQENVDVLIQQQLGYLG	0.9853	0.040	
HSFA7a	147-151	KQVLMMEIVSLRQQQQTTK	0.2203	0.223	
	224-228	GSSSMSELEVLALQMGGHG	0.4645	0.137	
	248-252	LEEDDHQLVVERELDDGFW	0.8066	0.000	
HSFA7b					
HSFA8	48-52	PSADNSFVILDTTVFSVQL	0.5936	0.637	
	57-61	DTTVFSVQLLPKYFKHSN	0.2683	0.413	
	139-143	KSGLWKEVDILKQDKQVLA	0.5268	0.918	
	153-157	KQVLAQELIKVRQYQEVTD	0.2937	0.001	
	219-223	KAGEGAKIVEEVTDEGESN	0.8335	-	
	234-238	GESNSYGLPLVTYQPPSDN	0.7591	0.021	
	275-279	FCLDENHVPLIIPDLYDDG	0.4979	0.193	
	360-364	TPERSRNLEILTEQMELLA	0.6286	0.119	
HSFA9	42-46	SSSKPDDVLLIKEEEDDA	0.8633	0.520	
	54-58	KEEEDDAVNLSLGFWKLHE	0.2786	-	
	66-70	GFWKLHEIGLITPFLRKT	0.2442	0.000	
	240-244	LAKDQRFVERLVKKRMMKI	0.3100	-	
	297-301	EHNPEPDIIVNNQSGNTRC	0.6156	0.078	
315-319	QLNSEDLLVDGGSMDVNG	0.8197	0.009		
Classe B	HSFB1	97-101	FRRGGEDLLTDIRRRKSVI	0.2127	-
		116-120	ASTAGKCVVVGSPSESNSG	0.9763	0.287
		184-188	AKKQRDELVTFLTGHKVR	0.5732	-
	HSFB2b	141-145	CFKRGEKILLRDIQRRKIS	0.7652	0.001
177-181		SAVTVAAPVVAHIVSPSN	0.8018	0.619	
208-212		PAAAAAIGGVVGGGSLQR	0.8375	-	
281-285		LLPEGKPLDLLPERQEMSE	0.2677	0.036	
318-322	LTPLRFGVISGVRARREE	0.4791	-		
HSFB3	210-214	TKKKCKQLMELVERYGED	0.8409	-	
Classe C	HSFC1	45-49	GPAHNSFIVVDP LDFSQRI	0.4565	0.895
		123-127	QDLEDGEIVREIERLKEEQ	0.6383	-

Légende : La séquence en rouge représente le motif SIM.

Parmi les 21 HSFs d'*Arabidopsis thaliana*, seules 4 HSFs soit HSFA1e, HSFA4a, HSFB2a et HSFB4 ne figurent pas dans les tableaux 3.1 et 3.2 en raison de l'application d'un seuil élevé pour les sites de potentiels de SUMOylation. La présence et l'abondance des sites de SUMOylation dans les HSFs varient en fonction des différentes classes de HSFs. Celles-ci (Classes A, B et C) peuvent avoir des motifs de séquence spécifiques qui favorisent ou limitent la SUMOylation. Les facteurs tels que la structure primaire, la conservation évolutive et d'autres caractéristiques propres à chaque classe peuvent influencer la propension à la SUMOylation. D'autre part, le nombre de sites de SUMOylation peut varier en fonction de la longueur de la protéine. En général, les protéines plus longues contiennent un plus grand nombre de résidus d'acides aminés, ce qui peut potentiellement offrir plus de sites susceptibles de subir une SUMOylation. Il semble que les HSFs de Classe A présentent un nombre plus élevé de site de SUMOylation par rapport aux deux autres classes.

La figure 3.2 offre une représentation détaillée de la structure des facteurs de choc thermique répertoriés dans le tableau 3.1. Ces catégories sont établies en fonction de leurs sites de SUMOylation prédits par les logiciels GPS-SUMO et JASSA. Chaque facteur de choc thermique est caractérisé par trois domaines fonctionnels majeurs : le domaine de liaison à l'ADN, le domaine d'oligomérisation et le domaine de localisation nucléaire. Une observation significative émerge concernant la localisation des lysines potentiellement SUMOylables. L'analyse de la figure révèle que la plupart de ces lysines se trouvent principalement dans le domaine d'oligomérisation, ainsi que dans la région désordonnée de la protéine, et quelques-unes dans le domaine de liaison à l'ADN. Le domaine d'oligomérisation, facilitant l'assemblage de la protéine en structures oligomériques telles que des dimères, des trimères ou des complexes plus importants, joue fréquemment un rôle crucial dans la fonction biologique des protéines en influençant leur activité enzymatique, leur stabilité structurale et leurs interactions avec d'autres protéines. La SUMOylation des lysines dans ce domaine pourrait ainsi altérer la capacité de la protéine à former des complexes oligomériques. Les lysines situées dans les régions désordonnées, souvent liées à la régulation de la signalisation cellulaire, à l'interaction avec d'autres protéines ou à la modulation de la conformation protéique, pourraient également subir la SUMOylation, jouant ainsi un rôle crucial dans la régulation de la fonction protéique, particulièrement dans le contexte des mécanismes de réponse au stress thermique associés aux

facteurs de choc thermique. En régulant la SUMOylation des lysines dans le domaine de liaison à l'ADN, la cellule peut finement ajuster les interactions protéine-ADN, influençant ainsi les processus cellulaires liés à la régulation génique. Cette régulation revêt une importance particulière dans des contextes tels que la réponse au stress thermique, où la modulation précise des mécanismes de transcription et de réparation de l'ADN joue un rôle essentiel dans la survie cellulaire. La distribution spécifique des lysines SUMOylables souligne ainsi l'importance potentielle de ces domaines fonctionnels dans les mécanismes de régulation thermique.

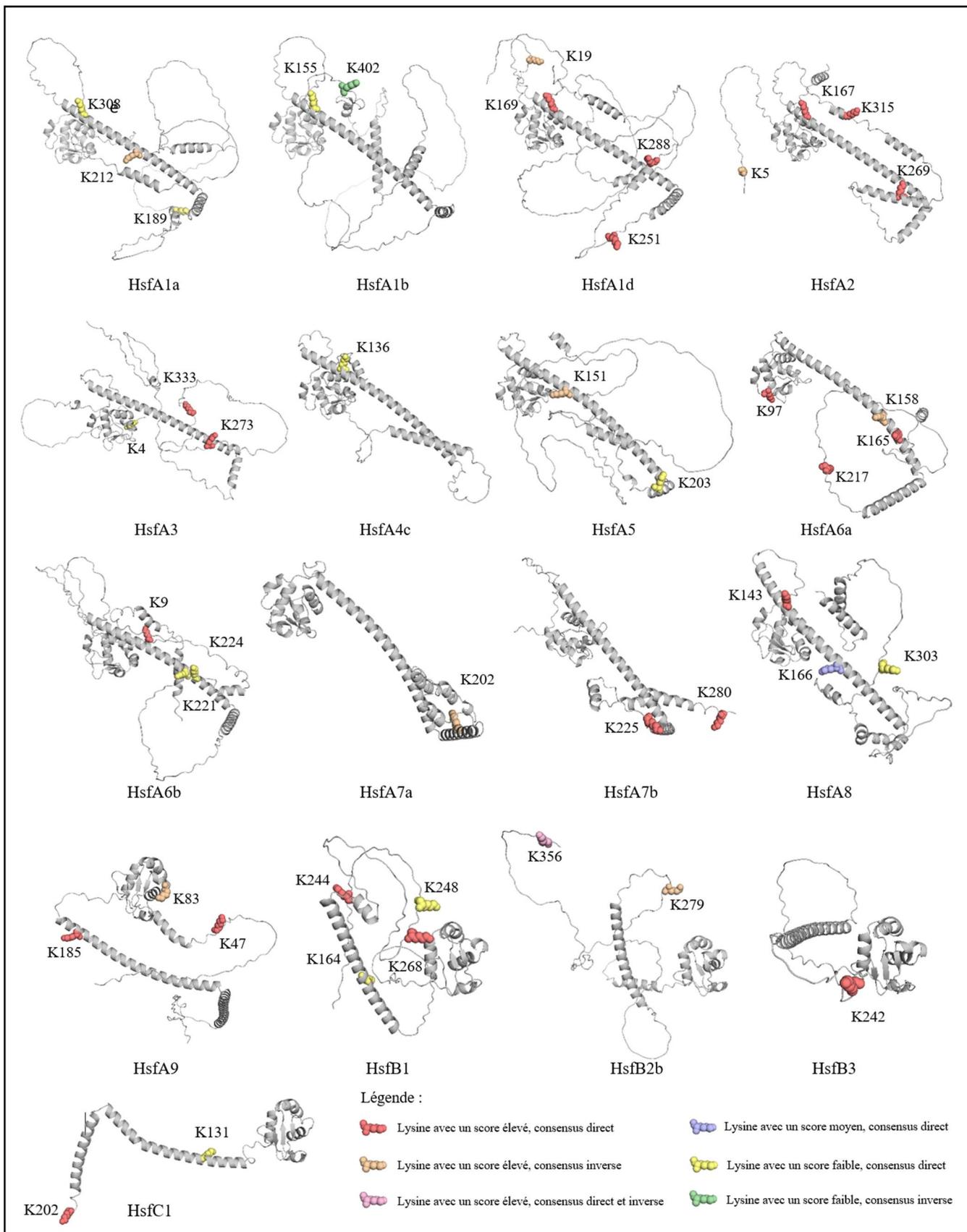
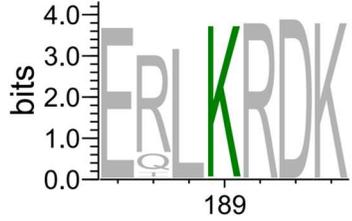
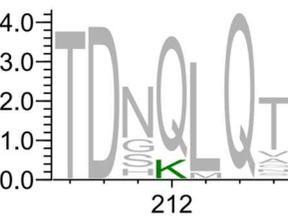
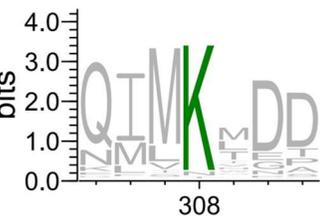
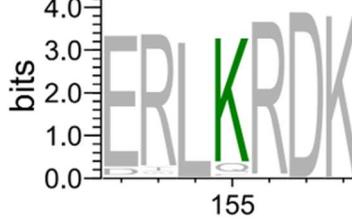
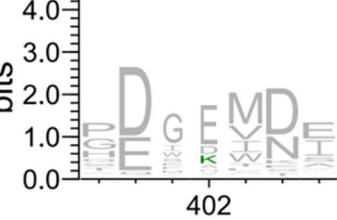
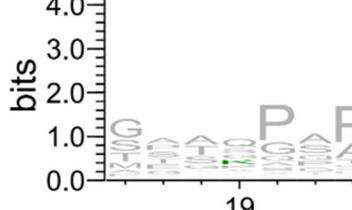
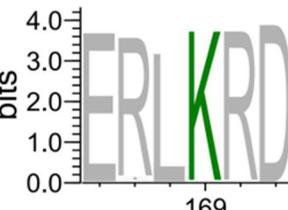
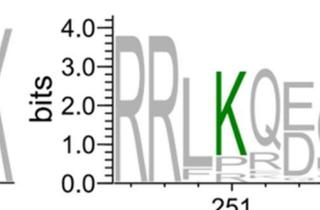
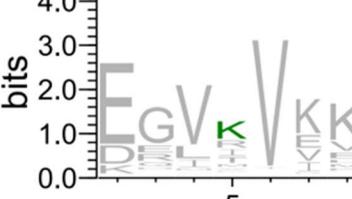
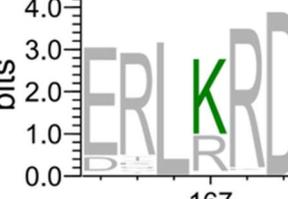
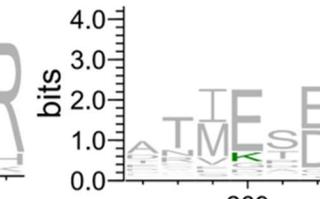
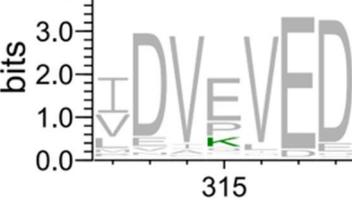


Figure 3.2 Positionnement des sites de SUMOylation prédits par les logiciels JASSA et GPS-SUMO. Les modèles pour chaque HSF proviennent d'AlphaFold et ont été représentés à l'aide de PyMol.

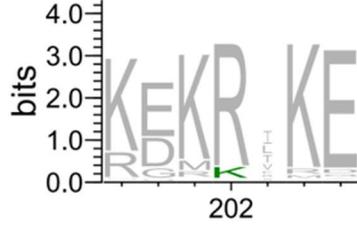
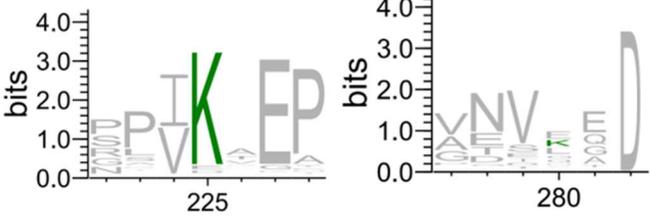
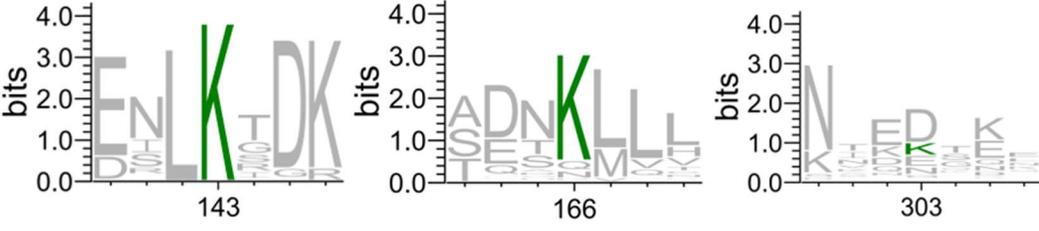
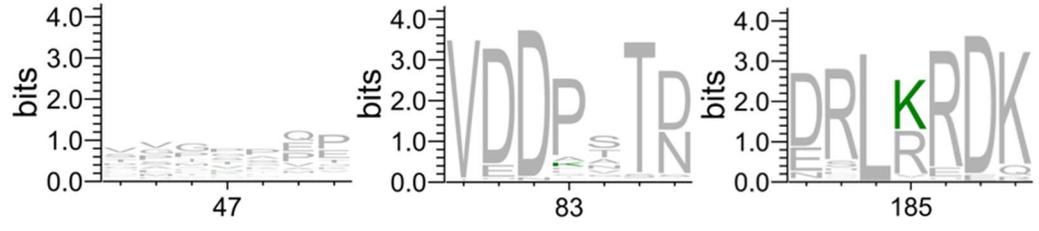
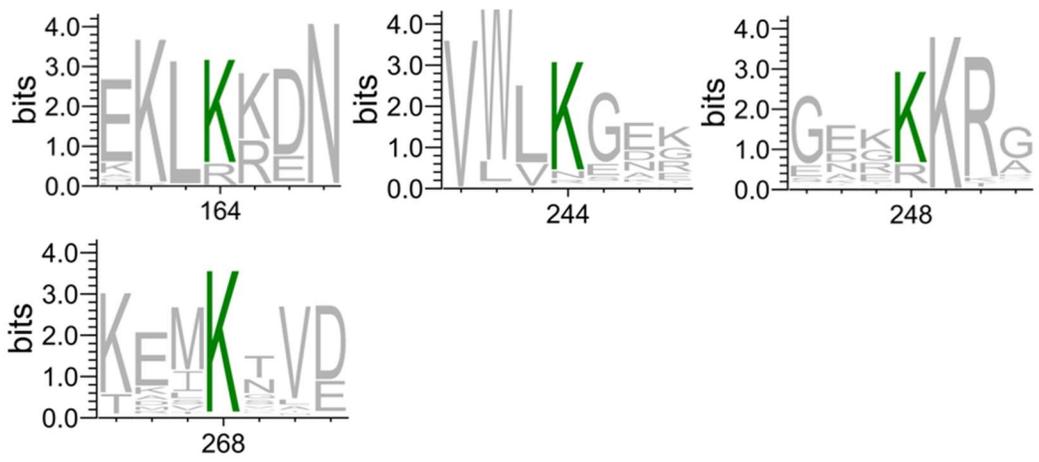
La figure 3.3 présente un WebLogo illustrant la conservation des lysines chez *Arabidopsis* au sein des sites de SUMOylation prédits par les logiciels GPS-SUMO et JASSA, comme indiqué dans le tableau 3.1 et la figure 3.2. Le WebLogo est un instrument visuel qui représente la fréquence relative des acides aminés à chaque position d'une séquence, permettant ainsi d'identifier des motifs ou des zones significativement conservées. Des logos ont été générés pour chaque site de SUMOylation des facteurs de transcription HSF, avec la lysine marquée en vert et les autres résidus en gris. La conservation des lysines dans les sites de SUMOylation fait référence à la préservation de lysines spécifiques au sein de ces sites. Sur le graphique, l'axe horizontal représente les positions dans la séquence, tandis que l'axe vertical indique la hauteur totale des lettres, reflétant ainsi le contenu d'information de chaque position en bits. La hauteur d'un acide aminé à une position donnée reflète son niveau de conservation : plus la hauteur est élevée, plus l'acide aminé est conservé à cette position.

En analysant la protéine HSFA1a, on remarque que les lysines des sites de SUMOylation K189 et K308 présentent une conservation plus élevée que K212. La hauteur de la lysine à la position 212 est notablement plus basse, affichant une valeur inférieure à 1 bit. De manière similaire, pour HSFA1b, la lysine K155 est beaucoup plus conservée que K402, comme indiqué par sa hauteur sur le WebLogo, qui est notablement plus basse. Pour la protéine HSFA1d, les lysines K169 et K288 sont fortement conservées par rapport à K251 et K19, bien que la lysine K19 présente une conservation moindre à cet emplacement. Pour HSFA2, toutes les lysines des sites montrent une conservation relativement faible, bien que K167 se démarque par une conservation plus élevée que les autres. Dans le cas de HSFA3, la lysine à la position 333 est beaucoup plus conservée que K273 et K4. La protéine HSFA4c présente une conservation notable au site K136. Pour HSFA5, K203 est plus conservée que K151. Concernant HSFA6a, les positions K165 et K217 montrent une conservation significativement plus élevée que K97 et K158, tandis que pour HSFA6b, seule K9 présente une forte conservation par rapport à K221 et K224. La lysine du site de SUMOylation de HSFA7a n'est pas conservée, tandis que pour HSFA7b, K225 est mieux conservée que K280. Les lysines K143 et K166 de HSFA8 sont considérablement conservées par rapport à K303. La protéine HSFA9 ne semble pas contenir de lysines conservées en aucun endroit de la séquence. Chez HSFB1, toutes les lysines des sites, à savoir K164, K244, K248 et K268, présentent une conservation, de même que pour HSFB2b, où les sites K279 et K356

montrent une conservation notable. La lysine K242 n'est pas aussi conservée chez HSFB3. Enfin, uniquement la lysine à la position 131 est conservée chez HSFC1.

HSFs	WebLogo		
HSFA1a	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>189</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>212</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>308</p>
HSFA1b	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>155</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>402</p>	
HSFA1d	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>19</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>169</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>251</p>
HSFA2	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>5</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>167</p>	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>269</p>
	 <p>bits</p> <p>4.0 3.0 2.0 1.0 0.0</p> <p>315</p>		

HSFA3	
HSFA4c	
HSFA5	
HSFA6a	
HSFA6b	

HSFA7a	
HSFA7b	
HSFA8	
HSFA9	
HSFB1	

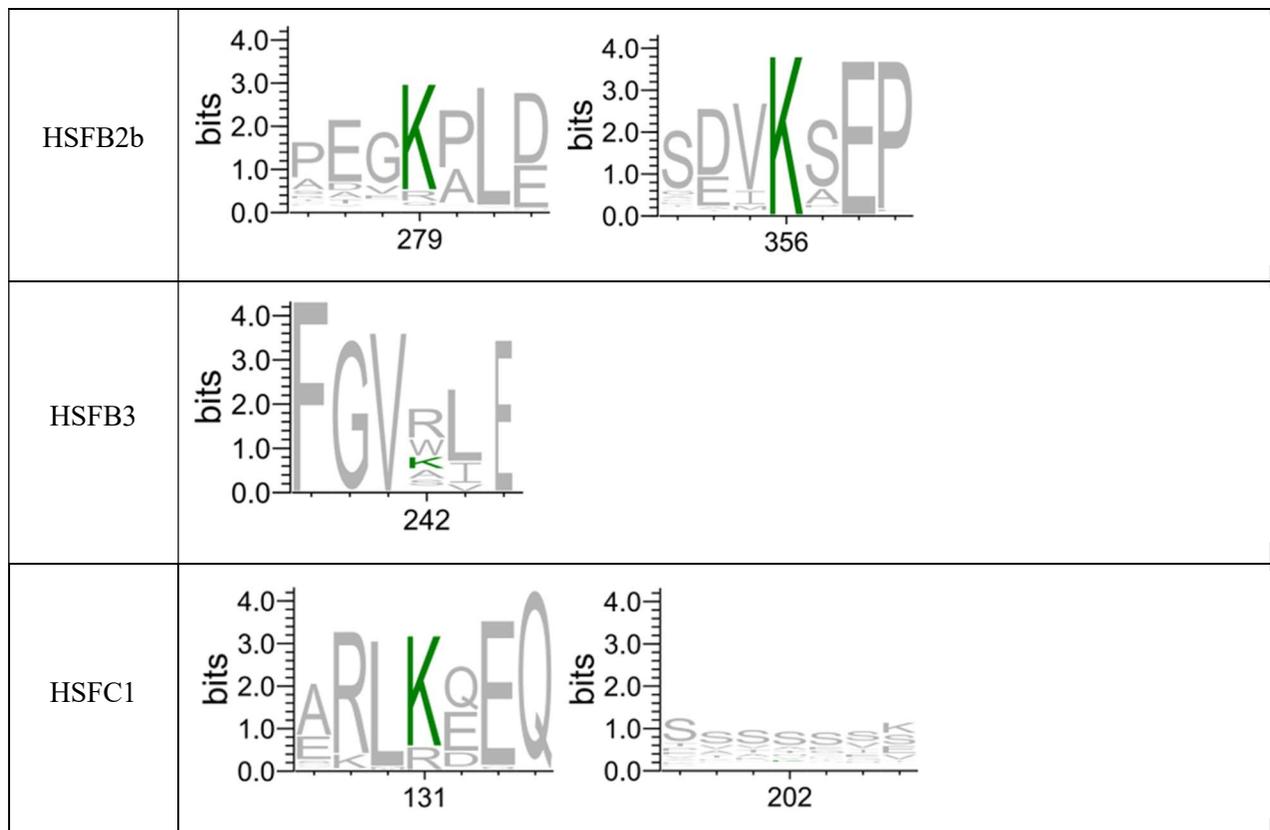


Figure 3.3 WebLogo de la conservation des sites de SUMOylation avec la lysine située au site prédit en vert et les autres résidus en gris des HSFs.

3.2 Conservation des lysines chez les 21 protéines HSFs

L'étude de la conservation des protéines HSFs dans la réponse des plantes au stress thermique offre un éclairage essentiel sur l'aspect évolutif de ces protéines. La conservation au fil du temps, observée dans différentes espèces, constitue un indicateur clé, suggérant un rôle crucial pour la survie ou les fonctions fondamentales des plantes. En effet, si une protéine est conservée de manière significative à travers l'évolution, cela implique généralement son importance biologique et sa contribution essentielle aux mécanismes adaptatifs des plantes. L'examen de la conservation des protéines revêt une double importance. Tout d'abord, il permet d'identifier des motifs ou des mécanismes de régulation qui transcendent les frontières des espèces. La préservation de ces éléments suggère leur rôle fonctionnel crucial dans la réponse au stress thermique. De plus, cette analyse de conservation offre une fenêtre précieuse pour explorer les caractéristiques

moléculaires partagées entre différentes plantes, éclairant ainsi les mécanismes communs qui sous-tendent la réponse au stress thermique dans le règne végétal. Pour approfondir cette analyse, les 21 facteurs de transcription HSF ont été soumis à un alignement minutieux à l'aide du logiciel MAFFT. Cette procédure a permis d'illustrer la conservation des séquences des protéines HSF à l'aide du logiciel ESPript, mettant en lumière des aspects spécifiques de la conservation. La Figure 3.4, résultant de cette démarche, focalise l'attention sur une conservation significative observée parmi les 21 HSFs. Notamment, une attention particulière est portée à la plage d'acides aminés 38 à 232 de la protéine HSFA1a. Cette plage spécifique englobe le domaine de liaison à l'ADN (DBD), s'étendant de l'acide aminé 50 à 144, ainsi que le domaine d'oligomérisation, qui s'étend de l'acide aminé 172 à 238. L'alignement des 21 protéines HSFs a révélé une conservation remarquable des lysines, notamment entre les positions 38 et 77 (figure 3.4.a), 78 et 137 (figure 3.4.b), 138 et 192 (figure 3.4.c), et 193 et 232 (figure 3.4.d). Cette mise en évidence des lysines fortement conservées suggère un rôle crucial de ces résidus dans les fonctions moléculaires partagées par les protéines HSFs.

La figure 3.5 présente une représentation du regroupement des 21 HSFs en fonction de leurs classes, à savoir la classe A, la classe B et la classe C, accompagnée de la moyenne du pourcentage de lysines SUMOylables. L'analyse révèle des variations distinctes dans la prévalence potentielle de la SUMOylation parmi ces classes. On constate que la moyenne des lysines SUMOylables diffère entre les classes, avec une moyenne de 6.12% pour la classe A, 5.92% pour la classe B, et 4.80% pour la classe C. Ces chiffres indiquent une tendance intéressante quant à la distribution des lysines potentiellement SUMOylables parmi les différentes classes de HSFs. Cette observation soulève des questions intéressantes sur les mécanismes de régulation spécifiques à chaque classe de HSFs et suggère des nuances fonctionnelles dans la SUMOylation parmi ces différents groupes de facteurs de transcription, offrant ainsi une perspective approfondie sur la régulation moléculaire des protéines HSFs en réponse au stress thermique.

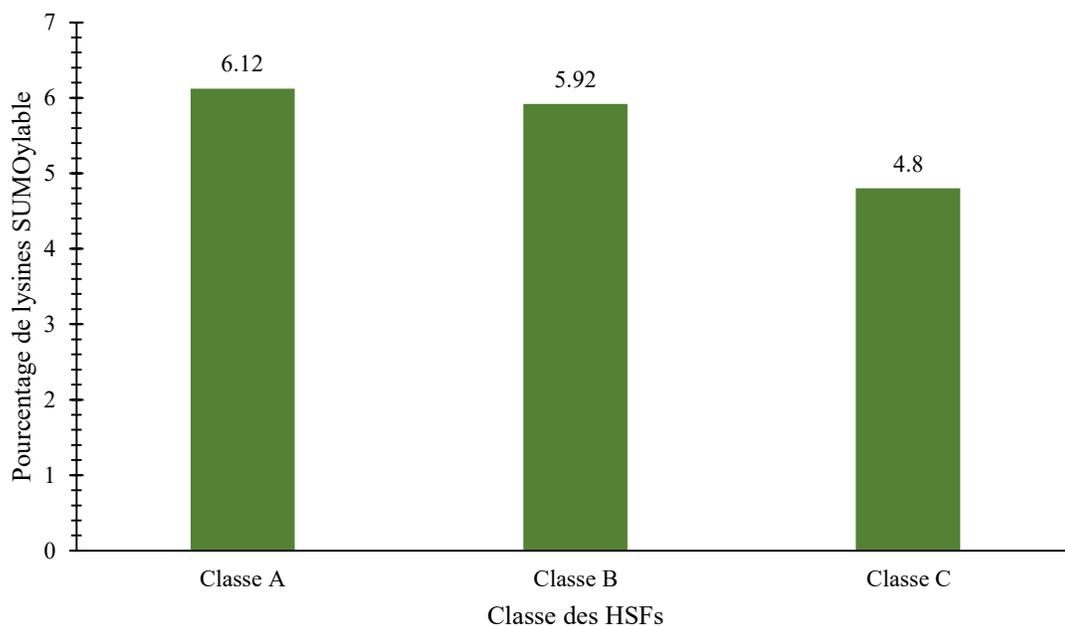


Figure 3.5 Répartition des lysines SUMOylables chez les HSFs en fonction de leurs classes.

3.3 Purification des protéines candidates HSFs

Suite aux analyses bio-informatiques qui ont identifié des protéines HSFs susceptibles d'être SUMOylées, il est important d'explorer la SUMOylation de protéines recombinantes. Afin de pouvoir comparer les sites de la SUMOylation des différentes protéines de choc thermique, il est important de purifier les protéines HSFs d'*Arabidopsis thaliana*. Lorsqu'on étudie les mécanismes moléculaires au sein des cellules végétales, il peut être extrêmement complexe de prendre en compte toutes les variables qui peuvent influencer un processus particulier. Les cellules végétales sont des environnements dynamiques et complexes où de nombreuses interactions moléculaires se produisent simultanément. L'utilisation des protéines recombinantes purifiées dans le cadre de cette étude offre plusieurs avantages en termes de simplification et de contrôle expérimental ainsi les protéines HSFs pourraient former des hétéro-oligomères ce qui compliqueraient les analyses. Les cellules végétales contiennent un grand nombre de protéines endogènes et de voies de signalisation, ce qui peut rendre difficile la compréhension des mécanismes moléculaires spécifiques. En utilisant des protéines recombinantes purifiées, on élimine la contribution de la plupart des autres composants cellulaires, permettant ainsi de se concentrer sur les interactions et les activités de la protéine d'intérêt de manière plus précise. D'ailleurs, en isolant des protéines spécifiques et en les purifiant, il est possible de tester les effets de chaque facteur de manière indépendante. Cela est particulièrement important pour comprendre des processus complexes tels que la SUMOylation, où plusieurs facteurs peuvent interagir simultanément. Par exemple, en testant l'effet de la machinerie de SUMOylation sur une protéine recombinante purifiée, il est possible de déterminer si la séquence de cette protéine est suffisante pour induire une SUMOylation efficace. De plus, dans de nombreux processus cellulaires, la liaison à l'ADN est cruciale, bien que les HSFs sont des facteurs de transcriptions qui lient l'ADN. En isolant et purifiant une protéine d'intérêt, il est possible d'étudier son interaction avec l'ADN de manière spécifique et déterminer l'importance de cette interaction dans la régulation de la SUMOylation. En outre, les expériences avec protéines recombinantes purifiées offrent un niveau élevé de contrôle expérimental. Les conditions de réaction, la concentration de la protéine, les réactifs utilisés, etc., peuvent être ajustés et contrôlés, permettant une interprétation plus précise des résultats expérimentaux.

La stratégie de clonage est une étape cruciale dans la manipulation génétique des organismes, notamment pour la production de protéines recombinante. Le processus de clonage implique l'amplification du gène d'intérêt à partir d'une source donnée, dans notre cas souvent de vecteurs obtenus de l'Arabidopsis Biological Resource Center (ABRC). Une fois le gène amplifié, il est inséré dans un vecteur d'expression, tel qu'un plasmide, à l'aide d'enzymes de restriction et de ligases à ADN. La transformation de bactéries avec ce plasmide modifié permet ensuite la production de la protéine fusion dans les bactéries. Une approche courante pour faciliter la purification des protéines recombinantes consiste à créer des protéines de fusion, où la protéine d'intérêt est fusionnée à des étiquettes qui permettent une expression et une purification efficace. Deux étiquettes couramment utilisées pour faciliter la purification sont l'étiquette 6x-His et l'étiquette thiorédoxine. L'étiquette 6x-His est une séquence de six résidus histidine qui a une forte affinité pour les ions métalliques, en particulier le nickel. Lorsque la protéine d'intérêt est exprimée sous forme de fusion avec l'étiquette 6x-His, elle peut être purifiée de manière efficace en utilisant une colonne de chromatographie d'affinité au nickel (Arnau et al., 2006). De son côté, la thiorédoxine est une petite protéine qui agit comme un modulateur de la solubilité des protéines (Née, 2011). En fusionnant la protéine d'intérêt avec l'étiquette thiorédoxine, on peut augmenter la solubilité de la protéine recombinante, ce qui est particulièrement utile lorsque la protéine d'intérêt a tendance à former des agrégats insolubles.

La figure 3.6 et le tableau 3.3 illustrent le processus de purification des protéines HSFs candidates d'Arabidopsis. La purification de la protéine HSFA1a a été réussie et elle apparaît sous une forme soluble, comme illustrée dans la figure 3.6.a. Cependant, il est important de noter que le clivage de cette protéine n'est pas complet, voire minimal. Sa taille approximative est de 56 kDa, tandis qu'elle atteint environ 67 kDa avec les étiquettes 6xHis-Trx. En ce qui concerne les protéines HSFA1e, HSFA7b, HSF3 et HSFC1, une fois de plus, ces protéines se révèlent solubles et subissent un clivage partiel à l'aide de la TEV, comme indiqué dans la figure 3.6.a-c, e-f. Leurs tailles respectives avec les étiquettes sont généralement de 52 kDa, 33 kDa, 28 kDa et 38 kDa, tandis qu'elles atteignent 64 kDa, 42 kDa, 40 kDa et 50 kDa, respectivement avec les étiquettes. En ce qui concerne la protéine HSFA9 (figure 3.6.d), elle présente une bonne solubilité, mais il n'y a pas eu de clivage à la TEV. Le clonage moléculaire de la protéine HSFA2 a été entravé par des difficultés rencontrées lors de la transformation bactérienne avec la

souche DH5 α . Par ailleurs, les tentatives d'amplification initiale des protéines HSFA7a et HSFB1 ont été infructueuses. Ces échecs peuvent être attribués à des problèmes potentiels liés aux amorces utilisées pour l'amplification.

Tableau 3.3 Résumé des clonages/expressions/purifications des protéines HSFs d'Arabidopsis

Protéine	Nom du plasmide ABRC	Amplification	Clonage moléculaire	Induit	Solubilité	Clivage à la TEV	Essai retard sur gel	Essai enzymatique
HSFA1a	DKLAT4G 17750	Oui	Oui	Oui	Oui	Partielle		
HSFA1e	U85727	Oui	Oui	Oui	Oui	Partielle		
HSFA2	DKLAT2G 26150	Oui	Non					
HSFA7a	U23517	Non						
HSFA7b	PENTR221	Oui	Oui	Oui	Oui	Partielle		
HSFA9	U18576	Oui	Oui	Oui	Oui	Non		
HSFB1	U16604	Non						
HSFB2b	U24153	Oui	Oui	Oui	Oui	Oui	Oui	Oui
HSFB3	U86661	Oui	Oui	Oui	Oui	Partielle		
HSFC1	U14335	Oui	Oui	Oui	Oui	Partielle		

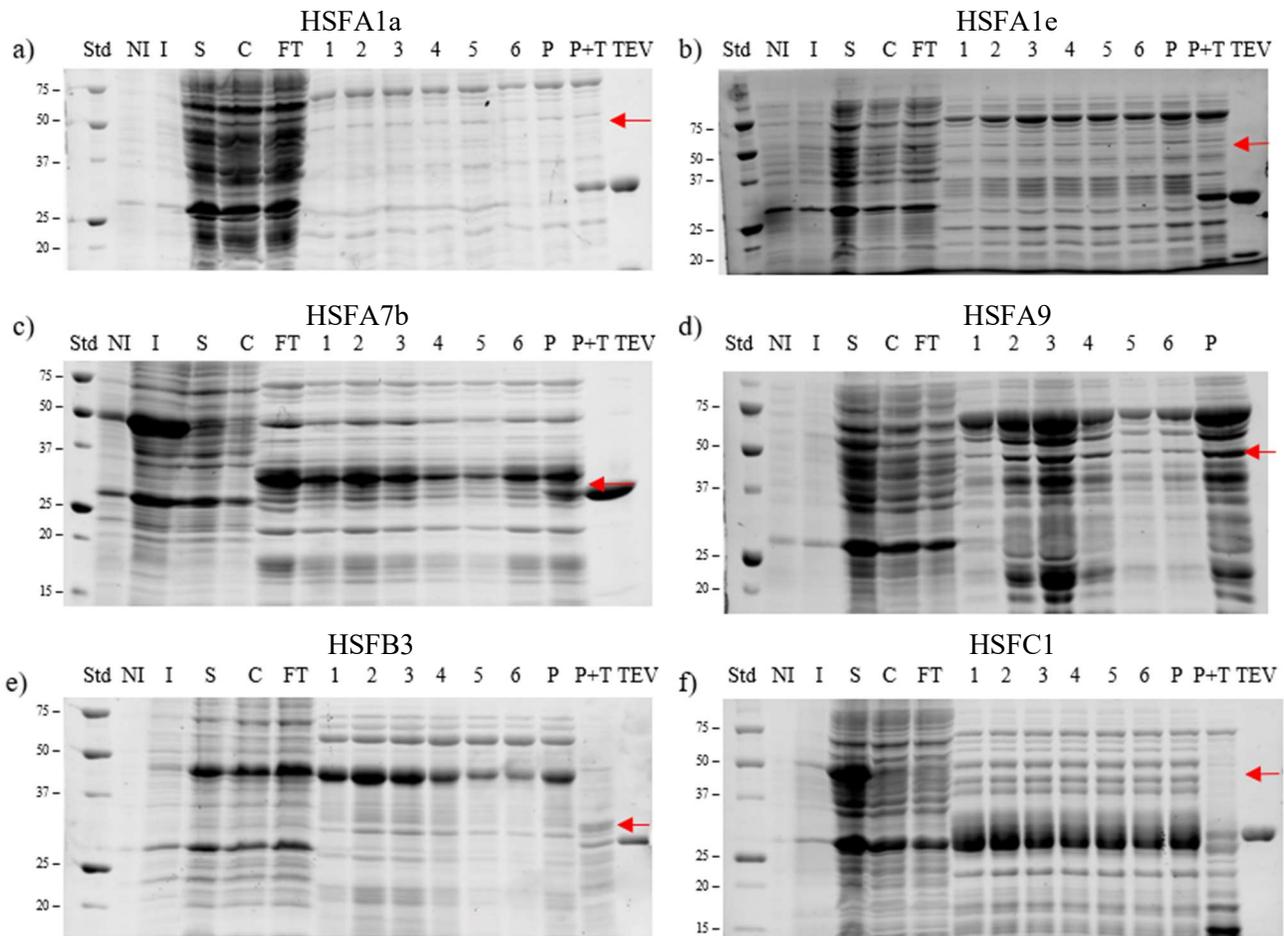


Figure 3.6 Gel SDS-PAGE de la purification des protéines HSFs candidates. a) Purification de la protéine HSFA1a. b) Purification de la protéine HSFA1e. c) Purification de la protéine HSFA7b. d) Purification de la protéine HSFA9. e) Purification de la protéine HSFB3. f) Purification de la protéine HSFC1. L'acquisition des images a été faite à l'aide du ChemiDoc MP. Échantillon standard (Std), non-induit (NI), induit (I), sonication (S), centrifugation (C), flow-through (FT), les échantillons purifiés en fraction 1 à 6, regroupement des fractions de protéines (P) et le regroupement avec le clivage à la TEV (P+T). Les flèches en rouge montrent la protéine partiellement clivée.

3.4 Purification de la protéine HSFB2b

Parmi les protéines HSF candidates, il a été possible d'exprimer et de purifier la protéine HSFB2b avec succès, marquant ainsi le début d'une étude approfondie de sa SUMOylation. Cette investigation a impliqué un clivage complet des étiquettes 6xHis-Trx, un test de retard sur gel et enfin, un essai enzymatique, comme détaillé dans le Tableau 3.3. La protéine HSFB2b, constituée de 377 acides aminés, présente une structure caractérisée par plusieurs domaines fonctionnels. Le

domaine de liaison à l'ADN de la protéine HSF2b s'étend de la position 57 à la position 151 des acides aminés. Ce domaine revêt une importance cruciale dans les interactions spécifiques avec l'ADN, soulignant ainsi le rôle potentiel de cette protéine dans la régulation génique. Par ailleurs, le domaine d'oligomérisation composé de résidus hydrophobes a été identifié entre les acides aminés 220 et 265, suggérant une possible implication dans des interactions hydrophobes essentielles à la fonction de la protéine. La figure 3.7 illustre la localisation précise des domaines fonctionnels de la protéine HSF2b. Notamment, un domaine de signal de localisation nucléaire a été identifié entre les positions 323 et 327 des acides aminés, suggérant que cette protéine peut être localisée spécifiquement dans le noyau cellulaire. Ces informations structurales détaillées fournissent une base solide pour comprendre la fonction potentielle de la protéine HSF2b dans les mécanismes de régulation cellulaire et renforcent la pertinence de son étude approfondie.

```

MPGEQTGETPTVAGVGGGGAGCSAGNSGGSSGCGAGGGGGGSGGGGGGGDSQRSIPTPF
LTKTYQLVEDPVYDELISWNEDGTTTFIVWRPAEFARDLLPKYFKHNNFSSFVRQLNTYGF
RKVVPDRWEFSNDCFKRGEKILLRDIQRRKISQPAMAAAAAAAAAAVAASAVTVAAVPVV
AHIVSPSNSGEEQVISSNSSPAAAAAIGGVVGGGSLQRTTSCTTAPELVEENERLRKDN
ERLRKEMTKLKGLYANIYTLMANFTPGQEDCAHLLPEGKPLDLLPERQEMSEAIMASEIE
TGIGLKLGEDLTPRLFGVSIGVKRRREEELGAAEEEDDDRREAAAQEGEQSSDVKAEPM
EENNSGNHNGSWLELGK

```

Figure 3.7 Séquence primaire de la protéine HSF2b. Le domaine de liaison à l'ADN (bleu), le domaine de l'oligomérisation (orange) et le domaine de signal de localisation nucléaire (vert). Les deux sites de SUMOylation K279 et K356 (rouge).

La figure 3.8 représente la purification de la protéine HSF2b. Les principaux sites prédits de SUMOylation de la protéine HSF2b sont K279 (en jaune) avec un score faible et K356 (en rouge) avec un score élevé (Figure 3.8.a). L'expression de la protéine a été faite chez les bactéries BL21 DE3 CodonPlus pRIL. Ces bactéries sont couramment employées pour l'expression de protéines recombinantes. Elles expriment l'ARN polymérase T7 qui permet une expression hautement efficace des gènes insérés dans des vecteurs compatibles. De plus, le plasmide pRIL

fournit des facteurs qui améliorent la stabilité et la synthèse des ARNt, ce qui favorise la production correcte des protéines cibles. En utilisant cette souche pour l'expression, on obtient généralement des niveaux de production de protéines recombinantes plus élevés. Les bactéries ont été cultivées dans du milieu LB et SB en tenant compte le temps, la température et la concentration d'IPTG. Le milieu LB est utilisé pour les pré-cultures. En revanche, le milieu SB est privilégié pour atteindre des niveaux plus élevés en bactéries en raison de son enrichissement nutritionnel. Un temps de culture adéquat est essentiel pour permettre une croissance cellulaire suffisante et l'expression optimale de la protéine. La température peut influencer la croissance cellulaire et l'efficacité de l'expression protéique. Dans notre cas, la pré-culture se fait à une température de 32 °C à 200 rpm pendant environ 18 heures, par la suite la culture se fait à 37 °C. La concentration d'IPTG doit être optimisée pour déclencher l'expression sans entraîner de toxicité cellulaire excessive. On induit l'expression protéique de la culture de 3 L à 37 °C avec 300 µM d'IPTG lorsque la OD600nm se situe entre 0.8 et 1.0. L'induction se fait à 18 °C à 200 rpm durant environ 18 heures. Par la suite, les bactéries ont été centrifugées et lysées afin de pouvoir être purifiées par chromatographie d'affinité sur une colonne de nickel à l'aide d'un FPLC (Thermo Fisher Scientific, Montréal, Canada). Ensuite, afin de pouvoir éliminer l'étiquette 6x-His et obtenir la protéine pure, un clivage à la TEV est fait. La protéase TEV est connu pour sa spécificité de couper après une séquence d'acides aminés particulière, ENLYFQ/G, afin de minimiser les risques de perturbation de la structure ou la fonction de la protéine (Harder et al., 2018).

L'étiquette 6x-His et l'étiquette thiorédoxine se retrouvent au N-terminal de la protéine HSF. Après avoir exprimé la protéine en grande quantité et effectué la purification sur colonne, il a été observé que la bande de "flow-through" était intense, indiquant ainsi qu'une importante quantité de protéine dans le surnageant n'a pas été retenue par la colonne. Par conséquent, un second passage de la protéine a été réalisé. Par exemple, lors de la purification d'un volume initial de 50 mL, 13 mg de la protéine HSFB2b a été éludée à la première édition avec la capacité de la colonne de 200 mg. Si l'on considère que la capacité théorique attendue de la résine est d'environ 40 mg de protéine par mL, la quantité de protéine récupérée semble être inférieure à la capacité maximale théorique de la résine, qui serait de 1500 mg. Ceci pourrait être expliqué par la saturation de la capacité de la colonne qui fait référence au point où la colonne de

chromatographie atteint sa capacité maximale de rétention des molécules cibles. Il est aussi possible d'observer sur le gel que la protéine a bien été clivée par la TEV en comparant la migration des protéines avant et après clivage (Figure 3.8.b). Afin de pouvoir se débarrasser des protéines indésirables, y compris la TEV, et la 6x-His thiorédoxine, une deuxième purification a été faite au moyen d'une colonne de tamis moléculaire. Cette méthode de purification permet d'obtenir la protéine d'étude sous sa forme la plus pure. Les échantillons provenant des purifications ont été mis dans un tampon Laemmli 2x et, par la suite, chargés sur un gel SDS-PAGE (Figure 3.8.c-d).

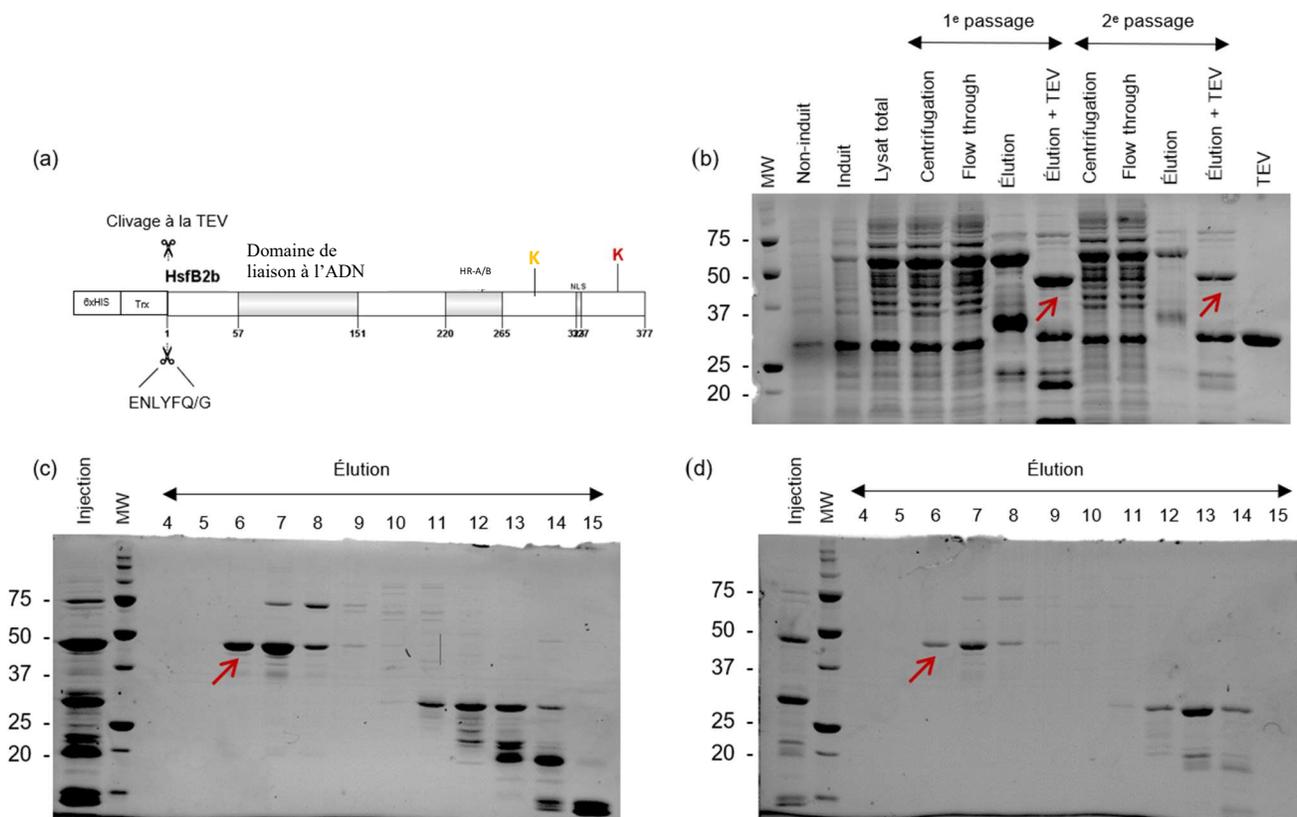


Figure 3.8 Exemple de la purification de la protéine HSF2b. A) Structure de la protéine HSF2b. B) Gel SDS-PAGE de la purification de HSF2b par chromatographie par affinité avec une colonne de Ni-NTA. La protéine 6xHis-Thiorédoxine-HSF2b a été clivée par la protéase TEV. C) Gel SDS-PAGE de la purification de la protéine sur tamis moléculaire. D) Gel SDS-PAGE de la purification de la protéine 2^e passage sur tamis moléculaire. L'acquisition des images a été faite à l'aide du ChemiDoc MP. Les flèches en rouge montrent la protéine clivée à différentes étapes de purification.

Le tableau 3.4 présente les résultats de récupération de la protéine HSFB2b en fournissant des détails sur la quantité initiale de bactéries, les volumes et les concentrations obtenus à chaque étape. Une culture bactérienne de 3 litres a été initiée, produisant un culot de bactéries de 26.7 grammes. Suite à la lyse cellulaire et à la centrifugation, un volume de 100 mL a été obtenu en vue de la purification. Lors du chargement de la protéine sur la colonne, il a été observé qu'une quantité significative de protéine était encore présentée dans l'échantillon de flow-through. Pour maximiser le rendement de la protéine HSFB2b, cet échantillon a été soumis à une seconde étape de chargement sur la colonne Ni-NTA. Les quantités protéiques obtenues sont 11.65 mg et 2.775 mg respectivement. Une purification supplémentaire par tamis moléculaire a été réalisée lors des deux passages pour éliminer les impuretés ainsi que les étiquettes 6x-His et la thiorédoxine. À la suite de la concentration des protéines, les concentrations protéiques ont été élevées à 60.94 μM et 37.51 μM pour le premier et le deuxième passage, respectivement, à partir des concentrations initiales de 208.90 μM et 37.70 μM obtenues lors de la purification par Ni-NTA. La purification de la protéine HSFB2b s'est révélée prometteuse, aboutissant à l'obtention d'une protéine HSFB2b pure, caractérisée par un clivage total de l'étiquette. De plus, une quantité substantielle de la protéine a été obtenue, ce qui ouvre la voie à la réalisation d'autres expériences. La qualité élevée de la protéine purifiée et la quantité adéquate obtenue offrent des bases solides pour entreprendre des expérimentations ultérieures.

Tableau 3.4 Exemple de la purification de la protéine HSFB2b

Étape de purification	Quantité de bactéries de départ (g)	Quantité de protéine (mg)	Concentration (μM)
Culture bactérienne	26,70	-	-
Chromatographie d'affinité			
Quantité récupérée Ni-NTA 1 ^{er} passage	-	13.98	-
Quantité récupérée Ni-NTA 2 ^e passage	-	3.33	-
Quantité récupérée			
Tamis moléculaire 1 ^{er} passage	-	0.445	-
Quantité récupérée			
Tamis moléculaire 2 ^e passage	-	0.080	-

Concentration finale HSFB2b 1 ^{er} passage			60.94
Concentration finale HSFB2b 2 ^e passage	-	-	37.51

3.5 HSFB2b se lie à la séquence spécifique HSE

Comme illustré précédemment, la protéine recombinante HSFB2b a été exprimée, purifiée et démontrée capable de se lier à sa séquence cible d'ADN à travers des expériences. La figure 3.9 démontre la capacité de la protéine HSFB2b à interagir tant avec des séquences spécifiques qu'avec des séquences non spécifiques d'ADN, révélant ainsi son fonctionnement actif et sa conformation correctement repliée. Les séquences HSE, également appelées éléments de réponse au choc thermique, agissent comme sites de liaison spécifiques pour les facteurs de transcription HSF activés, qui montrent une forte affinité pour ces séquences situées dans les régions promotrices des gènes de choc thermique. Contrairement à son interaction sélective avec l'ADN spécifique (figure 3.9a,b), la protéine HSFB2b manifeste une affinité relativement faible envers l'ADN non-spécifique. Les facteurs de transcription, dont les HSFs, montrent une préférence pour des séquences nucléotidiques spécifiques, régulant ainsi de manière sélective leur liaison à travers leur domaine de liaison à l'ADN (figure 3.9a,b). En conséquence, l'ADN spécifique, caractérisé par des motifs nucléotidiques particuliers, sert de site de liaison pour les HSFs, tandis que l'ADN non-spécifique, dépourvu de motifs spécifiques, peut être lié de manière arbitraire par des protéines, impactant ainsi des processus biologiques tels que la régulation génique et la structure de la chromatine. Celles-ci sont à 1 μ M de l'ADN non-spécifique et spécifique et à différentes concentrations de protéines trimériques soit; 0, 0.3, 0.7, 1.5, 3.1, 6.2, 12.5 et 25 μ M, il est possible d'observer la force de liaison à l'ADN (figure 3.9a,b). Malgré la capacité modérée de la protéine HSFB2b à se lier à l'ADN non-spécifique, il est notable que, dans les deux passages, cette protéine présente une affinité significativement plus élevée pour l'ADN spécifique. La tendance observée dans cette analyse est représentée graphiquement par des courbes sigmoïdales qui illustrent la liaison de l'ADN non spécifique et spécifique à différentes concentrations de la protéine HSFB2b (figure 3.9c). Ces courbes permettent de visualiser la force de liaison entre la protéine et l'ADN. L'analyse révèle qu'aucun Kd ne peut être déterminé pour l'ADN non

spécifique, avec un K_d estimé à plus de $25\mu\text{M}$, tandis que le K_d pour l'ADN spécifique est mesuré à 2.5 et $2.3\mu\text{M}$ pour le premier et le deuxième passage respectivement.

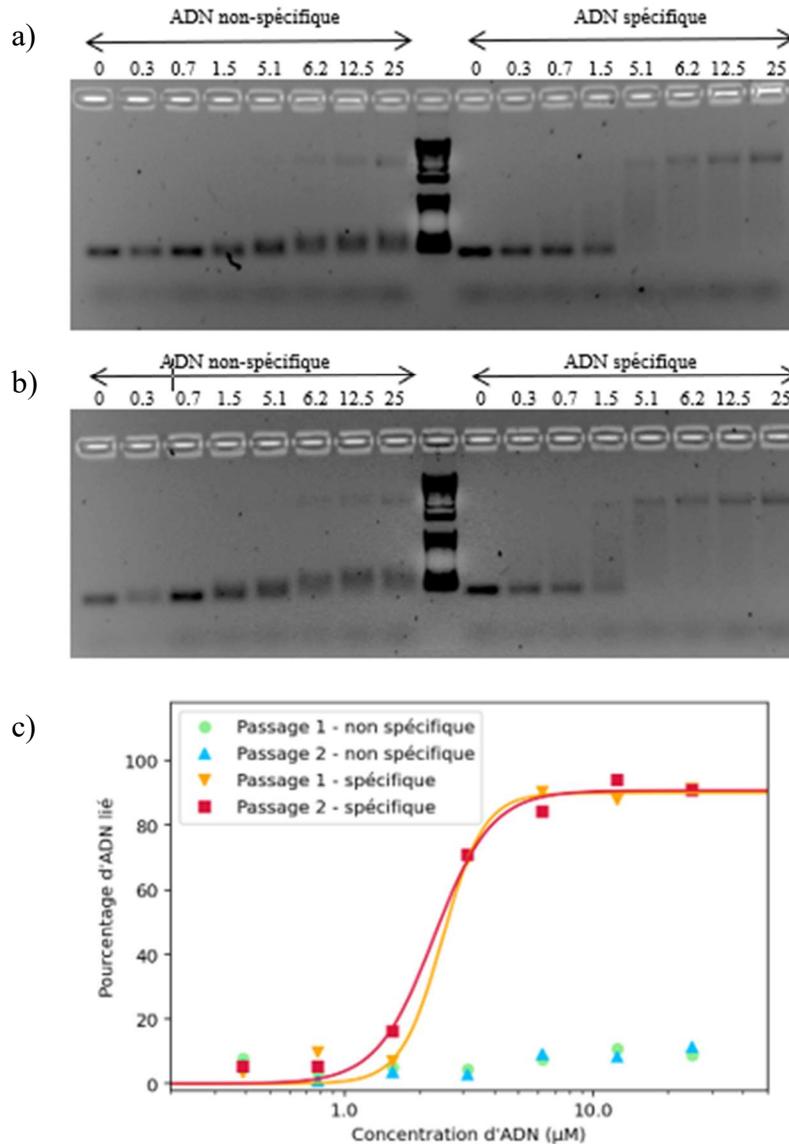


Figure 3.9 Retard sur gel de la protéine HSF2B2b présence d'ADN spécifique et non-spécifique. a) Gel d'agarose de la protéine HSF2B2b du 1^{er} passage. b) Graphique sigmoïdale de la liaison d'ADN non-spécifique à la protéine HSF2B2b. c) Courbe sigmoïdale représentant les interactions de liaison à l'ADN, non-spécifique et spécifique, observées aux passages 1 et 2. L'acquisition des images a été faite à l'aide du ChemiDoc MP.

Les essais de retard sur gel confirment la liaison de l'ADN par la protéine HSF2B2b et démontrent que la protéine adopte une conformation propice à son interaction avec l'ADN. La figure 3.10 présente l'essai enzymatique de la protéine HSF2B2b en présence d'ADN spécifique *in vitro* dans

le cadre de la réaction de SUMOylation. À droite de la figure, les différentes composantes de la machinerie de la SUMOylation sont observables, notamment la E1, la E2, la E3, et la protéine cible Hsfb2b. Du côté gauche de la figure, la réaction enzymatique de la SUMOylation est arrêtée après à différents intervalles d'incubation, à savoir 0, 5, 10, 20, 40, et 80 minutes. Il convient de noter que le gel SDS-PAGE est visualisé uniquement avec du TCE (2,2,2-trichloroéthane) et non avec le bleu de Coomassie, expliquant l'absence de visibilité du SUMO, qui nécessite la présence du bleu de Coomassie pour être détecté. Cependant, l'observation détaillée suggère que la protéine Hsfb2b subit la SUMOylation, car à partir de 40 minutes de réaction, le substrat disparaît progressivement. Il est cependant difficile de visualiser les formes SUMOylées de la protéine.

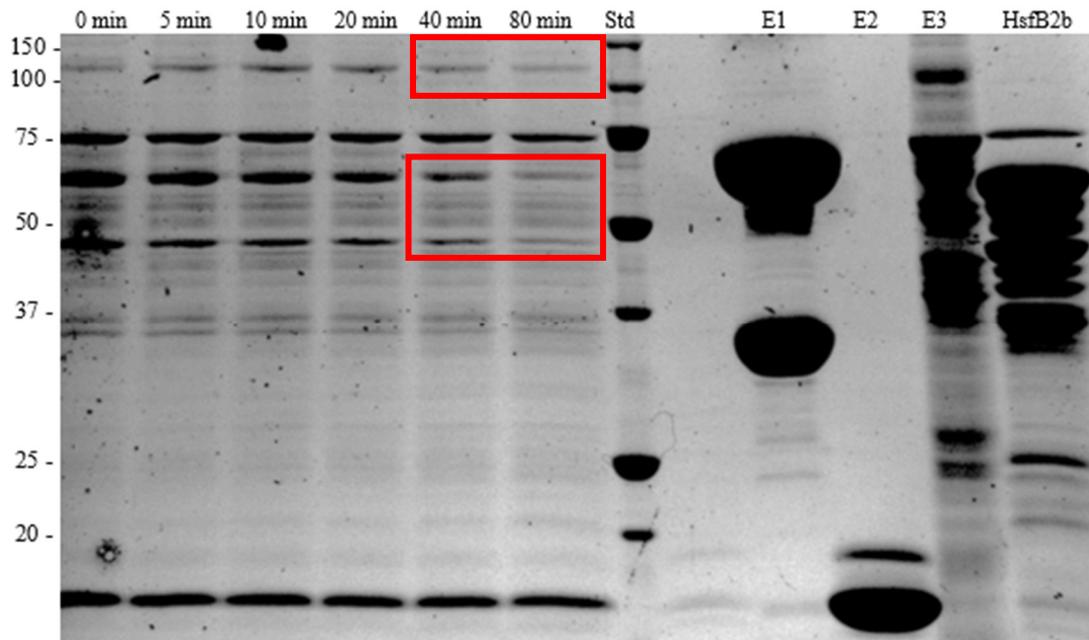


Figure 3.10 Essai enzymatique de la protéine Hsfb2b en présence d'ADN spécifique à SIZ-1. La réaction de SUMOylation contient la E1, la E2, la E3, SUMO et de l'ATP incubée pendant 5, 10, 20, 40 et 80 minutes. L'acquisition des images a été faite à l'aide du ChemiDoc MP.

Explorons la modélisation de la protéine Hsfb2b en interaction avec la machinerie de SUMOylation, telle que représentée dans la figure 3.11. Cette représentation a été élaborée en utilisant la technologie AlphaFold conjointement avec le logiciel PyMOL, afin de permettre une analyse approfondie des multiples interactions en jeu.

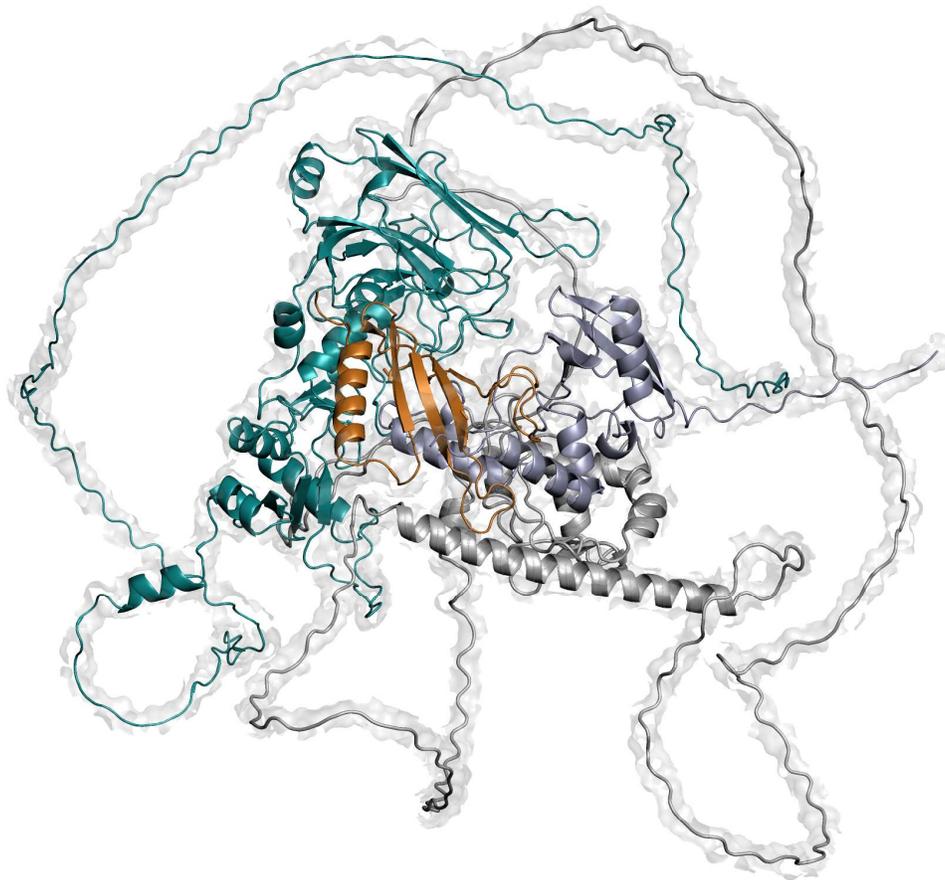


Figure 3.11 Modélisation de la protéine HSF2b avec la machinerie de la SUMOylation par AlphaFold. L'enzyme E2 est en mauve, la E3 ligase en bleu cyan, le facteur de choc thermique HSF2b en gris et la protéine SUMO en orange.

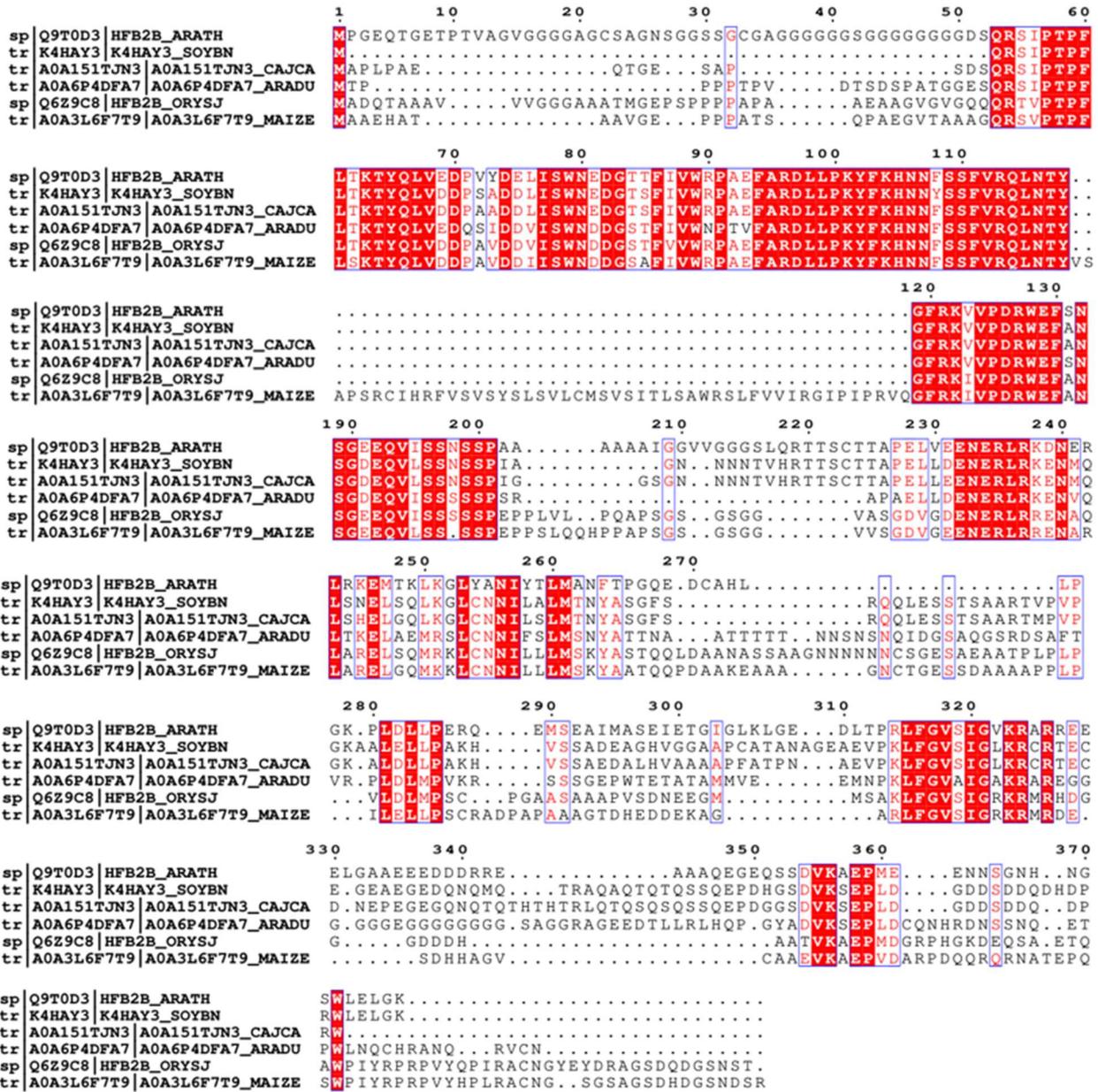


Figure 3.12 Conservation de la séquence protéique HSFB2b chez différentes espèces; le soja, le pois, les arachides, le riz et le maïs. Les résidus colorés en blanc sur fond rouge correspondent à des séquences hautement conservées tandis que les résidus colorés en rouge sur fond blanc sont semi-conservés.

L'investigation de la conservation de la protéine HSFB2b au sein de diverses cultures, notamment le soja, le pois, les arachides, le riz et le maïs, revêt une importance cruciale pour élucider son

rôle fonctionnel et son impact éventuel sur une variété d'organismes. Cette analyse comparative vise à établir la stabilité de la séquence de la protéine HSFB2b au fil de l'évolution, ainsi que l'existence de variations spécifiques entre différentes espèces. Les divergences dans la séquence primaire peuvent exister entre les organismes sans altérer la structure native des protéines. L'étude de la conservation de la protéine HSFB2b dans différentes cultures offre des indications précieuses quant à sa fonction biologique. La persistance de cette protéine à travers plusieurs espèces suggère une fonction essentielle, fournissant des indices cruciaux sur son rôle biologique. De surcroît, étant donné que différentes cultures sont souvent confrontées à des conditions environnementales distinctes, l'analyse de la conservation de la protéine HSFB2b dans ces contextes peut révéler des adaptations spécifiques à chaque espèce pour relever les défis environnementaux particuliers. La compréhension de la conservation de la protéine HSFB2b dans des cultures agricoles majeures telles que le soja, le pois, les arachides, le riz et le maïs peut avoir des implications significatives dans le domaine agricole. Cette étude peut fournir des informations cruciales sur la résistance de ces cultures au stress thermique ou à d'autres stress environnementaux. En ce qui concerne la méthodologie, pour réaliser l'alignement des séquences de la protéine HSFB2b et évaluer sa conservation, on a généralement recours à des méthodes d'alignement multiple. Ces méthodes comparent les séquences de protéines provenant de différentes espèces et identifient les positions où les acides aminés sont conservés, mettant en lumière les parties de la protéine demeurées inchangées au fil de l'évolution. Le processus d'alignement multiple peut être effectué à l'aide de programmes informatiques tels que MAFFT et EsPript, facilitant ainsi l'analyse des similitudes et des différences entre les séquences de protéines des différentes cultures.

CHAPITRE 4

DISCUSSION

4.1 Analyse biochimique

4.1.1 Purification des protéines HSFs

Parmi les 21 facteurs de transcription de la famille des HSFs identifiés chez *Arabidopsis thaliana*, nous avons réussi à cloner 10 protéines HSFs, à savoir HSFA1a, HSFA1e, HSFA2, HSFA7a, HSFA7b, HSFA9, HSFB1, HSFB2b, HSFB3 et HSFC1. Dans le cadre de notre étude, nous avons pu exprimer et purifier cinq de ces protéines, à savoir HSFA1a, HSFA1e, HSFA7b, HSFB3 et HSFC1, en utilisant des méthodes de chromatographie par affinité et par exclusion stérique. Malgré le succès de la purification, des défis subsistaient quant à l'expression optimale de ces protéines. En particulier, le clivage à la TEV, nécessaire pour libérer les protéines de leurs étiquettes de fusion, s'est révélé incomplet. Cette observation soulève des interrogations quant à l'efficacité de l'enzyme TEV ou à la possibilité d'une agrégation de protéines pendant la période prolongée de clivage d'environ 16 heures. Concernant la protéine HSFA9, bien qu'elle ait été exprimée avec succès et que la protéine soit soluble, le processus de clivage n'a pas fonctionné comme prévu. Les raisons de cet échec restent à élucider. Ce résultat suscite des interrogations sur la stabilité de la séquence de clivage ou d'autres facteurs qui pourraient influencer cette réaction spécifique. Malgré les succès dans l'expression et la purification de certaines protéines HSFs, des défis subsistent donc quant à l'optimisation de l'expression et aux processus de clivage. Ces observations soulignent l'importance de comprendre en détail les caractéristiques de chaque protéine et les conditions expérimentales pour garantir le succès des étapes cruciales de l'étude protéomique. De plus, certaines protéines semblent migrer à des poids moléculaires apparents différents de celui calculé. Cela peut être dû à une composition atypique en acides aminés, ce qui peut entraîner des comportements anormaux sur les gels, y compris un poids apparent différent du poids réel. La formation de structures secondaires ou tertiaires particulières peuvent également

affecter la mobilité des protéines lors de l'électrophorèse sur gel. Enfin, des interactions avec d'autres protéines ou des composés présents dans le milieu de l'électrophorèse peuvent aussi altérer la migration des protéines et conduire à des variations dans les poids moléculaires apparents observés sur le gel.

Des éléments récents suggèrent que la régulation de HSFA2 chez *Arabidopsis thaliana* diffère de celle observée chez la tomate, ce qui souligne la nécessité d'une approche précise et contextuelle dans l'étude des processus biologiques. Chez la tomate, HSFA1 est le régulateur principal, et sa réduction drastique a entraîné une diminution spectaculaire de HSFA2 (Mishra et al., 2002; von Koskull-Doring et al., 2007). En revanche, chez *Arabidopsis*, la suppression simultanée de HSFA1a/A1b n'a pas eu d'effet sur l'expression de HSFA2 (Busch et al., 2005). Il suggère que HSFA2 joue un rôle protecteur dans des conditions de stress thermique répété (Baniwal et al., 2004; Schramm et al., 2006; Nishizawa et al., 2006; Ogaqa et al., 2007). La protéine HSFA2 est une protéine clé dans l'induction des gènes du choc thermique chez les plantes et promouvait la production des HSP nécessaire pour la thermotolérance (Nishizawa et al., 2006; Schramm et al., 2006, 2008; Charng et al., 2007). Néanmoins, il ne nous a pas été possible d'exprimer la protéine HSFA2. Plusieurs facteurs peuvent contribuer à l'échec de la production de HSFA2. Parmi ceux-ci, la détérioration, la dégradation ou une concentration insuffisante de l'ADN peuvent être citées en exemple. Un ADN de qualité médiocre présente une intégration moins efficace dans les bactéries. Une raison peut être que les bactéries ne sont pas dans des conditions optimales lors de la transformation, par exemple la température ou le pH. Une exposition excessive à la chaleur pendant l'incubation peut également causer des dommages à l'ADN et aux bactéries. Il est possible qu'il y ait eu des erreurs humaines lors de la préparation des réactifs, de la manipulation des échantillons ou de l'exécution des étapes du protocole de transformation, entraînant ainsi un échec de la transformation.

Les gènes *hsfa7a* et *hsfb1* n'ont pas donné de résultats dès l'étape d'amplification, ce qui pourrait être attribuable à une mauvaise conception des amorces. Les amorces doivent être conçues pour s'apparier spécifiquement à la séquence cible, et leurs températures de liaison doivent être correctement adaptées pour permettre une hybridation efficace avec l'ADN cible. Une dégradation de l'ADN pendant la réaction enzymatique peut être due à l'exposition à la chaleur excessive ou des conditions de réactions incorrectes. Pour vérifier si la dégradation de l'ADN

observée pendant la réaction est due à des conditions expérimentales inappropriées, des contrôles positifs peuvent être utilisés. Un témoin non traité (NT) d'ADN permet de comparer l'intégrité de l'ADN avant et après la réaction, tandis qu'un témoin de dégradation contrôlée expose une partie de l'ADN à des conditions de dégradation connues. De plus, l'utilisation d'un témoin de dégradation induite, où une substance ou un agent connu pour induire la dégradation de l'ADN est introduit, permet d'évaluer la sensibilité du système expérimental. Ces contrôles positifs offrent une méthode fiable pour déterminer si la dégradation de l'ADN est le résultat des conditions expérimentales de la réaction ou d'autres facteurs externes. Des erreurs expérimentales peuvent ainsi causer des problèmes d'amplification tels que des erreurs dans la dilution ou une préparation d'échantillons inadéquate ainsi que l'oubli de différents réactifs dans la réaction. Finalement, la protéine HSF2b est la seule protéine qui a passé toutes les étapes : le gène codant pour cette protéine a été amplifié et cloné, la protéine a été exprimée de manière soluble et a pu être clivée au moyen de la protéase TEV. Une fois la protéine pure a été obtenue, il a été possible de faire des tests d'essai retard sur gel ainsi que des essais enzymatiques qui ont été un succès.

4.1.2 HSF2b : Purification complète

Le facteur de transcription HSF2b fait une taille de 377 acides aminés. Cette protéine est considérée comme un répresseur négatif des facteurs HSF. Elle agit en inhibant l'activité transcriptionnelle des HSF, ce qui réduit l'expression des gènes de protéines de choc thermique en l'absence de stress thermique (Ikeda et al., 2011). HSF2b contribue également à la régulation de la réponse au stress thermique. En absence de stress, elle empêche une activation excessive des gènes de protéines de choc thermique par les HSFs. Cependant, en cas de stress thermique, la protéine subit une régulation post-traductionnelle qui, en libérant son action inhibitrice, favorise l'activation des gènes de facteurs de choc thermique (Ikeda et al., 2011). Comme l'indique la figure 3.8 (b), la purification de la protéine a été efficace. D'ailleurs, il a même été possible de charger le volume de flow-through (partie de l'éluat qui ne s'est pas liée à la matrice de la colonne) de nouveau sur la colonne (2^e passage), étant donné qu'il restait beaucoup de protéines dans cette fraction. Ceci se fait par l'observation de bande intense pure et un bon rendement vers 55-60 kDa. Suite au clivage à la TEV, il a été possible d'utiliser la chromatographie par exclusion

stérique afin de pouvoir isoler la protéine d'intérêt HSFB2b à 40 kDa (Figure 3.8, c et d). La protéine pure et isolée est identifiable dans la fraction 6 dans les deux figures (Figure 3.8, c et d), cependant, la figure (d) dénote un rendement moindre qu'à la figure (c) tout simplement parce que l'échantillon du 2^e passage avait un rendement plus faible comparativement au 1^{er} passage. La fraction 7 dans le schéma présenté indique une concentration plus élevée de la protéine d'intérêt par rapport à la fraction 6, tout en présentant un rendement supérieur. Cependant, il convient de noter que cette fraction de purification contient une impureté qui pourrait être la HSP70, basée sur des caractéristiques telles que la taille apparente sur gel SDS-PAGE.

4.1.3 HSFB2b : Liaison spécifique à HSE

Il est crucial de noter que les HSFs (facteurs de transcription sensibles à la chaleur) et les HSPs (protéines de choc thermique) jouent un rôle significatif dans la réponse cellulaire au stress thermique, étant étroitement impliqués dans la régulation transcriptionnelle des gènes codant pour les protéines de choc thermique. Lorsque les cellules font face à un stress thermique, les HSFs sont activés. Ces HSFs agissent en tant que facteurs de transcription qui se lient spécifiquement à des séquences d'ADN appelées éléments de réponse à la chaleur (HSE - Heat Shock Element) situés dans les promoteurs des gènes codant pour les HSPs (Zhang et al., 2014). La spécificité de la liaison entre les HSFs activés et les HSEs dans les promoteurs des gènes de HSPs est renforcée par le fait que les HSFs adoptent une structure trimérique, formant ainsi une entité composée de trois sous-unités. Cette organisation trimérique des HSFs a des implications possibles sur la SUMOylation des protéines HSFs, influençant potentiellement leur activité transcriptionnelle et leur capacité à réguler efficacement les gènes de protéines de choc thermique en réponse au stress thermique.

Lors d'un choc thermique, HSF passe d'un équilibre monomère-dimère à l'état trimérique, acquérant ainsi une compétence de liaison à l'ADN et libérant l'ARN polymérase en pause pour la transcription des gènes de choc thermique (Kmiecik et al., 2021). La réponse au choc thermique est atténuée par la monomérisation médiée par Hsp70 et la dissociation de HSF de l'ADN (Kmiecik et al., 2021). Plusieurs mécanismes possibles expliquent comment la SUMOylation

pourrait influencer la transcription des gènes de choc thermique induite par HSF : (i) la SUMOylation pourrait empêcher la trimérisation de HSF, (ii) la SUMOylation pourrait interférer avec la liaison de HSF à sa séquence de reconnaissance (HSEs) dans l'ADN promoteur ou provoquer une dissociation spontanée de l'ADN ; (iii) la SUMOylation de HSF pourrait modifier la cinétique de la dissociation de HSF1 de l'ADN médiée par Hsp70 ; et (iv) l'interaction avec la machinerie de transcription pourrait être modulée, et la réinitiation de la transcription pourrait être inhibée après la SUMOylation de HSF (Kmiecik et al., 2021). La SUMOylation des trimères de HSF1, catalysée par le système minimal de SUMOylation composé des enzymes E1 (Aos1/Uba2) et E2 (Ubc9), accroît l'efficacité de SUMOylation deux à trois fois par rapport aux monomères de HSF HSF (Kmiecik et al., 2021). Cette SUMOylation renforce l'interaction des HSF avec les éléments de réponse au stress thermique (HSEs), favorisant ainsi la transcription des gènes codant pour les protéines de choc thermique (HSPs). Les HSPs, synthétisées en réponse à cette activation des gènes du HSP par les HSF SUMOylés, jouent un rôle crucial dans la protection cellulaire contre les dommages thermiques en préservant la stabilité protéique et en facilitant la réparation des protéines dénaturées par le stress thermique.

Suite à la purification par chromatographie par exclusion stérique, un essai de retard sur gel a été fait. Cette technique permet d'étudier la liaison d'une protéine à une séquence d'ADN spécifique. Une telle interaction a été détectée, entraînant un effet sur la mobilité de l'ADN. En termes de coopérativité, un phénomène intrigant a été observé lors de l'analyse du retard sur gel des HSFs. Normalement, on s'attendrait à observer trois retards distincts, correspondant à chaque protéine HSF qui se lie à l'ADN, rappelant ainsi que le complexe formé est un trimère. Cependant, une observation singulière se démarque sur le retard sur gel. Ce résultat suggère une forme de coopération ou d'interaction entre les protéines HSFs impliquées. Il est possible que la coopérativité entre les protéines HSFs découle de la liaison synchronisée à l'ADN par la forme trimérique de HSFb2b, dont la taille est compatible avec la formation de trimères plutôt que de monomères. Cette coordination pourrait entraîner un unique retard sur gel plutôt que trois retards distincts, suggérant une interaction entre les protéines HSF pendant le processus de liaison à l'ADN. Cela indique une régulation complexe et interdépendante de leur activité. Cette observation ouvre la porte à une exploration plus approfondie de la manière dont les protéines HSF interagissent et coopèrent lors de la liaison à l'ADN, soulignant la complexité du mécanisme

régulateur qui sous-tend la réponse au choc thermique. La figure 3.9.a démontre le gel sur retard de la fraction 6 de la chromatographie par exclusion stérique (1^{er} passage). À gauche, la protéine HSFB2b en interaction avec une séquence d'ADN non-spécifique et à droite, la protéine HSFB2b en interaction avec une séquence d'ADN spécifique.

Dans le contexte de la réponse au choc thermique et des éléments de choc thermique (Heat Shock Elements, HSEs), il existe une distinction entre les HSEs spécifiques et non spécifiques. Les HSEs font référence à une séquence d'ADN spécifiquement reconnue et liée par les facteurs de transcription du choc thermique (HSFs) qui se trouve généralement dans les promoteurs des gènes qui sont induits pendant la réponse au choc thermique. La liaison des HSFs aux HSEs spécifiques initie la transcription des protéines de choc thermique (HSPs), qui sont cruciales pour la protection cellulaire et la survie dans des conditions de stress, telles que l'exposition à des températures élevées. Les HSEs non-spécifiques se composent de séquences aléatoires qui peuvent ne pas correspondre parfaitement à la séquence consensuelle des HSE, mais qui restent identifiables et potentiellement liées par les HSFs. La présence de ces HSEs aléatoires favorise une réponse plus souple et adaptable au stress, permettant ainsi l'activation des gènes de choc thermique dans un éventail plus large de conditions. Cette flexibilité dans la reconnaissance des HSE non spécifiques contribue à l'adaptabilité des cellules face à divers facteurs de stress. Il est clairement observable que la protéine présente une affinité notable pour la séquence HSE spécifique, induisant ainsi une liaison spécifique à l'ADN. Cette interaction conduit à la formation d'un complexe protéine-ADN dont la mobilité dans le gel est significativement réduite par rapport à l'ADN non lié, comme illustré par la migration différenciée des bandes sur le gel électrophorétique du premier passage (Figure 3.9.a,b). Dans cette étude, l'accent est mis sur l'affinité entre les protéines HSF (Heat Shock Factors) et l'ADN, crucial pour la régulation des gènes impliqués dans la réponse au stress thermique. Les résultats graphiques présentent des courbes sigmoïdales, reflétant la relation entre la concentration de la protéine HSFB2b et la force de liaison à l'ADN, illustrant ainsi l'affinité entre les deux. Cette affinité est particulièrement significative car elle détermine la capacité de HSFB2b à se lier spécifiquement à l'ADN, régulant ainsi l'expression des gènes de réponse au stress thermique. L'analyse révèle des différences notables dans l'affinité des HSF pour l'ADN non spécifique par rapport à l'ADN spécifique. Alors que l'ADN non spécifique ne montre aucun Kd déterminable, suggérant une faible affinité, les

valeurs mesurées de K_d pour l'ADN spécifique sont significativement plus faibles, soit, 2.5 et 2.3 μ M, indiquant une forte affinité entre les HSFs et les séquences d'ADN spécifiques (figure 3.9c). Cette forte affinité est essentielle pour le ciblage précis des gènes impliqués dans la réponse au stress thermique, permettant ainsi une régulation efficace de cette voie critique de survie cellulaire. Cette compréhension de l'affinité des HSF pour l'ADN fournit des informations précieuses sur les mécanismes moléculaires sous-jacents à la réponse au stress thermique.

Ces observations soulignent l'importance de combiner différentes méthodes d'analyse pour obtenir une compréhension plus complète des interactions moléculaires. Il se peut que des facteurs tels que la spécificité de la séquence, la concentration ou d'autres conditions expérimentales influent sur les résultats, mettant en lumière la nécessité d'une approche intégrée pour interpréter avec précision les données expérimentales. La liaison à l'ADN et la SUMOylation des protéines HSFs (Heat Shock Factors) sont des processus interdépendants et régulatoires chez les humains. La liaison à l'ADN peut influencer la SUMOylation des HSFs, et réciproquement, la SUMOylation peut également modifier leur capacité à se lier à l'ADN. Lorsque les HSFs se lient à des séquences spécifiques d'ADN, telles que les éléments de réponse à la chaleur (HSEs) présents dans les promoteurs des gènes de choc thermique, cela peut induire des changements conformationnels dans les protéines HSFs. Ces changements conformationnels peuvent, à leur tour, affecter la susceptibilité des HSFs à la SUMOylation. En d'autres termes, la liaison à l'ADN peut influencer positivement ou négativement la SUMOylation des HSFs en modulant la disponibilité des sites de SUMOylation. À l'inverse, la SUMOylation des HSFs peut influencer sur leur liaison à l'ADN de diverses manières, comme illustré dans l'étude mettant en évidence l'effet inhibiteur de la SUMOylation sur l'interaction entre HSF2 et l'ADN (Tateishi et al., 2009). Cette démonstration repose sur une analyse biochimique utilisant HSF2 recombinant isolé (Tateishi et al., 2009). L'investigation par résonance magnétique nucléaire (RMN) du domaine de liaison à l'ADN de HSF2 SUMOylé révèle la flexibilité de la partie SUMO par rapport au domaine de liaison à l'ADN, sans établir d'interface non covalente ni d'effet structural sur ce dernier (Tateishi et al., 2009). Ces constatations suggèrent que l'ajout de SUMO a un impact négatif sur la formation du complexe protéine-ADN, induisant une interférence stérique distribuée de manière aléatoire (Tateishi et al., 2009). La SUMOylation pourrait réguler la trimérisation des HSFs, affectant ainsi leur capacité à se lier de manière optimale à l'ADN. De

plus, la SUMOylation peut interférer directement avec l'interaction entre les HSFs et les séquences spécifiques d'ADN, entravant ainsi la liaison aux HSEs. Cette modification post-traductionnelle peut également influencer la dynamique de dissociation des HSFs de l'ADN, modifiant ainsi leur efficacité dans la régulation de la transcription. De plus, il est envisageable de conduire des expériences analogues à celles décrites par Tateishi en 2009, afin de déterminer l'impact de la SUMOylation de HSFb2b sur sa liaison à l'ADN. Ceci permettrait de vérifier si les observations faites pour la protéine HSF2, une protéine humaine où les sites de SUMOylation sont positionnés différemment, sont également valides pour HSFb2b.

La protéine HSFb2b a bien été purifiée avec un bon rendement, on a également démontré qu'elle interagit avec de l'ADN spécifique, comprenant des séquences HSEs. À ce moment, il était intéressant d'étudier la SUMOylation des protéines HSFs, car elle module leur activité durant le stress thermique. La SUMOylation des HSFs peut également moduler la liaison à l'ADN en modifiant leur structure tridimensionnelle. Cela permet une interaction plus stable avec les HSEs dans les promoteurs des gènes de HSP qui peut alors favoriser une réponse au stress thermique. L'exploration de la SUMOylation de la protéine HSFb2b en présence de la séquence HSE revêt une importance significative dans notre étude. Le processus de SUMOylation de la protéine HSFb2b semble être intensifié lorsqu'elle interagit avec la séquence HSE, ce qui suggère une corrélation directe entre la reconnaissance de cette séquence par la protéine et son processus de SUMOylation. Cependant, il est important de souligner que cette conclusion découle de l'observation de l'interaction entre la protéine et la séquence HSE, sans prendre en considération l'essai en l'absence d'ADN, lequel pourrait apporter des données supplémentaires pour nuancer cette corrélation. Ces résultats suggèrent que la séquence HSE pourrait agir comme un signal régulateur qui module la modification post-traductionnelle de la protéine HSFb2b. L'analyse de la SUMOylation en présence de la séquence HSE offre ainsi un aperçu plus approfondi des mécanismes de régulation moléculaire impliquant la protéine HSFb2b. Cette observation suggère également que la séquence HSE pourrait jouer un rôle clé dans la modulation des activités fonctionnelles de la protéine HSFb2b via la SUMOylation, ce qui ouvre la voie à de futures investigations visant à élucider les implications physiologiques de cette régulation fine dans le contexte du stress thermique ou d'autres stimuli cellulaires.

Il existe plusieurs E3 ligases, toutefois, il y a seulement quatre SUMO E3 ligases impliquées dans la SUMOylation des protéines. La SIZ-1 (*SAP et MIZ-1*) est une E3 ligase de type PIAS (Protein Inhibitor of Activated STAT) qui joue un rôle essentiel dans la SUMOylation de diverses protéines chez *Arabidopsis*, régulant ainsi plusieurs processus biologiques, y compris la réponse au stress (Miura et al., 2005; Jmii et Cappadocia, 2021). Les protéines inhibitrices de l'activation de STAT 1/2 de type PIAS-like 1 et 2 sont des E3 ligases impliquées dans l'association à la régulation du développement des organes floraux (Tomanov et al., 2014; Jmii et Cappadocia, 2021). MMS21 (Methyl Methanesulfonate Sensitivity 21) est une autre E3 ligase de type PIAS (Protein Inhibitor of Activated STAT) chez *Arabidopsis thaliana*. La protéine MMS21 joue un rôle important dans la régulation de la SUMOylation des protéines. Elle est impliquée dans plusieurs processus cellulaires, y compris la réponse au stress et la régulation de la croissance et du développement chez la plante (Ishida et al., 2009; ; Jmii et Cappadocia, 2021). Ces dernières régulent la conjugaison d'au moins 100 protéines (Rytz et al., 2018; ; Jmii et Cappadocia, 2021). Le choix de SIZ-1 (Ligase 1 Contenant le Domaine SAP et MIZ 1) en tant qu'E3 ligase SUMO pour les HSFs (Facteurs de Choc Thermique) repose sur plusieurs considérations. Tout d'abord, SIZ-1 a démontré une spécificité élevée pour la SUMOylation de divers substrats, suggérant sa capacité à catalyser efficacement cette modification sur les HSFs. De plus, SIZ-1 est la principale SUMO E3 ligase qui répond à un choc thermique (Viestra, 2018). Étant donné le rôle crucial des HSFs dans la régulation de la réponse au choc thermique, le choix de SIZ-1 comme E3 ligase SUMO est logique pour moduler spécifiquement l'activité transcriptionnelle des HSFs et sa contribution à la régulation de la réponse cellulaire au choc thermique. Par exemple, HSFB2b est influencé par SIZ-1 liée au stress thermique avec un p value <0.01 ayant un niveau très élevé de significativité (Rytz et al., 2018). Étant donné que notre protéine d'intérêt, HSFB2b, se lie au HSE, et que la SIZ-1 se lie également au HSE et à l'ADN. La SIZ-1, jouant un rôle crucial dans la réparation de l'ADN, favorise la SUMOylation en se liant à des protéines transcriptionnelles. Dans ce processus, la protéine HSFB2b subit une SUMOylation, un phénomène possiblement stimulé par la liaison de HSFB2b à une séquence HSE spécifique.

La SIZ-1 est connue pour son rôle dans la SUMOylation de divers substrats, y compris d'autres protéines, et elle peut également subir une auto-SUMOylation. L'auto-SUMOylation de la SIZ-1 n'est pas inhabituelle dans le contexte de la SUMOylation. De nombreuses protéines E3 ligases

de SUMOylation, y compris la SIZ-1, ont la capacité de s'auto-SUMOyler en plus de SUMOyler d'autres protéines (Casto et al., 2022). L'auto-SUMOylation peut réguler l'activité de la protéine elle-même, modifiant potentiellement sa stabilité, son interaction avec d'autres protéines, ou d'autres aspects de sa fonction biologique (Casto et al., 2022). SIZ-1 présente davantage de sites de sites de SUMOylation potentiels et la réaction se fait directement à même le complexe ternaire E3/E2~SUMO tandis que la SUMOylation de HSF2 nécessite quant à lui la formation d'un complexe HSF2/E3/E2~SUMO. Dans le contexte de la SUMOylation, la SIZ-1 peut être responsable de la SUMOylation de plusieurs substrats, y compris elle-même. Cela peut être lié à sa propre régulation et à son rôle dans la régulation d'autres protéines au sein de la cellule. La protéine HSF, en revanche, peut subir la SUMOylation en réponse à divers stimuli, tels que le stress thermique, ce qui peut moduler son activité transcriptionnelle et son interaction avec d'autres protéines. Les deux processus sont des mécanismes importants de régulation post-traductionnelle qui contribuent à la diversité fonctionnelle des protéines dans la cellule. La purification des protéines offre une base solide pour les analyses ultérieures, notamment dans le contexte des prédictions de la structure protéique. Une fois les protéines isolées et caractérisées, l'utilisation de la machinerie de la SUMOylation devient cruciale pour comprendre les mécanismes de régulation fine, modulant ainsi les interactions moléculaires et la fonction biologique des protéines concernées.

4.2 Analyses bio-informatiques : Prédiction des sites de SUMOylation potentiels

L'étude de la SUMOylation chez les facteurs de transcription d'*Arabidopsis thaliana* a également été faite avec une étude d'approches bio-informatiques. Pour prédire ou identifier les sites de SUMOylation potentiels, plusieurs indicateurs ou prédicteurs ont été utilisés. Avec l'aide des logiciels JASSA (Joint Analysis of SUMOylation and Acetylation) (<http://www.jassa.fr>) et GPS-SUMO (Group based Prediction System for SUMOylation Sites) (<https://sumo.biocuckoo.cn/>). Bien que ces deux outils bio-informatiques ont l'objectif similaire de prédire les sites de SUMOylation, il peut y avoir des différences notoires dans leur prédiction en raison de leur méthodologie, des algorithmes et des ensembles de données utilisées. L'utilisation d'informations sur l'acétylation par JASSA est justifiée car l'acétylation est une modification post-traductionnelle

courante qui peut interagir avec la SUMOylation. En intégrant des données sur l'acétylation dans le processus de prédiction des sites de SUMOylation, JASSA peut tenir compte des interactions ou des effets synergiques entre ces deux types de modifications. Par exemple, certaines régions acétylées peuvent favoriser ou inhiber la SUMOylation en modifiant la structure ou la charge de la protéine. Ainsi, en considérant les informations d'acétylation, JASSA peut améliorer la précision de ses prédictions en capturant les relations complexes entre différentes modifications post-traductionnelles et en identifiant les sites susceptibles d'être SUMOylés en fonction de ces interactions. (Beauclair et al., 2015). Les informations fournies englobent les séquences contenant les sites consensus ainsi que leurs scores respectifs de SUMOylation. Chaque site se voit attribuer un score en fonction de plusieurs critères, incluant les caractéristiques intrinsèques de la séquence protéique, la structure secondaire locale, et la conservation évolutive. Ces éléments jouent un rôle crucial dans les prédictions et doivent être pris en compte pour une prédiction optimale. Les prédictions concernant les structures protéiques et les modifications post-traductionnelles, traitées par GPS-SUMO, sont souvent sujettes à la complexité des interactions moléculaires et des influences environnementales. Les diverses conformations que peuvent adopter les protéines en fonction des conditions expérimentales, ainsi que les multiples variables cellulaires qui peuvent influencer les modifications post-traductionnelles, ajoutent à la complexité du processus de prédiction. Certains motifs ou combinaisons d'acides aminés peuvent favoriser la formation de structures spécifiques, ce qui est souvent pris en compte dans la prédiction de la structure secondaire, tels que les hélices alpha, les feuilletts bêta et les boucles. La capacité à distinguer avec précision entre ces différentes structures est essentielle pour une prédiction idéale. Les régions conservées au cours de l'évolution sont également d'une grande importance, car elles sont souvent cruciales pour la fonction de la protéine. Par conséquent, comparer une séquence donnée avec des séquences similaires chez d'autres espèces peut être bénéfique pour identifier les régions conservées. Un score élevé indique une forte probabilité de SUMOylation à un site donné, tandis qu'un score faible indique une probabilité moindre de SUMOylation à ce site.

GPS-SUMO démontre son efficacité en évaluant à la fois la sensibilité et la spécificité. La sensibilité de GPS-SUMO se réfère à sa capacité à détecter correctement les vrais positifs, c'est-à-dire les sites de SUMOylation réels dans une protéine. En d'autres termes, elle mesure à quel point GPS-SUMO est efficace pour ne pas manquer de sites de SUMOylation présents. D'un

autre côté, la spécificité de GPS-SUMO se rapporte à sa capacité à éliminer les faux positifs, c'est-à-dire les sites prédits comme étant de la SUMOylation mais qui ne le sont pas en réalité. Elle indique à quel point GPS-SUMO peut distinguer avec précision les vrais sites de SUMOylation des autres sites potentiellement similaires. En évaluant à la fois la sensibilité et la spécificité, on obtient une mesure plus complète de la performance de GPS-SUMO dans la prédiction des sites de SUMOylation. Cette même approche s'applique également au logiciel JASSA, où la sensibilité et la spécificité sont des critères essentiels pour évaluer son efficacité dans la prédiction des sites de SUMOylation. Il est crucial de prendre en compte la diversité et la représentativité des ensembles de données utilisées pour entraîner et tester chaque prédicteur. Ces résultats sont obtenus par des approches basées sur des algorithmes novateurs. GPS-SUMO utilise une approche basée sur un groupe de motifs plutôt que sur un motif de consensus unique pour la prédiction des sites de SUMOylation. Cette méthode tient compte de la diversité des motifs de SUMOylation observés dans différentes protéines et contextes biologiques. En utilisant cette approche basée sur un groupe de motifs, GPS-SUMO tente de capturer la variabilité des motifs de SUMOylation plutôt que de se limiter à un motif de consensus rigide. Cela permet au logiciel d'identifier des sites de SUMOylation potentiellement plus diversifiés et de s'adapter à la complexité biologique des processus de SUMOylation. Il est important de noter que l'efficacité de la prédiction dépend de la qualité et de la diversité des motifs de SUMOylation disponibles dans la base de données d'entraînement de GPS-SUMO. Les bases de données utilisées pour entraîner l'outil sont cruciales pour garantir des prédictions précises et fiables.

Il a donc été possible de déterminer des sites de SUMOylation en respectant le motif consensus [Ψ -K-X-E], où [Ψ] représente un acide aminé hydrophobe, [K] étant une lysine, [X] est un acide aminé quelconque et [E] est un acide glutamique, avec un score élevé, moyen et faible. Cependant, notre analyse bio-informatique se focalise sur les sites de SUMOylation présentant un score élevé. Cette sélection découle des sites de SUMOylation affichant des scores élevés, expliquant ainsi l'absence de sites de SUMOylation pour certains HSFs. Le tableau 3.1 met en évidence les sites de SUMOylation au sein des protéines HSFs, caractérisés par des scores significativement élevés. Les protéines HSFs arborant des motifs consensus en rouge révèlent des sites de SUMOylation particulièrement importants, avec un consensus direct, tandis que les sites identifiés en rose pâle indiquent des sites où la lysine associée affiche un score élevé et un

consensus bidirectionnel. Les positions clés de lysine extraites de ce tableau incluent K169, K251 et K288 du HSFA1d, K167, K269 et K315 du HSFA2, K273 et K333 du HSFA3, K97, K165 et K217 du HSFA6a, K9 du HSFA6b, K225 et K280 du HSFA7b, K143 du HSFA8, K47 et K185 du HSFA9, K244 et K268 du HSFB1, K356 du HSFB2b, K242 du HSFB3, et enfin K202 du HSFC1. Généralement, plus la séquence de la protéine est longue plus il y a des sites de SUMOylation potentiels. De plus, ces sites de SUMOylation se retrouvent proches des domaines de liaison à l'ADN (DBD) et le domaine de HR-A/B (oligomérisation). La proximité des sites de SUMOylation avec des domaines cruciaux tels que le domaine de liaison à l'ADN et le domaine d'oligomérisation suggère diverses hypothèses quant à leur fonction potentielle. Ces hypothèses, bien qu'exigeant une validation expérimentale pour confirmation, indiquent que la SUMOylation des protéines peut influencer leur activité transcriptionnelle. Les sites de SUMOylation voisins du domaine de liaison à l'ADN pourraient moduler l'interaction protéine-ADN, régulant ainsi la transcription des gènes cibles. De même, ceux situés près du domaine d'oligomérisation pourraient jouer un rôle dans la régulation de la formation de complexes protéiques, impactant ainsi la fonction biologique de la protéine. La proximité des sites de SUMOylation avec des domaines critiques suggère également une régulation potentielle de la stabilité protéique ou de sa localisation subcellulaire. En outre, ces sites, en étant à proximité de domaines névralgiques, pourraient agir comme des intégrateurs de signaux de stress, participant ainsi à la régulation de la réponse cellulaire, notamment dans des conditions de stress.

Les sites de SUMOylation sont localisés à des positions différentes sur chaque protéine HSFs. Par exemple, on observe des sites sur des résidus de lysine particuliers, tels que K169, K251 et K288 du HSFA1d, K167, K269 et K315 du HSFA2, etc. Certains sites de SUMOylation sont spécifiquement identifiés par des motifs consensus en vert pour un site de SUMOylation consensus direct (Ψ -K-x-E), le consensus inverse en orange (E-x-K- Ψ) et le consensus bidirectionnel en rose. Chaque protéine HSF semble présenter plusieurs sites de SUMOylation, démontrant la complexité de la régulation post-traductionnelle de ces protéines. Des exemples spécifiques de multiples sites incluent K97, K165 et K217 du HSFA6a, et K244 et K268 du HSFB1, illustrant la diversité des sites sur une même protéine. Certains sites de SUMOylation se regroupent sur des régions spécifiques des protéines HSFs, ce qui pourrait suggérer une régulation coordonnée ou une fonction particulière de ces régions. La présence de motifs

consensus en rouge pourrait indiquer une concentration de sites importants dans des régions spécifiques des protéines HSFs, ce qui souligne potentiellement des zones cruciales pour la régulation de ces protéines. Les sites de SUMOylation semblent être distribués de manière spécifique sur les protéines HSFs, avec des positions clés et des regroupements dans certaines régions. Cette distribution suggère une régulation complexe et spécifique des protéines HSFs par la SUMOylation, avec des implications potentielles pour leur fonction biologique.

La SUMOylation des protéines est fortement stimulée en réponse à des conditions de stress thermique, comme mis en évidence dans une étude de Rytz en 2018. Dans ce contexte, il a été observé que davantage de protéines subissent la SUMOylation, particulièrement à des positions considérées comme "non-consensus". Ces positions non-consensus, lorsqu'elles subissent la SUMOylation, semblent être spécifiquement associées à des situations de stress thermique. Cette tendance souligne une régulation accrue de la SUMOylation dans des conditions de stress thermique, avec une préférence marquée pour des sites non-consensus, ouvrant ainsi des perspectives sur les mécanismes de régulation fine des protéines en réponse à ces contraintes environnementales. Cette observation renforce l'idée que la SUMOylation peut jouer un rôle crucial dans la réponse cellulaire au stress thermique en modulant la fonction des protéines impliquées. Face au stress, la cellule pourrait privilégier certaines voies de signalisation ou certaines protéines. Il est envisageable que la SUMOylation se concentre sur des lysines spécifiques au détriment d'autres, influençant ainsi la proportion de lysines appartenant à des motifs consensus qui sont effectivement SUMOylés. Des mécanismes de régulation négative pourraient également être déclenchés en réponse au stress, délibérément limitant la SUMOylation de certaines lysines associées à des motifs consensus (Enserink, 2015).

Plusieurs investigations ont été entreprises concernant le site de SUMOylation « K315 » du facteur de choc thermique HSFA2, tel que présenté dans le tableau 3.1. La SUMOylation de HSFA2 exerce une influence substantielle sur sa fonction. Parmi les divers substrats de SUMO, le facteur de choc thermique HSFA2 joue un rôle positif dans les conditions de stress thermique. La SUMOylation de HSFA2, médiée par SUMO1, induit un état inactif dans le noyau cellulaire, et la déSUMOylation de HSFA2 conduit à une forme active de cette protéine lorsqu'elle est exposée au stress thermique (Cohen-Peer et al., 2010). De plus, le résidu situé à la position 315 de cette protéine représente la cible principale pour l'interaction avec SUMO-1 (Cohen-Peer et al.,

2010). Une substitution de la lysine par une arginine est amplement suffisante pour abolir totalement l'interaction avec SUMO-1 (Cohen-Peer et al., 2010). Il est également crucial de noter que le site K315 est le principal, voire le seul, site de SUMOylation de HSFA2 qui subit cette modification *in vivo* (Cohen-Peer et al., 2010). Comme l'illustre la présente étude bio-informatique, la lysine en position 315 (K315) revêt une importance particulière, comme en témoigne le tableau 3.1 où elle affiche un score élevé avec un consensus direct. Les données du tableau révèlent que cette lysine obtient un score élevé de 0.92. Ces constatations soulignent l'importance cruciale de la lysine en position 315 dans le contexte de la SUMOylation de HSFA2, confirmée par ses performances remarquables dans les analyses avec un consensus direct.

Une étude analogue à la nôtre a été menée dans l'organisme du blé. Pour déterminer les résidus potentiels de conjugaison SUMO dans TaHSFA1, les auteurs ont employé les outils SUMO_{LOT} (<http://www.abgent.com/sumoplot>) et JASSA (<http://www.jassa.fr/index.php?m=jassa>) afin de prédire les sites de SUMOylation. Globalement, SUMO_{LOT} a révélé 7 sites potentiels (Lys77, Lys98, Lys103, Lys164, Lys175, Lys255, et Lys459), tandis que JASSA en a annoté 4 (Lys98, Lys175, Lys241, et Lys255) ils ont alors utilisé dorénavant K77, K98, K103, K164, K175, K241, K255, et K459 pour désigner ces acides aminés respectivement) (Wang et al., 2023). Ce même article révèle plusieurs conclusions significatives. Il identifie des résidus spécifiques, notamment le site K459, comme étant responsables de la modification SUMO, malgré leur caractère non canonique de conjugaison SUMO. De plus, une interaction directe entre TaHsfA1 et l'enzyme E3 ligase SUMO TaSIZ1 est mise en évidence, suggérant un rôle potentiel de TaSIZ1 dans la SUMOylation de TaHsfA1. La dynamique de la SUMOylation de TaHsfA1 en réponse au stress thermique est également observée, montrant une accumulation rapide de protéines conjuguées à SUMO après un traitement court de chaleur, suivie d'une diminution significative après un traitement prolongé. Par ailleurs, l'importance cruciale de la SUMOylation du résidu K459 dans la régulation de l'activité transcriptionnelle de TaHsfA1 est soulignée, où des mutations bloquant cette SUMOylation compromettent considérablement son activité transcriptionnelle. En outre, il est suggéré l'existence probable de protéases SUMO non caractérisées dans le blé, impliquées dans un cycle thermosensible de HsfA1 crucial pour l'adaptation du blé aux températures élevées. Enfin, plusieurs sites potentiels de SUMOylation dans TaHsfA1 sont identifiés, le site K459 étant confirmé comme site de modification SUMO, malgré des scores de prédiction inférieurs à

d'autres sites (Wang et al., 2023). Par ailleurs, des investigations antérieures ont exploré d'autres membres de la famille HSF dans notre plante modèle, Arabidopsis. Ces études ont abouti à l'identification de diverses protéines HSFs, dont HSFb2b, HSFA2, HSFA1d, HSFA5, HSFA1b, HSFA4a, et HSFC1, grâce à l'utilisation de la spectrométrie de masse (Rytz, et al., 2018).

À la suite de l'analyse des sites de SUMOylation à l'aide des logiciels GPS-SUMO et JASSA, il a paru particulièrement instructif d'explorer l'alignement des 21 facteurs de choc thermique (HSFs) afin d'examiner la conservation des lysines. Bien que l'étude de la conservation puisse offrir des indications pertinentes lors de la prédiction des sites de SUMOylation, elle ne constitue pas un indicateur direct des sites de SUMOylation propres à des protéines spécifiques. La conservation des gènes HSF dans différentes espèces est due à leur présence répandue et à leur similarité au niveau de la séquence ou de la structure. Cette présence généralisée suggère leur rôle crucial dans la réponse au stress thermique et leur importance évolutive. De plus, leur similitude structurelle ou séquentielle indique une fonctionnalité conservée dans la régulation de la réponse au stress thermique à travers les différentes lignées évolutives. L'examen de la conservation des HSFs permet de mieux appréhender l'évolution et la fonctionnalité dans la réponse au stress thermique. La figure 3.4 présente l'alignement des 21 HSFs, mettant en évidence les régions où les lysines avoisinantes semblent être particulièrement bien préservées. Notamment, la séquence des divers HSFs semble principalement afficher une plus grande conservation dans la figure 3.4.b, qui correspond au domaine de liaison à l'ADN où la plupart des lysines sont également conservées qui pourraient alors s'expliquer par leur participation à la liaison de l'ADN. Cependant, on observe également la conservation de lysines dans d'autres régions des séquences, notamment le domaine d'oligomérisation.

Étant donné que la SUMOylation des protéines est conditionnée par leur interaction avec les protéines E2 et/ou E3, nous avons développé des modèles pour ces interactions. L'analyse des structures E3/HSF, E2/HSF et E3/E2/HSF dans le contexte de la SUMOylation nécessite l'examen des complexes formés entre les protéines E3 ligase, E2, les facteurs de transcription HSF, et les protéines cibles soumises à la SUMOylation. Dans ce processus, l'E3 ligase facilite le transfert de SUMO de l'E2 à une protéine cible, comme HSF peut être une cible, modifiant ainsi son activité transcriptionnelle. Le complexe E2/HSF participe directement au transfert de SUMO, et l'E2 joue un rôle central en formant un lien covalent entre SUMO et la protéine cible,

influençant l'interaction de HSF avec d'autres protéines et modulant son activité. La formation du complexe E3/E2/HSF dans la SUMOylation implique une collaboration précise entre l'E3 ligase, l'E2 et le facteur de transcription HSF, orchestrant ainsi l'addition régulée de SUMO à HSF et influençant sa fonction biologique. La forme monomérique, caractérisée par des actions relativement indépendantes, peut conduire à une régulation moins coordonnée de la SUMOylation de HSF, tandis que la forme trimérique, avec la collaboration de E3, E2 et HSF, permet une modulation plus intégrée de ce processus, crucial pour la réponse cellulaire au stress (Kmiecik et al., 2021). Les domaines de liaison spécifiques aux substrats des E3 ligases, les sites d'interaction E3/E2 et les motifs de SUMOylation spécifiques aux HSF déterminent la spécificité des interactions et, par conséquent, les protéines cibles. En résumé, l'analyse des structures E3/HSF, E2/HSF et E3/E2/HSF dans la SUMOylation souligne la coordination essentielle entre ces composants pour réguler efficacement l'addition de SUMO à HSF, modulant ainsi la réponse cellulaire au stress.

La modélisation de la machinerie de SUMOylation avec HSF peut certainement aider à comprendre la SUMOylation préférentielle de certains résidus ou régions (Figure 3.11). HSF est connu pour son rôle dans la régulation du stress thermique et peut également être impliqué dans la régulation de la SUMOylation. En modélisant la machinerie de SUMOylation avec HSF, il est possible de déterminer comment HSF interagit avec les enzymes de la voie de SUMOylation et comment cela pourrait influencer la spécificité de la SUMOylation pour certains résidus ou régions protéiques. Cependant, il est important de noter que la SUMOylation est un processus complexe qui implique plusieurs enzymes et cofacteurs, et la modélisation avec HSF ne capture qu'une partie de cette complexité. D'autres facteurs, tels que la séquence spécifique des résidus cibles et la disponibilité des enzymes de SUMOylation, peuvent également influencer la SUMOylation préférentielle. Par conséquent, bien que la modélisation avec HSF puisse fournir des informations précieuses, elle ne peut pas expliquer entièrement la SUMOylation préférentielle sans prendre en compte d'autres aspects du processus.

Ces connaissances sont fondamentales pour comprendre comment les plantes modulent leur réponse face à des conditions environnementales changeantes, et cette compréhension est renforcée par l'identification de la famille des facteurs de choc thermique humains (HSFs), comprenant au moins sept membres fonctionnels, à savoir HSF1, HSF2, HSF3, HSF4, HSF5,

HSFX et HSFY (Joutsen et Sistonen, 2019; Roos-Mattjus et Sistonen, 2021; Kovacs et al., 2022). Parmi eux, HSF1 et HSF2 sont les membres les plus étudiés et sont clairement associés à la réponse au stress thermique, tandis que HSF4 est principalement impliqué dans des processus liés au développement (Pirkkala et al. 2001; Prahlad and Morimoto 2009). Ces HSFs jouent un rôle central dans la régulation de l'expression de divers gènes, notamment; ceux codant pour les chaperons, influençant ainsi la protéostasie cellulaire. Au-delà de leur impact sur la protéostasie cellulaire, les HSFs participent à une diversité de processus biologiques tels que le vieillissement, le développement, la prolifération cellulaire et la différenciation cellulaire (Kovacs et al., 2022). De plus, ils sont impliqués dans plusieurs pathologies, notamment la neurodégénérescence et le cancer (Kovacs et al., 2022). Il est crucial de noter que les HSFs ont subi une diversification fonctionnelle significative, ce qui signifie que chaque membre de la famille peut avoir des fonctions spécifiques distinctes des autres membres. Par exemple, HSF1 est impliqué dans la régulation de la réponse au stress thermique, la réponse aux protéines mal repliées, la réponse aux dommages à l'ADN et la régulation de l'apoptose (Kovacs et al., 2022). D'un autre côté, HSF2 est associé à la régulation de la prolifération cellulaire, de la différenciation cellulaire (Vihervaara et al. 2013; Xing et al. 2005) et de la spermatogenèse (Wang et al. 2004; Kovacs et al., 2022). Les HSFs humains forment une famille de facteurs de transcription essentiels qui modulent l'expression de nombreux gènes, en particulier ceux liés aux chaperons, et exercent ainsi une influence majeure sur la protéostasie cellulaire. Leur implication dans une variété de processus biologiques et pathologies souligne la complexité de leurs fonctions, chaque membre de la famille ayant des rôles spécifiques qui contribuent à la diversité de leurs impacts biologiques. L'étude comparative des HSFs chez *Arabidopsis* met en lumière de manière significative la complexité des fonctions des HSFs chez les humains, où chaque membre de la famille joue des rôles spécifiques dans divers processus biologiques.

L'analyse comparative des sites de SUMOylation entre les membres de la famille des HSFs chez les humains et chez *Arabidopsis* révèle des similarités et des différences importantes dans leur position, leur abondance et leur fonction. Chez les humains, notamment avec HSF1 et HSF2, les sites de SUMOylation sont généralement localisés sur des résidus spécifiques au sein de la séquence protéique, souvent dans des régions conservées associées à la régulation de la réponse au stress thermique. Ces sites de SUMOylation sont relativement abondants et sont impliqués

dans la modulation de la protéostasie cellulaire, la régulation de l'expression des chaperons et d'autres processus biologiques. En revanche, chez *Arabidopsis*, les données sur les sites de SUMOylation des HSFs sont moins bien caractérisées, mais il est probable qu'ils présentent des caractéristiques similaires en termes de positionnement sur la séquence protéique et de rôle dans la réponse au stress thermique. Cependant, en raison de la diversité fonctionnelle des HSFs chez les plantes, il est possible que certains membres de cette famille présentent des profils de SUMOylation différents, reflétant leurs fonctions spécifiques dans divers processus biologiques. Par conséquent, bien que les HSFs humains et ceux d'*Arabidopsis* partagent des similitudes dans leurs sites de SUMOylation, notamment en ce qui concerne leur position et leur fonction dans la réponse au stress thermique, des différences spécifiques peuvent exister en raison de l'évolution divergente des régulations cellulaires chez les animaux et les plantes. Une analyse approfondie des sites de SUMOylation des HSFs dans les deux systèmes permettrait de mieux comprendre les mécanismes régulateurs impliqués dans la réponse au stress thermique et la conservation fonctionnelle de ces facteurs de transcription à travers les règnes animal et végétal.

En explorant les différences et les similarités entre les HSFs d'*Arabidopsis* et en mettant l'accent sur leurs lysines SUMOylables, des implications importantes se dessinent, tant dans le domaine de l'agronomie que dans la compréhension fondamentale de la biologie des plantes. Les similitudes observées dans les motifs conservés de SUMOylation entre les HSFs d'*Arabidopsis* et d'autres espèces végétales et animales soulignent une conservation évolutive dans la régulation de la réponse au stress thermique. Ces motifs, définis par des résidus lysines spécifiques, témoignent de l'importance fondamentale de la SUMOylation dans la modulation de l'activité des HSFs, assurant ainsi une réponse adaptative aux variations de température. La figure 3.12 présente un alignement entre la protéine HSFB2b de notre organisme modèle et cinq autres cultures majeures au Canada, à savoir le maïs, le soja, le riz, les arachides et le pois. L'objectif de cette comparaison était d'étudier la conservation des séquences à travers l'évolution au sein de ces cultures importantes. Les résultats mettent en évidence une conservation significative des séquences, ce qui témoigne d'une certaine stabilité dans le contexte évolutif.

La relation entre la conservation des résidus et l'évolution des sites de SUMOylation se trouve dans leur trajectoire évolutive. En règle générale, les résidus hautement conservés à travers les espèces sont considérés comme plus anciens sur le plan évolutif, ayant été préservés au fil des

diverses périodes de sélection. En contrepartie, les résidus moins conservés sont souvent perçus comme plus récents sur le plan évolutif, ayant pu apparaître ou se différencier plus récemment au cours de l'évolution. Appliqué aux sites de SUMOylation, ce principe suggère que les sites hautement conservés, où la structure ou la fonction sont maintenues à travers une diversité d'espèces, ont probablement joué un rôle dans des processus biologiques cruciaux sur de longues périodes. Ces sites conservés pourraient être essentiels dans la régulation des fonctions protéiques ou des interactions entre protéines. D'autre part, les sites moins conservés pourraient représenter des changements évolutifs plus récents, apparaissant peut-être en réponse à des pressions sélectives spécifiques dans des lignées ou des conditions environnementales particulières. Ces sites moins conservés pourraient être plus dynamiques et plus sujets à la variation entre les espèces ou même au sein d'une même espèce, ce qui refléterait des adaptations plus récentes ou des changements dans la régulation des fonctions protéiques. L'étude de la conservation des résidus dans les sites de SUMOylation peut offrir des informations sur leur évolution et leur rôle biologique, nous permettant de mieux comprendre comment ces sites ont émergé et se sont adaptés au fil du temps dans différents contextes évolutifs et environnementaux.

CONCLUSION

En conclusion, ce mémoire a apporté des éclairages significatifs sur plusieurs aspects cruciaux de la biologie moléculaire des plantes, en particulier en mettant en lumière l'étude comparative des sites de SUMOylation des facteurs de choc thermique (HSFs) chez *Arabidopsis thaliana* face au stress thermique. L'exploration approfondie des mécanismes moléculaires impliqués dans la réponse au choc thermique a révélé des connexions complexes entre les HSFs et la machinerie de SUMOylation.

Par ailleurs, cette étude fournit des informations sur deux sites potentiels de SUMOylation pour la protéine HSFB2b, à savoir K279 et K356. Une observation importante résultant de l'alignement est la conservation de la lysine à la position 356 chez toutes les cultures étudiées. Cette constance au niveau de la séquence suggère une possible importance fonctionnelle de ce site de SUMOylation dans la protéine HSFB2b, soulignant ainsi une conservation significative à travers les différentes espèces végétales étudiées. Ces constatations renforcent l'idée d'une régulation conservée de la SUMOylation, suggérant des implications fonctionnelles potentielles au sein de ces cultures variées. Cette conservation évolutive met en lumière l'importance cruciale de ces facteurs de transcription dans la physiologie des plantes et dans leur capacité à faire face aux défis environnementaux. Cependant, les différences spécifiques dans les sites de SUMOylation entre les HSFs d'*Arabidopsis* et d'autres espèces végétales reflètent des adaptations évolutives propres à cette plante modèle. Ces variations peuvent influencer la spécificité de la réponse d'*Arabidopsis* au stress thermique dans son environnement naturel. La diversité des sites de SUMOylation peut influencer la capacité des HSFs d'*Arabidopsis* à interagir avec des régions spécifiques de l'ADN, régulant ainsi l'expression de gènes de choc thermique particuliers. Ces différences offrent une opportunité unique pour adapter les stratégies agronomiques visant à améliorer la résilience des cultures à des conditions environnementales changeantes, particulièrement cruciales dans le contexte du changement climatique. La régulation fine de la réponse au stress thermique, orchestrée par les HSFs et modulée par la SUMOylation, est directement liée à la productivité agricole. Comprendre les mécanismes moléculaires qui sous-

tendent la réponse d'*Arabidopsis* au stress thermique peut orienter le développement de cultures plus résilientes face aux conditions climatiques changeantes, un aspect essentiel pour assurer la sécurité alimentaire mondiale. En outre, la connaissance des sites de SUMOylation spécifiques des HSFs d'*Arabidopsis* peut conduire à des applications biotechnologiques novatrices. La manipulation ciblée de ces sites pourrait offrir des moyens de moduler sélectivement l'activité des HSFs, permettant ainsi de renforcer la résistance des cultures à divers stress environnementaux. Cette approche peut être cruciale pour le développement de variétés de plantes adaptées à des conditions de croissance spécifiques, optimisant ainsi les rendements agricoles et contribuant à la durabilité des systèmes alimentaires.

En somme, l'étude comparative des HSFs chez *Arabidopsis* et de leurs sites de SUMOylation élargit notre compréhension des mécanismes de régulation de la réponse au stress thermique dans le règne végétal, révélant des aspects fondamentaux de l'évolution et de l'adaptation des plantes à leur environnement. Ces connaissances sont précieuses non seulement pour la recherche fondamentale, mais offrent également des pistes pour des applications pratiques dans l'agronomie, contribuant ainsi à relever les défis contemporains liés au changement climatique et à la sécurité alimentaire. La recherche sur la SUMOylation offre des perspectives prometteuses pour améliorer la sécurité alimentaire à plusieurs niveaux. En comprenant comment ce processus biologique régule les réponses des plantes aux stress abiotiques, aux pathogènes et aux maladies, nous pourrions développer des cultures plus résistantes et moins vulnérables aux pertes de récolte. De plus, en explorant comment la SUMOylation influence la biosynthèse des métabolites essentiels, nous pourrions augmenter la qualité nutritionnelle des cultures, fournissant ainsi des aliments plus riches en éléments nutritifs. L'identification et la caractérisation des sites de SUMOylation ont offert des perspectives nouvelles sur les mécanismes de régulation post-traductionnelle dans la réponse aux stress environnementaux, avec des implications potentielles pour l'adaptation à long terme des plantes à des environnements changeants. Cette mémoire a significativement enrichi notre compréhension des mécanismes moléculaires sous-tendant la réponse des plantes au stress thermique, ouvrant des perspectives passionnantes pour des recherches futures visant à optimiser les stratégies d'adaptation des plantes aux changements climatiques, avec des implications potentielles pour l'amélioration des cultures et la durabilité agricole.

BIBLIOGRAPHIE

Arnau J., Lauritzen C., E. Petersen G., Pedersen J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, *Protein Expression and Purification*, 48 (1): 1-13, ISSN 1046-5928, <https://doi.org/10.1016/j.pep.2005.12.002>.

Arrigo A-P. (2005). Chaperons moléculaires et repliement des protéines. L'exemple de certaines protéines de choc thermique Heat shock proteins as molecular chaperones. 21, (6) : 619-625. <https://www.erudit.org/fr/revues/ms/2005-v21-n6-7-ms951/011194ar/>

Bouffard D. (2021). Les aléas affectés par les changements climatiques : effets sur la santé, vulnérabilités et mesures d'adaptation. Institut national de santé publique. <https://www.inspq.qc.ca/sites/default/files/publications/2771-aleas-changements-climatiques-effets-sante-vulnerabilite-adaptation.pdf>

Brackett C., Blagg BSJ. (2021). Current status of SUMOylation Inhibitors. *Curr Med Chem*. 28(20): 3892-3912. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8483067/>

Braun L., Cannella D., Pinheiro M. A., Kieffer S., Belrhali H., Garin J., Hakimi M-A. (2009). The small ubiquitin-like modifier (SUMO)- conjugating system of *Toxoplasma gondii*. 39 (1): 81-90. <https://pubmed.ncbi.nlm.nih.gov/18761012/>

Bryant P., Pozzati G., Zhu W., Shenoy A., Kundrotas P., Elofsson. (2022). Predicting the structure of large complexes using AlphaFold and Monte Carlo tree search. *Nature Communication*. 13: 6028. <https://www.nature.com/articles/s41467-022-33729-4>

Castro, P., Couto D., Santos, M., Freitas, A., Lourenço, T., Dias, E., Huguet, S., Silva, J., Tavares, R., Bejarano, E. et Azevedo, H. (2022). SUMO E3 ligase SIZ-1 connects SUMOylation and reactive oxygen species homeostasis processes in *Arabidopsis*. *Plant Physiology*. 189 (2): 934-954. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9157161/>

Chang YY, Liu HC, Liu NY, Chi WT, Wang CN, Chang SH, Wang TT. (2007). A heat-inducible transcription factor, HSFA2, is required for extension of acquired thermotolerance in *Arabidopsis*. *Plant Physiology*. (143): 251–262 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1761974/>

Chen K., Lu J., Liu A. (2000). Chapter 11 – The Activation of Trans-Acting Factors in Response to Hypo- and Hyper-Osmotic Stress in Mammalian Cells. *Cell and Molecular Response to Stress*. (1): 141-155 <https://www.sciencedirect.com/science/article/abs/pii/S1568125400800133>

Cheong M. S., Park H. C., Hong M. J., Lee J., Choi W., Jin J. B., et al. (2009). Specific domain structures control abscisic acid-, salicylic acid-, and stress-mediated SIZ-1 phenotypes. *Plant Physiology*. 151, 1930–1942. 10.1104/pp.109.143719

Cohen-Peer R, Schuster S, Meiri D, Breiman A, Avni A. (2010). Sumoylation of Arabidopsis heat shock factor A2 (HSFA2) modifies its activity during acquired thermotolerance. *Plant Mol Biol.* 74:33–45. <https://pubmed.ncbi.nlm.nih.gov/20521085/>

Couëdel A., Rattalino Edreira J.-I., Pisa Lollato R., Archontoulis S., Sadras V., Grassini P. (2021). Assessing environment types for maize, soybean, and wheat in the United States as determined by spatio-temporal variation in drought and heat stress, *Agricultural and Forest Meteorology.* (307): 108513, ISSN 0168-1923, <https://doi.org/10.1016/j.agrformet.2021.108513>.

Das, Saudamini and R.Vincent J. (2009). Mangroves protected villages and reduced death toll during Indian Super Cyclone. *Proceedings of the National Academic of Science.* 106(18) : 7357–7360. https://www.academie-sciences.fr/pdf/rapport/rads_270617.pdf#page=23

Di Gregorio E. S., Duennwald L. M. (2018). ALS Yeasts Models – Past Success Stories and New Opportunities. *Front Mol Neurosci.* 11: 394. [10.3389/fnmol.2018.00394](https://doi.org/10.3389/fnmol.2018.00394)

Duval D, Duval G, Kedinger C, Poch O, Boeuf H. (2003). The ‘PINIT’ motif, a newly identified conserved domain of the PIAS protein family, is essential for nuclear retention of PIAS3L. *FEBS Lett.* 554: 111–118

Enserink, J.M. (2015). Sumo and the cellular stress response. *Cell Div* 10, 4. <https://doi.org/10.1186/s13008-015-0010-1>

Fan, L., Song, J., Forney, C.F., et Jordan, M.A. (2011). Fruit maturity affects the response of apples to heat stress. *Postharvest Biology and Technology.* 62(1), p. 35-42. doi : 10.1016/j.postharvbio.2011.04.007 <https://profil-profiles.science.gc.ca/fr/publication/fruit-maturity-affects-response-apples-heat-stress-0>

Farhad Md., Kumar U., Tomar V., Bhati P-K., Krishnan N., Mustarin K-E., Barel V., Brestic M., Hossain A. (2023). Heat stress in wheat: a global challenge to feed billions in the current era of the changing climate. *Front. Sustain. Food Syst., Sec. Climate-Smart Food Systems.* Volume 7. <https://doi.org/10.3389/fsufs.2023.1203721>

Fiebelkorn D., Rahman M. (2016). Development of a protocol for frost-tolerance evaluation in rapeseed/canola (*Brassica napus* L.). *The Crop Journal.* 4(2): 147-152. <https://doi.org/10.1016/j.cj.2015.11.004>

Filippopoulou C, Simos G, Chachami G. (2020). The Role of Sumoylation in the Response to Hypoxia: An Overview. *Cells;* 9(11):2359. <https://doi.org/10.3390/cells9112359>

Fortunato S., Lasorella C., Dipierro N., Vita F., Concetta de Pinto M. (2023). Redox Signaling in Plant Heat Stress Response. *Antioxidants (Basel).* 12(3) : 605. doi: [10.3390/antiox12030605](https://doi.org/10.3390/antiox12030605)

Garneau JR., Lima CD. (2010). The SUMO pathway : emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol.* 11(12): 861-871. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3079294/>

Ghosh UK., Islam MN., Siddiqui MN., Khan MAR.(2021). Understanding the roles of osmolytes for acclimatizing plants to changing environment: a review of potential mechanism. *Plant Signal Behav.* 16(8):1913306. doi: 10.1080/15592324.2021.1913306.

Gomez-Pastor R., Burchfiel T. E., Thiele J. D. (2018). Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat Rev Mol Cell Biol.* 19(1): 4-19. doi: [10.1038/nrm.2017.73](https://doi.org/10.1038/nrm.2017.73)

Guerra D., Crosatti C., Khoshro H., Mastrangelo A., Mica E., Mazzucotelli E. (2015). Post-transcriptional and post-translational regulations of drought and heat response in plants : a spider's web of mechanisms. *Front. Plant Sci. Sec Plant Genetics and Genomics.* Volume 6. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2015.00057/full>

Guo M., Liu J-H., Ma X., Luo D-X., Gong Z-H., Lu M-H. (2016). The plant Heat Stress Transcription Factors (HSFs) : Structure, Regulation, and Function in Response to Abiotic Stresses. *Front Plant Sci.* 7: 114. doi: [10.3389/fpls.2016.00114](https://doi.org/10.3389/fpls.2016.00114)

Hasanuzzaman M., Nahar K., Mahabub Alam Md., Roychowdhury R., Fuhita M. (2013). Physiological, Biochemical, and Molecular Mechanisms of Heat Stress Tolerance in Plants. *Int J Mol Sci.* 14(5) : 9643-9684. doi: [10.3390/ijms14059643](https://doi.org/10.3390/ijms14059643)

Heng W. (2015). Global trend of forest ecosystem services valuation: a comprehensive review. *Forest Ecosystems* 2(1): 1-14. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4476178/>

Hirata F., Thibodeau L., Hirata A. (2010). Ubiquitination and SUMOylation of annexin A1 and helicase activity. *Biochim Biophys Acta.* 1800(9): 899-905 <https://pubmed.ncbi.nlm.nih.gov/20359522/>

Hu C., Yang J., Qi Z., Wu H., Wang B., Zou F., Mei H., Liu J., Wang W., Liu Q. (2022). Heat shock proteins: Biological functions, pathological roles, and therapeutic opportunities. *MedComm.* 3(3):e161. doi: 10.1002/mco2.161

Hu S., Ding Y., Zhu C. (2020). Sensitivity and Responses of Chloroplasts to Heat Stress in Plants. *Front. Plant Sci., Sec. Plant Physiology.* Volume 11. <https://doi.org/10.3389/fpls.2020.00375>

Hu Y., Mivechi N.F. (2023). Association and Regulation of Heat Shock Transcription Factor 4b with both Extracellular Signal-Regulated Kinase Mitogen-Activated Protein Kinase and Dual-Specificity Tyrosine Phosphatase DUSP26, *Molecular and Cellular Biology.* 10.1128/MCB.26.8.3282-3294.2006, 26, 8, (3282-3294).

Ishida, T., Yoshimura, M., Miura, K., and Sugimoto, K. (2012). MMS21/HPY2 and SIZ-1, two Arabidopsis SUMO E3 ligases, have distinct functions in development. *PLoS One* 7: e46897. <https://pubmed.ncbi.nlm.nih.gov/23056518/>

Jardin P. et Lejour C. (1991). Les protéines de choc thermique : un modèle moléculaire de réponse des génomes végétaux aux stress abiotiques. pp. 75-90. https://bibliotheque.auf.org/doc_num.php?explnum_id=376

Joutsen J, Sistonen L. (2019). Tailoring of Proteostasis Networks with Heat Shock Factors. *Cold Spring Harb Perspect Biol.* <https://doi.org/10.1101/cshperspect.a034066>

Kam J., Gresshoff P., Shorter R., Xue GP. (2008). The Q-type C2H2 zinc finger subfamily in *Populus tomentosa*: Structure, expression, and association with leaf shape traits. *Frontiers in Plant Science.* 12: 652170.

<https://www.frontiersin.org/articles/10.3389/fpls.2021.652170/full#ref74>

Katiyar-Agarwal S., Gao S., Vivan-Smith A., Jin H. (2007). A novel class of bacteria induced small RNAs in *Arabidopsis*. *Genes & Development* 21(23): 3123-3134.

<https://pubmed.ncbi.nlm.nih.gov/12072434/>

Kmiecik SW., Drzewicka K., Melchior F., Mayer MP. (2021). Heat shock transcription factor 1 is SUMOylated in the activated trimeric state. *J Biol Chem.* 296:100324. doi: 10.1016/j.jbc.2021.100324. Epub 2021 Jan 23. PMID: 33493517; PMCID: PMC7949154.

Kong X., Hong Y., Hsu Y.-F., Huang H., Liu X., Song Z., and Zhu J.-K. (2020). SIZ-1-Mediated SUMOylation of ROS1 Enhances Its Stability and Positively Regulates Active DNA Demethylation in *Arabidopsis*. *Mol. Plant.* 13: 1816–1824.

Kovács, D., Kovács, M., Ahmed, S., Barna J. (2022). Functional diversification of heat shock factors. *Biologia Futura.* 73: 427–439. <https://doi.org/10.1007/s42977-022-00138-z>

Kurepa J., Walker JM., Smalle J., Gosink MM., Davis SJ., Durham TL., Sung DY., Vierstra RD. (2003). The small ubiquitin-like modifier (SUMO) protein modification system in *Arabidopsis*. Accumulation of SUMO1 and-2 conjugates is increased by stress. *J Biol Chem.* 278(9): 6862-72.

<https://pubmed.ncbi.nlm.nih.gov/12482876/>

Lamaoui M., Jemo M., Datla R., Bekkaoui F. (2018). Heat and Drought Stresses in Crops and Approches for Their Mitigation. *Front. Chem., Sec. Crop Biology and Sustainability.* Volume 6. <https://doi.org/10.3389/fchem.2018.00026>

Larkindale J, Vierling E. (2008). Core genome responses involved in acclimation to high temperature. *Plant Physiol.* 146: 748–761

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2245833/>

Lee, JH., Kim, W.T. (2011). Regulation of abiotic stress signal transduction by E3 ubiquitin ligases in *Arabidopsis*. *Mol Cells.* 31 : 201–208. <https://doi.org/10.1007/s10059-011-0031-9>

Lescouzères L. et Bomont P. (2020). E3 Ubiquitin Ligases in Neurological Diseases : Focus on Gigaxonin and Autophagy. *Front Physiol.* 11 :1022

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7642974/>

Li C., Chen Q., Gao X., Qi B., Xu S., Chen J., Wang X. (2005). AtHsfA2 modulates expression of stress responsive genes and enhances tolerance to heat and oxidative stress in *Arabidopsis*.

Sci China C life Sci. 48(6):540-50 <https://pubmed.ncbi.nlm.nih.gov/16483133/>

- Li J., Song Q., Zuo ZF., Liu L. (2022). MicroRNA398-facilitated, RD21-complex-mediated arabinogalactan protein degradation to modulate plant drought tolerance. *Plant Cell and Environment*. 23(18): 10803. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9502370/>
- Ma CS., Zhang W., Peng Y., Zhao F., Chang XQ., Xing K., Zhu L., Ma G., Yang HP., Rudoll V. (2021). Climate warming promotes pesticide resistance through expanding overwintering range of a global pest. *Nat Commun*. 12:5351. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8429752/>
- Masclé HX., Germain-Desprez D., Huynh P., Estéphan P., Aubry M. (2007). Sumoylation of the transcriptional intermediary factor 1beta (TIF1beta), the Co-repressor of the KRAB Multifinger proteins, is required for its transcriptional activity and is modulated by the KRAB domain. *J Biol Chem*. 282(14):10190-202. <https://pubmed.ncbi.nlm.nih.gov/17298944/>
- Mandal S., Ghorai M., Anand U., Samanta D., Kant N., Mishra T., Rahman MH., Jha NK., Jha SK., Lal MK., Tiwari RK., Kumar M., Radha, Prasanth DA., Mane AB., Gopalakrishnan AV., Biswas P., Proćków J., Dey A. (2022). Cytokinin and abiotic stress tolerance -What has been accomplished and the way forward? *Front Genet*. 13:943025. doi: 10.3389/fgene.2022.943025.
- Meiri D, Breiman A. (2009). Arabidopsis ROF1 (FKBP62) modulates thermotolerance by interacting with HSP90.1 and affecting the accumulation of HSFA2-regulated sHSPs. *Plant Journal*. 59: 387–399 <https://pubmed.ncbi.nlm.nih.gov/19366428/>
- Mishra S., Spaccarotella K., Gido J., Samantha I., Chowdary G. (2023). Effects of Heat Stress on Plant-Nutrient Relations : An Update on Nutrient Uptake, Transport, and Assimilation. *Int J Mol Sci*. 24(21): 15670. doi: [10.3390/ijms242115670](https://doi.org/10.3390/ijms242115670)
- Miura, K., Renhu, N., Suzaki, T. (2020). The PHD finger of *Arabidopsis* SIZ-1 recognizes trimethylated histone H3K4 mediating SIZ-1 function and abiotic stress response. *Commun Biol* 3, 23. <https://doi.org/10.1038/s42003-019-0746-2>
- Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A. S., Raghothama, K. G., Baek D., Koo Y., Jin J., Bressan R., Yun DJ., Hasegawa P. (2005). The Arabidopsis SUMO E3 ligase SIZ-1 controls phosphate deficiency responses. *Proc. Natl. Acad. Sci. U. S. A.* 102(21): 7760–7765. <https://pubmed.ncbi.nlm.nih.gov/15894620/>
- Morano K., Thiele D. (1999). Heat Shock Factor Function and Regulation in Response to Cellular Stress, Growth, and Differentiation Signals. *Gene Expr*. 7 (4-5-6): 271-282. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6174667/>
- Nilson E. S., Assmann M. S. (2007). The Control of Transpiration. Insights from Arabidopsis. *Plant Physiol*. 143(1): 19-27. doi: [10.1104/pp.106.093161](https://doi.org/10.1104/pp.106.093161)
- Nishizawa A., Yabuta Y., Yoshida E., Maruta T., Yoshimura K., Shigeoka S. (2006). Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant Journal*. 48: 535-547 <https://pubmed.ncbi.nlm.nih.gov/17059409/>

Nitika, Porter CM., Truman WA., Truttmann MC. (2020). Post-translation modification of Hsp70 family proteins : Expanding the chaperone code. *J Biol. Chem.* 295(31): 10689-10708.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7397107/>

Niu GJ., Xu JD., Yuan WJ., Sun JJ., Yang MC., He ZH., Zhao XF., Wang JX. (2018). Protein Inhibitor of Activated STAT (PIAS) Negatively Regulates the JAK/STAT Pathway by Inhibiting STAT Phosphorylation and Translocation. *Front Immunol.* 9:2392.
doi:10.3389/fimmu.2018.02392

Orlowsky B. Seneviratne S. (2012). Global changes in extreme events : regional and seasonal dimension. *Climatic Change.* 110: 669-696. <https://link.springer.com/article/10.1007/s10584-011-0122-9>

Orosa B., Yates G., Verma V., Srivastava A., Srivastava M., Campanaro A., De Vega D., Fernandes A., Zhang, Lee J., Bennett M., Sasanandom A. (2018). SUMO conjugation to the pattern recognition receptor FLS2 triggers intracellular signaling in plant innate immunity. *Nature Communications* 7: 1-14. <https://pubmed.ncbi.nlm.nih.gov/30518761/>

Palacios AV., Acharya P., Peidi AS., Beck MR., Blanco E., Mishra A., Bawa-Khalfe T, Pakrin SC. (2024). SumoPred-PLM: human SUMOylation and SUMO 2/3 sites Prediction using Pre-tained Protein Language Model. *NAR Genom Bioinform.* 6(1): Iqae011.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10849187/>

Park, H.C., Choi, W., Park, H.J., Cheong MS., Koo YD., Shin G., Chung WK., Kim WY., Kim MG., Bressan RA., Bohnert H., Lee SY., Yun DJ. (2011). Identification and molecular properties of SUMO-binding proteins in Arabidopsis . *Mol Cells.* 32: 143–151.
<https://doi.org/10.1007/s10059-011-2297-3>

Pearce S. R. (2001). Plant Freezing and Damage. *Annals of Botany* 87(4): 417-424
doi:[10.1006/anbo.2000.1352](https://doi.org/10.1006/anbo.2000.1352)

Pirkkala L., Nykänen P., Sistonen L. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB J.* 15:1118–1131.
<https://doi.org/10.1096/f00-0294rev>

Pommier M. (2000). Protéines du choc thermique (HSP70, HSP27) et tumeurs. *Sciences pharmaceutiques.* ffdumas-01261342f <https://dumas.ccsd.cnrs.fr/dumas-01261342/document>

Prahlad V., Morimoto RI. (2009). Integrating the stress response: lessons for neurodegenerative diseases from *C. elegans*. *Trends Cell Biol.* 19:52–61. <https://doi.org/10.1016/j.tcb.2008.11.002>

Qu J., Ma C., Feng J., Xu S., Wang L., Li F., Li Y., Zhang R., Zhang X., Xue J., Guo D. (2016). Transcriptome dynamics during maize endosperm development. *PLoS Genetics.* 11 (10): e1005612. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5047526/>

Roos-Mattjus P., Sistonen L. (2021). Interplay between mammalian heat shock factors 1 and 2 in physiology and pathology. *FEBS J.* <https://doi.org/10.1111/febs.16178>

Roy D., Sadanandom A. (2021). SUMO mediated regulation of transcription factors as a mechanism for transducing environmental cues into cellular signaling in plants. *Cell Mol Life Sci.* 78(6):2641-2664. doi: 10.1007/s00018-020-03723-4. Epub 2021 Jan 16. PMID: 33452901; PMCID: PMC8004507.

Rytz C. T., Feng J., Barros J., Vierstra R. (2023). Arabidopsis-expressing lysine-null SUMO1 reveals a non-essential role for secondary SUMO modification in plants. *Plant Direct.* 7(7):e506. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10350450/>

Rytz C. T., Miller J. N., McLoughlin F., Augustine C. R., Marshall S. R., Yu-ting J., Charng Y-Y., Scalf M., Smith M. L., Vierstra D. R. (2018). SUMOylome Profiling Reveals a Diverse Array of Nuclear Targets Modified by the SUMO Ligase SIZ-1 during Heat Stress, *The Plant Cell.* 30(5): 1077–1099, <https://doi.org/10.1105/tpc.17.00993>

Saracco S., Hansson M., Scalf M., Walker J., Smith L., Vierstra R. (2009). Tandem affinity purification and mass spectrometric analysis of ubiquitylated proteins in Arabidopsis. *Plant J.* 59(2): 344-358. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3639010>

S. Latchman D. (2008). Chapter 8 – Regulation of transcription factor activity, *Eukaryotic Transcription Factors* (Fifth edition), Academic Press, Pages 312-VI, ISBN 9780123739834, <https://doi.org/10.1016/B978-012373983-4.50016-X>.

Sarge K.-D., Shawn P. M., Morimoto P.-M. (1993). Activation of Heat Shock Gene Transcription by Heat Shock Factor 1 Involves Oligomerization, Acquisition of DNA-Binding Activity, and Nuclear Localization and Can Occur in the Absence of Stress, *Molecular and Cellular Biology.* 13(3): 1392-1407, DOI: [10.1128/mcb.13.3.1392-1407.1993](https://doi.org/10.1128/mcb.13.3.1392-1407.1993)

Schramm F., Ganguli A., Kiehlmann E., Englich G., Walch D., von Koskull-Döring P. (2006). The heat stress transcription factor HSFA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in *Arabidopsis*. *Plant Mol Biol.* 60: 759–772 <https://pubmed.ncbi.nlm.nih.gov/16649111/>

Schramm F., Larkindale J., Kiehlmann E., Ganguli A., Englich G., Vierling E., von Koskull-Döring P. (2008). A cascade of transcription factor DREB2A and heat stress transcription factor HSFA3 regulates the heat stress response of Arabidopsis. *Plant Journal.* 53: 264–274 <https://pubmed.ncbi.nlm.nih.gov/17999647/>

Sharma M., Sidhu AK., Samota MK., Gupta M., Koli P., Choudhary M. (2023). Post-Translational Modifications in Histones and Their Role in Abiotic Stress Tolerance in Plants. *Proteomes.* 11(4):38. doi: 10.3390/proteomes11040038

Smoyer-Tomic K.E., Kuhn R., Hudson A. (2003). Heat Wave Hazards : An Overview of Heat Wave Impacts in Canada. *Natural Hazards.* 28: 465-486. <https://doi.org/10.1023/A:1022946528157>

Song Y., Chen Q., Ci D., Shao X., Zhang D. (2014). Effects of high temperature on photosynthesis and related gene expression in poplar. *BMC Plant Biol.* 14: 111. doi: [10.1186/1471-2229-14-111](https://doi.org/10.1186/1471-2229-14-111)

Su S., Zhang Y., Liu P. Curr. (2020). Roles of Ubiquitination and SUMOylation in DNA Response. *Issues Mol. Biol.* 35: 59-84 <https://doi.org/10.21775/cimb.035.059>

Swindell WR., Huebner M., Weber PA. (2007). Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics.* 8: 125. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1887538/>

Szczepanowski RH., Filipek R., Boehrler M. (2005). Crystal Structure of A Fragment of Mouse Ubiquitin-activating Enzyme. *JBC.* 280(23): 22006-22011.

Toma-Fukai S., Shimizu T. (2021). Structural Diversity of Ubiquitin E3 Ligase. *Molecules.* 26(21): 6682. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8586995/>

Tomanov K., Zeschmann A., Hermkes R., Eifler K., Ziba I., Grieco M., Novatchkova M., Hofmann K., Hesse H., Bachmair A. (2014). Arabidopsis PIAL1 and 2 Promote SUMO Chain Formation as E4-Type SUMO Ligases and Are Involved in Stress Responses and Sulfur Metabolism . *The Plant Cell.* 26(11): 4547–4560, <https://doi.org/10.1105/tpc.114.131300>

Vangelisti A., Natali L., Bernardi R., Sbrana C., Turrini A., Hassani-Pak K., Hugues D., Cavallini A., Giovannetti M., Giordani T. (2018). Transcriptome changes induced by arbuscular mycorrhizal fungi in sunflower (*Helianthus annuus* L.) roots. *Scientific Reports.* (11)1: 1-12. <https://www.nature.com/articles/s41598-017-18445-0>

Vihervaara A., Mahat D. B., Guertin M.-J., Chu T., Danko C-G., Lis J-T., Sis-tonen L. (2017). Transcriptional response to stress is pre-wired by promoter and enhancer architecture. *Nat Commun.* 8:255. <https://doi.org/10.1038/s41467-017-00151-0>

Wang G., Ying Z., Jin X., Tu N., Zhang Y., Phillips M., Moskophidis D., Mivechi N-F. (2004). Essential requirement for both HSF1 and HSF2 transcriptional activity in spermatogenesis and male fertility. *Genesis.* 38:66–80. <https://doi.org/10.1002/gene.20005>

Wang H., Feng M., Jiang Y., Du D., Dong C., Zhang Z., Wang W., Liu J., Liu X., Li S., Chen Y., Guo W., Xin M., Yao Y., Ni Z., Sun Q., Peng H., Liu J. (2023). Thermosensitive SUMOylation of TaHsfA1 defines a dynamic ON/OFF molecular switch for the heat stress response in wheat. *Plant Cell Physiol.* 64(10): 3889-3910. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10533334/>

Wang J., Chen L., Long Y., Si W., Cheng B. Jiamg H. (2021). A Novel Heat Shock Transcription Factor (*ZmHsf08*) Negatively Regulates Salt and Drought Stress Responses in Maize. *Int J Mol Sci.* 22(21): 11922. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8584904/>

Wang Y., Dasso M. (2009). SUMOylation and deSUMOylation at a glance. *J Cell Sci.* 122 (23): 4249–4252. doi: <https://doi.org/10.1242/jcs.050542>

Whittam R., Bartley W., Weber G. (1955). The kinetics of the exchange of the phosphate groups of adenosine triphosphate during oxidative phosphorylation. *Biochem J.* 59 (4): 590–599. doi: <https://doi.org/10.1042/bj0590590>

Wong K-A., Kim R., Christofk H., Gao J., Lawson G., Wu H. (2004). Protein inhibitor of activated STAT Y (PIASy) and a splice variant lacking exon6 enhance sumoylation but are not essential for embryosis and adult life. *Mol Cell Biol.* 24: 5577–5586

Song X., Liu G. Duan W., Liu T., Huang Z., Ren J., Li Y., Hou X. (2014). Genome-wide identification, classification, and expression analysis of the C3H/BFL family in cabbage (*Brassica oleracea* var. capitata). *Mol. Genet Genomics.* 289(4): 541-51 <https://pubmed.ncbi.nlm.nih.gov/24609322/>

Xing H., Wilkerson D-C., Mayhew CN., Lubert EJ., Skaggs HS., Goodson ML., Hong Y., Park-Sarge O-K., Sarge KD. (2005). Mechanism of hsp70i gene bookmarking. *Science.* 307:421–423. <https://doi.org/10.1126/science.1106478>

Xiong R., Wang A. (2013). SCE1, the SUMO-Conjugating Enzyme in Plants That Interacts with Nib, the RNA-Dependent RNA Polymerase of *Turnip Mosaic Virus*, Is Required for Viral Infection. *J Virol.* 87(8): 4704-4715. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624346/>

Yang, Q., Zhao, J., Chen, D., Wang Y. (2021). E3 ubiquitin ligases: styles, structures, and functions. *Mol Biomed.* 2(1): 23. <https://doi.org/10.1186/s43556-021-00043-2>

Yau T., Sander W., Eidson C., Courey A. (2021). SUMO Interacting Motifs : Structure and Function. *Cells.* 10 (11): 2825 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8616421/>

Yoo CY., Miura K., Jin JB., Lee J., Park HC., Salt DE., Yun DJ., Bressan RA., Hasegawa PM. (2006). SIZ-1 small ubiquitin-like modifier E3 ligase facilitates basal thermotolerance in Arabidopsis independent of salicylic acid. *Plant Physiol.* 142(4):1548-58. doi: 10.1104/pp.106.088831.

Yoshida T., Sakuma Y., Todaka D., Maruyama K., Qin F., Mizoi J., Kidokoro S., Fujita Y., Shinozaki K., Yamaguchi-Shinozaki K. (2008). Functional analysis of an Arabidopsis heat-shock transcription factor HSFA3 in the transcriptional cascade downstream of the DREB2A stress-regulatory system. *Biochem Biophys Res Commun.* 368: 515–521 <https://pubmed.ncbi.nlm.nih.gov/18261981/>

Zhang Y., Chou SD., Murshif A., Prince T., Stevenson MA., Calderwood S.K. (2011). The Role Of Heat Shock Factors In stress-induced Transcription. *Methods Mol. Biol.* 787: 21-32. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4088327/>

Zhang S., Wang S., Jinlian L., Liu Z., Wang Y., Ma N., Meng Q. (2018). SUMO E3 Ligase SIZ-1 Facilitates Heat Tolerance in Tomato, *Plant and Cell Physiology.* 59(1): 58–71, <https://doi.org/10.1093/pcp/pcx160>

Zhao J., Lu Z., Wang L., Jin B. (2021). Plant Responses to Heat Stress : Physiology, Transcription, Noncoding RNAs, and Epigenetics. *Int J Mol Sci.* 22(1):117. doi: [10.3390/ijms22010117](https://doi.org/10.3390/ijms22010117)

Zhao P., Javed S., Shi X., Wu B., Zhang D., Xu S., Wang X. (2020). Varying Architecture of Heat Shock Elements Contributes to Distinct Magnitudes of Target Gene Expression and Diverged Biological Pathways in Heat Response of Bread Wheat *Front. Genet., Evolutionary and Population Genetics.* Volume 11. <https://doi.org/10.3389/fgene.2020.00030>

Zhao Q., Ma Y., Li Z., Zhang K., Zheng M., Zhang S. (2020). The Function of SUMOylation and Its Role in the Development of Cancer Cells under Stress Conditions : A Systematic Review. *Stems Cells Int.* 8835714. doi: [10.1155/2020/8835714](https://doi.org/10.1155/2020/8835714)

Zhao X., Blobel G. (2005). A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proceedings of the National Academy of Sciences.* 102(13): 4777-4782. <https://doi.org/10.1073/pnas.0500537102>

Zheng Y., Schumaker K., Guo Y. (2012). Sumoylation of the transcription factor AtMYB30 by small ubiquitin-like modifier E3 ligase SIZ-1 mediates abscisic acid response in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences.* 109 (31):12822-7. <https://pubmed.ncbi.nlm.nih.gov/22814374/>