

## **Chapitre 2 : Analyses des sHsps nucléaires chez les insectes**

## 2.1 Introduction

Le noyau est l'organelle cellulaire qui distingue les eucaryotes des procaryotes. Il est séparé du cytoplasme par une double membrane, la membrane extérieure étant en continuité avec le réticulum endoplasmique. Une grande partie des activités biologiques se passe au noyau, entre autres la réplication et la réparation des dommages de l'ADN, la biogenèse des ribosomes, la transcription, l'épissage alternatif, les modifications de l'ARN et de l'ADN. La translocation des protéines en continu entre le noyau et le cytoplasme est connue. Ce transport nucléocytoplasmique se fait via une famille de facteurs bien conservés à travers les espèces : les karyophérines  $\beta$ . Celles-ci peuvent participer à l'import nucléaire via les importines, ou à l'export nucléaire par les exportines.

Les petites protéines de stress/Small heat shock proteins (sHsps) sont des chaperons moléculaires ATP-indépendants ubiquitaires retrouvées chez tous les règnes du vivant (Maaroufi and Tanguay, 2013, Kriehuber et al., 2010). Pour les sHsps cytosoliques, la relocalisation dans le noyau est observée dans certaines conditions de stress chez pratiquement tous les mammifères (Arrigo and Ahmad-Zadeh, 1981, Borrelli et al., 2002, van de Klundert and de Jong, 1999), mais ce comportement est distinct des sHsps nucléaires qui sont présentes tout le temps dans le noyau chez les plantes et *Drosophila melanogaster*, grâce à un signal de ciblage nucléaire (NLS : Nuclear Localization Signal) (Siddique et al., 2008, Siddique et al., 2003, Scharf et al., 2001, Michaud et al., 2008). La différence entre la localisation nucléaire des sHsps chez les mammifères, plantes et *D.melanogaster* suggère qu'elles s'adaptent à leurs environnements et à leurs besoins cellulaires spécifiques. Alors qu'on connaît l'existence de plusieurs membres de sHsps avec un NLS chez les plantes, le seul membre décrit chez les insectes jusqu'à aujourd'hui reste DmHsp27. Le but de ce chapitre est de répondre à la question s'il existe d'autres membres de sHsps avec une localisation nucléaire chez les insectes et d'étudier par la suite la particularité de cette classe de sHsps.

## **2.2 Matériel et méthodes**

### **2.2.1 Extraction des séquences de petites protéines de choc thermique nucléaire chez les insectes**

Toutes les séquences de protéines d'insectes contenant le domaine alpha-cristallin (ACD numéro d'accession PF00011) sont extraites avec les outils HMMER (<http://hmmer.org/>) et pfam (<http://pfam.xfam.org/family/HSP20#tabview=tab1>) (Finn et al., 2011, Finn et al., 2014). Les protéines contenant le terme "fragment" dans leur description sont enlevées. Par la suite les séquences potentiellement nucléaires sont extraites suite à un alignement avec DmHsp27 pour déterminer la présence ou non du NLS.

### **2.2.2 Alignement et LOGOS**

Les séquences sont alignées avec les paramètres par défaut MUSCLE (McWilliam et al., 2013). Les structures secondaires de l'alignement généré par ESPript (Robert and Gouet, 2014), selon la structure prédite du domaine ACD de DmHsp27 (Moutaoufik et al., 2016). Les numéros d'accession UniProt des séquences utilisées dans les alignements sont indiqués. Pour identifier les régions conservées, j'ai utilisé le programme WebLogo 3 (Crooks et al., 2004) pour créer des blocs (LOGOS) de résidus conservés à partir de l'alignement des séquences multiples de protéines .

### **2.2.3 Phylogénétique moléculaire**

Les arbres phylogénétiques ont été construits avec MEGA7 (Tamura et al., 2013) en utilisant des alignements multiples de séquences. Les relations phylogénétiques sont obtenues par maximum de vraisemblance en utilisant le modèle de substitution WAG (Whelan and Goldman, 2001) et 1000 répliques bootstrap.

## 2.3 Résultats et discussion

La détermination de la localisation cellulaire des protéines est importante pour la compréhension des fonctions biologiques. Dans cette étude, je me suis intéressé à la localisation nucléaire des petites protéines de choc thermique chez les insectes. DmHsp27 de *D.melanogaster* a été le premier sHsp nucléaire étudiée (Beaulieu et al., 1989b). Depuis cette découverte d'autres sHsps en particulier chez les plantes montrent une localisation nucléaire tels que *Solanum peruvianum* LpHsp16.1, *Arabidopsis thaliana* AtHsp17.4, *Gossypium arboreum* GaHsp17.3, *Medicago truncatula* MtHsp17.1, *Lactuca sativa* LsHsp17.8, *Triticum aestivum* TaHsp19.0, *Hordeum vulgare* HvHsp19.1, et *Oryza sativa* OsHsp18.6 (Siddique et al., 2008, Siddique et al., 2003, Scharf et al., 2001).

L'importation de protéines dans le noyau dépend d'un signal de ciblage nucléaire (NLS) et les membres de la superfamille du récepteur de transport nucléaire (importine). Chez les plantes la localisation nucléaire dépend de la voie classique appelée (NLS riche en lysine) grâce à une courte séquence (N(G)KRKR) entre les brins  $\beta$  5 et 6 dans le domaine ACD conservée (Siddique et al., 2003). Ce mode nécessite l'importine  $\alpha$ , pour servir de pont entre le NLS et le récepteur d'importine  $\beta$  (Gorlich, 1997, Nigg, 1997, Ohno et al., 1998, Palmeri and Malim, 1999). Chez DmHsp27, le motif NLS (appelée riche en arginine) est composé de trois arginines (54-55-56) dans la région N-terminale non conservée (Michaud et al., 2008). Le comportement différent des simples, doubles et triples mutants suggère que ces arginines forment deux groupes fonctionnels. Un groupe contient les arginines 54 et 55 alors que l'autre groupe est formé par l'arginine 56. Si au moins un groupe est actif (pas de mutation d'arginine) la localisation est nucléaire. L'inactivation des deux groupes par mutation conduit à une localisation cytoplasmique de DmHsp27. Le NLS riche en arginine utilise l'importine  $\beta$  directement pour l'importation par un mécanisme indépendant de l'importine  $\alpha$  (Palmeri and Malim, 1999).

L'intérêt d'étudier la localisation nucléaire de sHsps chez les insectes et particulièrement la drosophile, se base sur le fait que la région NTR a été rapportée comme une région désordonnée non conservée chez les sHsps (Poulain et al., 2010, McHaourab et al., 2009). Par conséquent, la localisation d'un NLS dans cette région est surprenante. Dans ce travail, j'ai identifié et caractérisé les sHsps nucléaires chez les

insectes afin de définir la particularité de ces protéines.

### **2.3.1 Les sHsps nucléaires chez les insectes**

Il est admis que la présence du domaine ACD conservée est un critère suffisant pour l'attribution d'une séquence à la famille de sHsp (Caspers et al., 1995). J'ai obtenu 628 séquences contenant ACD chez les insectes. Seulement 28 séquences montrent un NLS riche en arginine dans la partie NTR (Tableau 2.1). La potentielle localisation nucléaire des séquences obtenues est confirmée par deux outils : ProtComp v. 9.0 (<http://www.softberry.com/berry.phtml?topic=protcomppl&group=help&subgroup=proloc>) basée sur une comparaison directe avec une base de protéines homologues ayant une localisation connue et Euk-mPLOC (Chou and Shen, 2007, Chou and Shen, 2008) un outil puissant développé en hybridant les informations des domaines fonctionnels, et de l'information de l'évolution séquentielle.

Nom	Uniprot ID	Espèce	Poids moléculaire théorique	Localisation de NLS
Hsp 27	P02518	<i>Drosophila melanogaster</i>	23616.6	Arg 54,55,56
l(2)efl	A0A0J7K2I6	<i>Lasius niger</i>	3398.42	Arg 53-54
l(2)efl	A0A0N0BG57	<i>Melipona quadrifasciata</i>	26244.03	Arg 52-53
ND	K7IXH9	<i>Nasonia vitripennis</i>	22277.49	Arg 50-52
Hsp27	B3GK93	<i>Ceratitis capitata</i>	23944.24	Arg 54-55-56
Hsp27	A0A0K8UP97	<i>Bactrocera latifrons</i>	24037.31	Arg 53-54-55
Hsp20	I1T1H1	<i>Bactrocera dorsalis</i>	23831.06	Arg 51-52-53
Hsp27	A0A0A1XCZ4	<i>Bactrocera cucurbitae</i>	23920.18	Arg 53-54-55
Hsp20	T1PAE8	<i>Musca domestica</i>	24364.49	Arg 52-53
ND	Q16S79	<i>Aedes aegypti</i>	21912.46	Arg 57-58
Hsp25	A1E385	<i>Sarcophaga crassipalpis</i>	24982.00	Arg 60-61
Hsp27	A0A0L0CSL4	<i>Lucilia cuprina</i>	23330.55	Arg 57-58-58
Hsp27	F5B962	<i>Drosophila repletoides</i>	24503.6	Arg 48,49,50
Hsp27	A0A0M4EGE5	<i>Drosophila busckii</i>	24444.62	Arg 51-52-52
GI13087	B4L112	<i>Drosophila mojavensis</i>	23459.5	Arg 51,52,53
GJ13835	B4LGS9	<i>Drosophila virilis</i>	23791.9	Arg 51,52,53
Hsp27	A9UEZ2	<i>Drosophila buzzatii</i>	23694.7	Arg 51,52,53
Hsp27	F5B960	<i>Drosophila albomicans</i>	24770	Arg 52,53,54

Hsp27	F5B961	<i>Drosophila sulfurigaster albostrigata</i>	24674.9	Arg 52,53,54
GE20833	B4PEX9	<i>Drosophila yakuba</i>	23593.6	Arg 54,55,56
GM25143	B4HKP4	<i>Drosophila sechellia</i>	23624.6	Arg 54,55,56
GG15371	B3NCF2	<i>Drosophila erecta</i>	23649.6	Arg 54,55,56
GD14177	B4QN53	<i>Drosophila simulans</i>	23565.5	Arg 54,55,56
GH17009	B4IYY1	<i>Drosophila grimshawi</i>	24441.5	Arg 54,55,56
GF10692	B3M6F4	<i>Drosophila ananassae</i>	24489.6	Arg 57,58,59
GA18205	Q29F98	<i>Drosophila pseudoobscura pseudoobscura</i>	25041.1	Arg 58,59,60
GK17637	B4MN48	<i>Drosophila willistoni</i>	25284.5	Arg 58,59,60
GL22445	B4H1L8	<i>Drosophila persimilis</i>	25041.1	Arg 58,59,60

**Tableau 2-1 : Les sHsps nucléaires chez les insectes.**

### **2.3.2 Analyse des sHsps nucléaires chez les insectes**

L'analyse par alignement des séquences a révélé que les motifs caractéristiques de sHsps sont conservés. Tel est le cas de l'arginine associée à des pathologies humaines dans le brin  $\beta$  6+7 (figure 2.1 : étoile noire) (Vicart et al., 1998, Wang et al., 2001) et le motif conservé CAM (C-terminal anchoring motif), impliqué dans l'inter-dimère (de Jong et al., 1998, van Montfort et al., 2001b, Stamler et al., 2005).



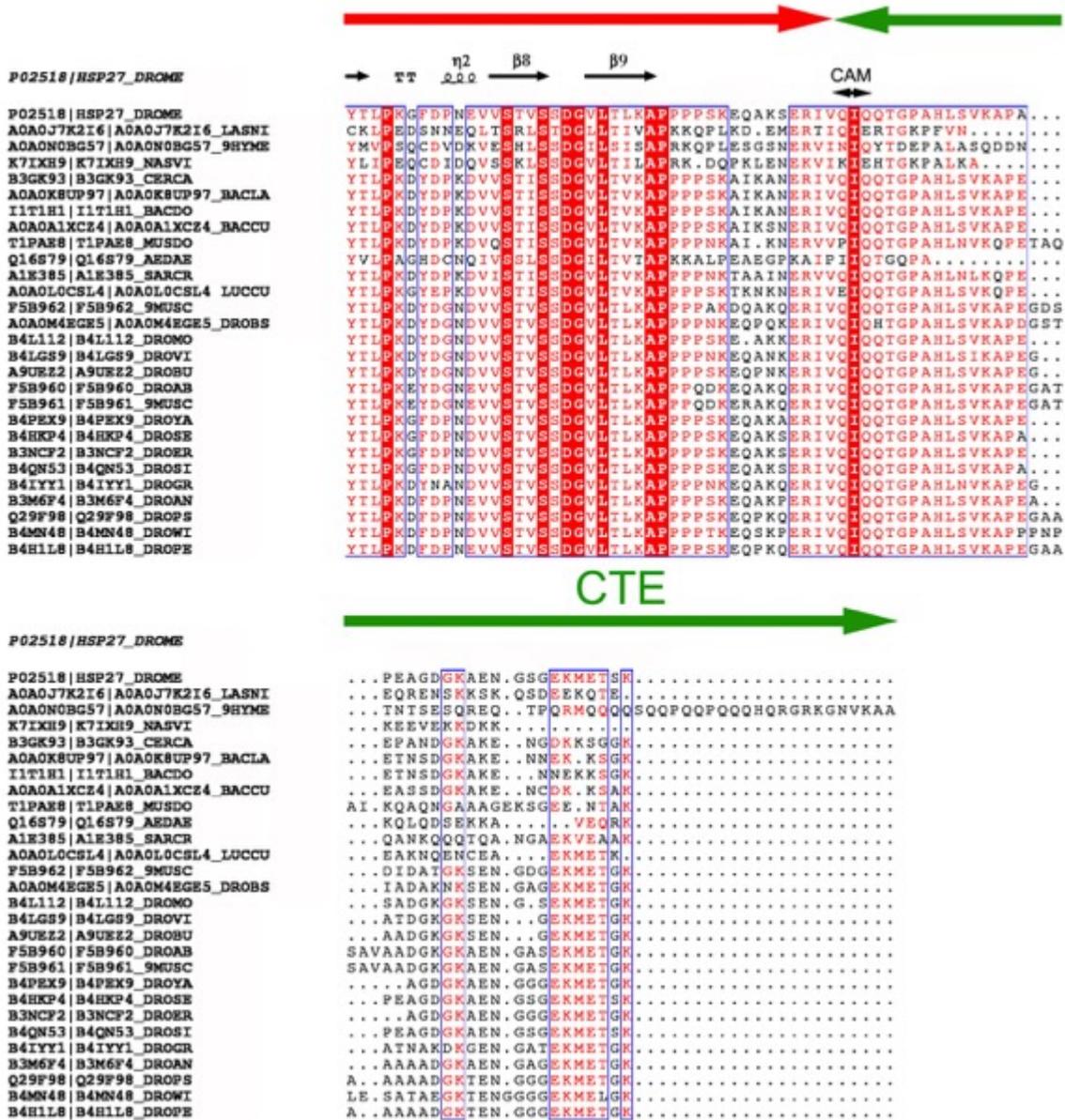
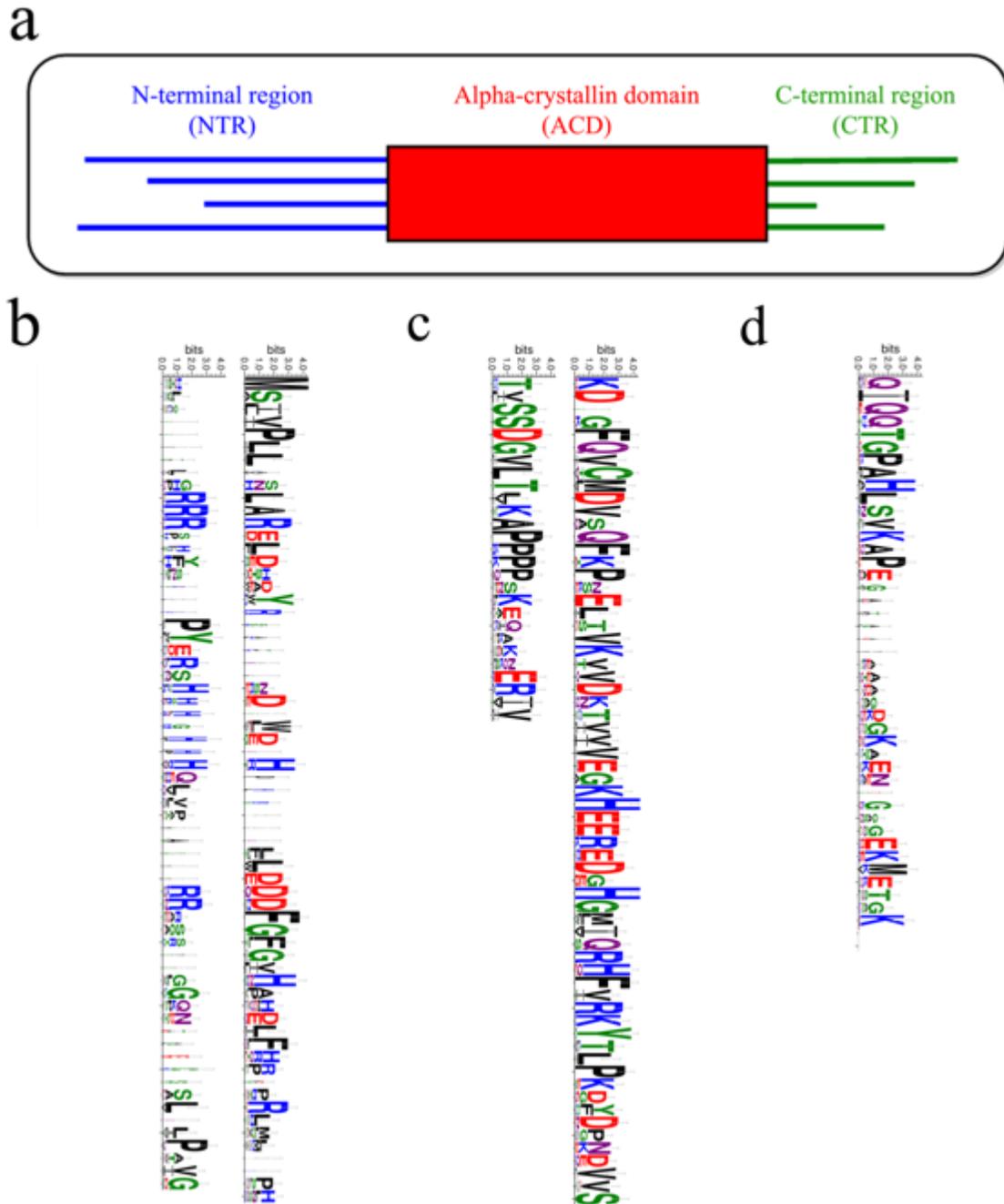


Figure 2.1 : Alignement de séquence des petites protéines de stress chez les insectes.

L'alignement est généré en utilisant MUSCLE. Les numéros d'accèsion UniProt des séquences utilisées dans les alignements sont indiqués. La structure secondaire du domaine ACD prédit de DmHsp27 (Moutaoufik et al., 2016) est indiquée en haut. La région N-terminale montre la conservation des arginines composant le NLS. Le CAM I/V-X-I/L/V dans l'extension C -terminale et l'arginine conservée dans le brin Beta 6+7 est indiquée par une étoile noire.

En général les sHsps se caractérisent par la présence d'un domaine central alpha-cristallin ACD conservé entouré d'une région N-terminale (NTR) et d'une extension C-terminale (CTE) (Figure 2.2 a). L'analyse de la région NTR montre une longueur moyenne de 88.5 ( $\pm 3.4$ ) avec surreprésentation de phénylalanine, leucine, histidine et arginine et absence de cystéine et lysine dans la majorité des séquences (Figure 2.2 b). Il est important de souligner que les régions riches en phénylalanine dans la région N-terminale, ont été proposées pour être impliquées dans la liaison de substrat (Haslbeck et al., 2004b, Van Montfort et al., 2001a).

Le domaine ACD présente une longueur moyenne de 91.9 ( $\pm 0.25$ ), tel que rapporté dans la littérature ; la longueur de l'ACD est plus constante que les deux autres régions. En revanche, dans le domaine ACD un enrichissement significatif des résidus chargés positivement et négativement est observé, alors que les résidus polaires et aromatiques sont sous-représentés et il y a absence de tryptophane dans la majorité des séquences (Figure 2.2 c). Toutes les sHsps nucléaires étudiées présentent une cystéine dans le domaine ACD suggérant que dans des conditions oxydantes des ponts disulfures entre deux ACD peuvent être impliqués dans l'interface du dimère (Baranova et al., 2011). L'extension CTE, la région la plus courte montre une longueur moyenne de 37.5 ( $\pm 2.9$ ) avec une sous-représentation des acides aminés aromatiques et de cystéine (Figure 4.2d). Ces données sont en accord avec une étude précédente par (Kriehuber et al., 2010) et cela montre une conservation des caractéristiques générales des sHsps.



**Figure 2.2 : Analyses des petites protéines de stress nucléaires chez les insectes.**

a- Architecture des petites protéines de stress (sHsps). En général les sHsps contiennent un domaine alpha-cristallin (ACD, en rouge) entouré d'une région N-terminale non conservée (NTR, en bleu) et une extension C-terminale (CTE, en vert). Présentation logo de la composition en amino-acide de NTR (b), ACD (c) et CTE (d). La hauteur de chaque lettre est proportionnelle à sa fréquence. Les acides aminés sont colorés en fonction de

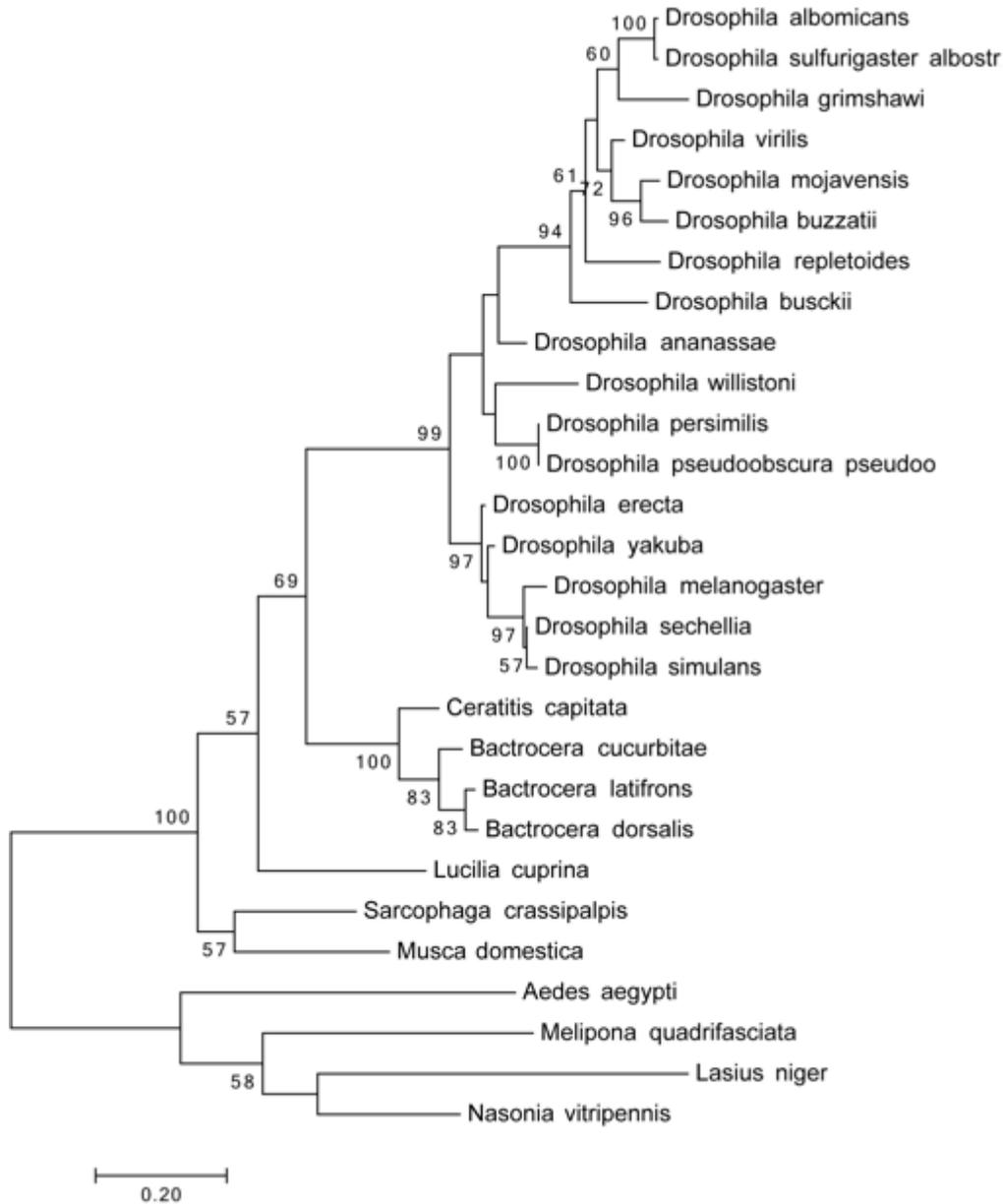
leurs propriétés chimiques : acides (D et E) en rouge, basique (K, R , H) en bleu, polaire (G, S, T, Y, C ) en vert et (N , Q) en violet, hydrophobiques (A , V , L , I, P , W , F , M ) en noir.

### 2.3.3 Evolution des sHsps nucléaires chez les insectes

Les séquences obtenues appartiennent à trois groupes :

- Le premier groupe contient les espèces : *Lasius niger*, *Melipona quadrifasciata*, et *Nasonia vitripennis*. Toutes ces espèces ont en commun le même ordre Hymenoptera et sous ordre Apocrita.
- Le deuxième groupe regroupe : *Ceratitis capitata*, *Bactrocera latifrons*, *Bactrocera dorsalis*, *Bactrocera cucurbitae*, *Musca domestica*, *Sarcophaga crassipalpis*, *Lucilia cuprina*, *Drosophila melanogaster*, *Drosophila mojavensis*, *Drosophila repletoides*, *Drosophila buzzatii*, *Drosophila albomicans*, *Drosophila sulfurigaster albostrigata*, *Drosophila yakuba*, *Drosophila sechellia*, *Drosophila virilis* et *Drosophila erecta*. Ces espèces partagent le même ordre Diptera et le sous-ordre Brachycera.
- Le dernier groupe contient seulement *Aedes aegypti* qui appartient à l'ordre Hymenoptera et sous-ordre Nematocera.

L'arbre phylogénétique obtenu (Figure 2.3) montre qu'avec les sHsps nucléaires, on est capable d'avoir une séparation des différents groupes et même à l'intérieur du même genre comme le cas du genre *Drosophila*, ce qui signifie que les sHsps sont des bons marqueurs pour suivre l'histoire évolutive.



**Figure 2.3 : Analyse phylogénétique des sHsps nucléaires d’insectes par la méthode de maximum de vraisemblance.**

L'histoire évolutive est déduite en utilisant la méthode du maximum de vraisemblance sur la base du modèle Whelan et Goldman (Whelan and Goldman, 2001) avec le logiciel MEGA7 (Tamura et al., 2013).

### 2.3.4 Les fonctions de DmHsp27 dans le noyau

Les fonctions de DmHsp27 dans le noyau restent inconnues. Une étude avait rapporté une association de DmHsp22, DmHsp23, DmHsp26 et DmHsp27 aux hnRNP (heterogeneous nuclear ribonucleoprotein) (Kloetzel and Bautz, 1983). En plus, des sHsps cytoplasmiques et DmHsp27 de drosophile sont co-purifiées avec le prosome lors de choc thermique et quatre petits ARN ont été séquencés de ces complexes prosomes-sHsps (Arrigo et al., 1985). L'alpha-cristallin de souris est capable de lier les gènes D/E/F-cristallin (Pietrowski et al., 1994) et l'alpha-cristallin bovin reconnaît le simple et double brin d'ADN (Bloemendal et al., 1999, Singh et al., 1998). Les sHsps peuvent aussi protéger les ARN messenger (ARNm) lors du stress en induisant l'arrêt de la traduction (Kedersha et al., 1999, Scharf et al., 1998). Il n'est pas encore clair si cette interaction se fait directement ou via des intermédiaires.

La recherche d'interactions de DmHsp27 met en lumière certaines corrélations potentiellement intéressantes pour sa fonction. L'ensemble des données disponibles dans les bases de données DroID (Drosophila Interaction Database, (Yu et al., 2008), BioGRID3.4 (Chatr-Aryamontri et al., 2017) et FlyBase (Marygold et al., 2016) sont groupées dans le tableau 2.2. Outre l'interaction attendue de DmHsp27 avec les facteurs de maturation et de liaison à l'ARN (TH1 et thoc7), DroID met en évidence l'interaction potentielle entre DmHsp27 et une protéine de liaison à d'ADN (Alh) (Giot et al., 2003), en plus d'une protéine impliquée dans l'initiation de la réplication d'ADN Mcm3 (Guruharsha et al., 2011). De façon intéressante, DmHsp27 pourrait interagir avec des kinases, phosphatases, facteurs de transcription ou régulation des histones (Tableau 2.2). Ces interactions nous orientent vers un rôle de DmHsp27 dans les processus nucléaires de base tels que la réplication d'ADN, transcription et maturation de l'ARN en plus de sa fonction de chaperon moléculaire.

Protéine partenaire de DmHSP27WT	Fonctions	Références
Transcriptional Adaptor 2b (Ada2b)	Régulation de l'acétylation/méthylation des histones	(Giot et al., 2003)
Alhambra (Alh)	Liaison d'ADN.	(Giot et al., 2003)
Arginine methyltransferase 2 (Art2)	Régulation de l'acétylation/méthylation des histones	(Guruharsha et al., 2011)
Burgundy (bur)	Activité GMP synthase	(Giot et al., 2003)
Beadex (Bx)	Protéine à doigt de zinc	(Guruharsha et al., 2011)
Cyclin-dependent kinase 1 (Cdk1)	Protéine kinase	(Guruharsha et al., 2011)
CG18004	ND	(Guruharsha et al., 2011)
CG10984	ND	(Guruharsha et al., 2011)
E(Z)	Régulation de méthylation des histones	(Kang et al., 2015)
Xport-A	ND	(Rosenbaum et al., 2011)
Histone deacetylase 3 (HDAC3)	Régulation de l'acétylation/méthylation des histones	(Guruharsha et al., 2011)
Kayak (kay)	Facteur de transcription	(Guruharsha et al., 2011)
Lisencephaly-1 (Lis-1)	Activité ATPase	(Giot et al., 2003)
Lk6 kinase (Lk6)	Protéine kinase	(Giot et al., 2003)
LWR	Enzyme de conjugaison SUMO	(Joanisse et al., 1998)
MED19	Régulation transcription	(Giot et al., 2003)
Minichromosome maintenance 3 (Mcm3)	Réplication d'ADN	(Guruharsha et al., 2011)
Mediator complex subunit 19 (MED19)	Facteur de transcription	(Giot et al., 2003)
microtubule star (mts)	phosphatase	(Guruharsha et al., 2011)
Nemo (nmo)	Protéine kinase	(Guruharsha et al., 2011)
ovo	Facteur de transcription	(Guruharsha et al., 2011)
P53	Facteur de transcription	(Guruharsha et al., 2011)
RPT1	Activité ATPase	(Guruharsha et al., 2011)
Slbo	Activation de la transcription	(Rorth et al., 1998)

Spt7	Régulation de l'acétylation/méthylation des histones	(Guruharsha et al., 2011)
Su(var)3-9	Régulation de la méthylation des histones	(Giot et al., 2003)
TH1	Liaison d'ARNm	(Guruharsha et al., 2011)
thoc7	Maturation d'ARNm	(Guruharsha et al., 2011)

**Tableau 2-2 : Interactions de DmHsp27.**

## 2.4 Conclusion

Dans ce chapitre j'ai montré que le signal NLS riche en arginine n'est pas une particularité de *D.melanogaster*. La progression des projets de séquençage de génomes chez d'autres insectes pourra confirmer la présence d'autres NLS (du même type ou non) dans d'autres espèces. Le rôle exact de la localisation des sHsps dans le noyau reste inconnu. Le réseau d'interaction de DmHsp27 nous laisse croire que cette protéine ne joue pas seulement le rôle de chaperon moléculaire, mais qu'elle est probablement impliquée dans différents processus nucléaires. Cette étude nous ouvre les perspectives d'étudier l'oligomérisation de DmHsp27WT et les mutants au niveau du domaine ACD conservé (chapitre 3). En plus, une analyse plus approfondie des séquences obtenues dans ce chapitre a permis de choisir les sites de mutations et d'étudier l'importance et la flexibilité de la région N-terminale (chapitre 4).

## Références

- ARRIGO, A. P. & AHMAD-ZADEH, C. 1981. Immunofluorescence localization of a small heat shock protein (hsp 23) in salivary gland cells of *Drosophila melanogaster*. *Mol Gen Genet*, 184, 73-9.
- ARRIGO, A. P., DARLIX, J. L., KHANDJIAN, E. W., SIMON, M. & SPAHR, P. F. 1985. Characterization of the prosome from *Drosophila* and its similarity to the cytoplasmic structures formed by the low molecular weight heat-shock proteins. *EMBO J*, 4, 399-406.
- BARANOVA, E. V., WEEKS, S. D., BEELEN, S., BUKACH, O. V., GUSEV, N. B. & STRELKOV, S. V. 2011. Three-dimensional structure of alpha-crystallin domain dimers of human small heat shock proteins HSPB1 and HSPB6. *J Mol Biol*, 411, 110-22.
- BEAULIEU, J. F., ARRIGO, A. P. & TANGUAY, R. M. 1989. Interaction of *Drosophila* 27000 Mr Heat-Shock Protein with the Nucleus of Heat-Shocked and Ecdysone-Stimulated Culture Cells. *Journal of Cell Science*, 92, 29-36.
- BLOEMENDAL, M., TOUMADJE, A. & JOHNSON, W. C., JR. 1999. Bovine lens crystallins do contain helical structure: a circular dichroism study. *Biochim Biophys Acta*, 1432, 234-8.
- BORRELLI, M. J., BERNOCK, L. J., LANDRY, J., SPITZ, D. R., WEBER, L. A., HICKEY, E., FREEMAN, M. L. & CORRY, P. M. 2002. Stress protection by a fluorescent Hsp27 chimera that is independent of nuclear translocation or multimeric dissociation. *Cell Stress Chaperones*, 7, 281-96.
- CASPERS, G. J., LEUNISSEN, J. A. & DE JONG, W. W. 1995. The expanding small heat-shock protein family, and structure predictions of the conserved "alpha-crystallin domain". *J Mol Evol*, 40, 238-48.
- CHOU, K. C. & SHEN, H. B. 2007. Euk-mPLOC: a fusion classifier for large-scale eukaryotic protein subcellular location prediction by incorporating multiple sites. *J Proteome Res*, 6, 1728-34.
- CHOU, K. C. & SHEN, H. B. 2008. Cell-PLOC: a package of Web servers for predicting subcellular localization of proteins in various organisms. *Nat Protoc*, 3, 153-62.
- CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Res*, 14, 1188-90.
- DE JONG, W. W., CASPERS, G. J. & LEUNISSEN, J. A. 1998. Genealogy of the alpha-crystallin--small heat-shock protein superfamily. *Int J Biol Macromol*, 22, 151-62.
- FINN, R. D., BATEMAN, A., CLEMENTS, J., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., HEGER, A., HETHERINGTON, K., HOLM, L., MISTRY, J., SONNHAMMER, E. L., TATE, J. & PUNTA, M. 2014. Pfam: the protein families database. *Nucleic Acids Res*, 42, D222-30.
- FINN, R. D., CLEMENTS, J. & EDDY, S. R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*, 39, W29-37.

- GIOT, L., BADER, J. S., BROUWER, C., CHAUDHURI, A., KUANG, B., LI, Y., HAO, Y. L., OOI, C. E., GODWIN, B., VITOLS, E., VIJAYADAMODAR, G., POCHART, P., MACHINENI, H., WELSH, M., KONG, Y., ZERHUSEN, B., MALCOLM, R., VARRONE, Z., COLLIS, A., MINTO, M., BURGESS, S., MCDANIEL, L., STIMPSON, E., SPRIGGS, F., WILLIAMS, J., NEURATH, K., IOIME, N., AGEE, M., VOSS, E., FURTAK, K., RENZULLI, R., AANENSEN, N., CARROLLA, S., BICKELHAUPT, E., LAZOVATSKY, Y., DASILVA, A., ZHONG, J., STANYON, C. A., FINLEY, R. L., JR., WHITE, K. P., BRAVERMAN, M., JARVIE, T., GOLD, S., LEACH, M., KNIGHT, J., SHIMKETS, R. A., MCKENNA, M. P., CHANT, J. & ROTHBERG, J. M. 2003. A protein interaction map of *Drosophila melanogaster*. *Science*, 302, 1727-36.
- GORLICH, D. 1997. Nuclear protein import. *Curr Opin Cell Biol*, 9, 412-9.
- GURUHARSHA, K. G., RUAL, J. F., ZHAI, B., MINTSERIS, J., VAIDYA, P., VAIDYA, N., BEEKMAN, C., WONG, C., RHEE, D. Y., CENAJ, O., MCKILLIP, E., SHAH, S., STAPLETON, M., WAN, K. H., YU, C., PARSA, B., CARLSON, J. W., CHEN, X., KAPADIA, B., VIJAYRAGHAVAN, K., GYGI, S. P., CELNIKER, S. E., OBAR, R. A. & ARTAVANIS-TSAKONAS, S. 2011. A protein complex network of *Drosophila melanogaster*. *Cell*, 147, 690-703.
- HASLBECK, M., IGNATIOU, A., SAIBIL, H., HELMICH, S., FRENZL, E., STROMER, T. & BUCHNER, J. 2004. A domain in the N-terminal part of Hsp26 is essential for chaperone function and oligomerization. *J Mol Biol*, 343, 445-55.
- KEDERSHA, N. L., GUPTA, M., LI, W., MILLER, I. & ANDERSON, P. 1999. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J Cell Biol*, 147, 1431-42.
- KLOETZEL, P. M. & BAUTZ, E. K. 1983. Heat-shock proteins are associated with hnRNA in *Drosophila melanogaster* tissue culture cells. *EMBO J*, 2, 705-10.
- KRIEHLER, T., RATTEI, T., WEINMAIER, T., BEPPERLING, A., HASLBECK, M. & BUCHNER, J. 2010. Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J*, 24, 3633-42.
- MAAROUFI, H. & TANGUAY, R. M. 2013. Analysis and phylogeny of small heat shock proteins from marine viruses and their cyanobacteria host. *PLoS One*, 8, e81207.
- MCHAOURAB, H. S., GODAR, J. A. & STEWART, P. L. 2009. Structure and mechanism of protein stability sensors: chaperone activity of small heat shock proteins. *Biochemistry*, 48, 3828-37.
- MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y. M., BUSO, N., COWLEY, A. P. & LOPEZ, R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res*, 41, W597-600.
- MICHAUD, S., LAVOIE, S., GUIMOND, M. O. & TANGUAY, R. M. 2008. The nuclear localization of *Drosophila* Hsp27 is dependent on a monopartite arginine-rich NLS and is uncoupled from its association to nuclear speckles. *Biochim Biophys Acta*, 1783, 1200-10.
- MOUTAOUFIK, M. T., MORROW, G., MAAROUFI, H., FÉRARD, C., FINET, S. & TANGUAY, R. M. 2016. Oligomerization and chaperone-like activity of *Drosophila melanogaster* small heat shock protein DmHsp27 and three arginine mutants in the alpha-crystallin domain. *Cell Stress and Chaperones*, 1-12.

- NIGG, E. A. 1997. Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature*, 386, 779-87.
- OHNO, M., FORNEROD, M. & MATTAJ, I. W. 1998. Nucleocytoplasmic transport: the last 200 nanometers. *Cell*, 92, 327-36.
- PALMERI, D. & MALIM, M. H. 1999. Importin beta can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin alpha. *Mol Cell Biol*, 19, 1218-25.
- PIETROWSKI, D., DURANTE, M. J., LIEBSTEIN, A., SCHMITT-JOHN, T., WERNER, T. & GRAW, J. 1994. Alpha-crystallins are involved in specific interactions with the murine gamma D/E/F-crystallin-encoding gene. *Gene*, 144, 171-8.
- POULAIN, P., GELLY, J. C. & FLATTERS, D. 2010. Detection and architecture of small heat shock protein monomers. *PLoS One*, 5, e9990.
- ROBERT, X. & GOUET, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, 42, W320-4.
- SCHARF, K. D., HEIDER, H., HOHFELD, I., LYCK, R., SCHMIDT, E. & NOVER, L. 1998. The tomato Hsf system: HsfA2 needs interaction with HsfA1 for efficient nuclear import and may be localized in cytoplasmic heat stress granules. *Mol Cell Biol*, 18, 2240-51.
- SCHARF, K. D., SIDDIQUE, M. & VIERLING, E. 2001. The expanding family of Arabidopsis thaliana small heat stress proteins and a new family of proteins containing alpha-crystallin domains (Acid proteins). *Cell Stress Chaperones*, 6, 225-37.
- SIDDIQUE, M., GERNHARD, S., VON KOSKULL-DORING, P., VIERLING, E. & SCHARF, K. D. 2008. The plant sHSP superfamily: five new members in Arabidopsis thaliana with unexpected properties. *Cell Stress Chaperones*, 13, 183-97.
- SIDDIQUE, M., PORT, M., TRIPP, J., WEBER, C., ZIELINSKI, D., CALLIGARIS, R., WINKELHAUS, S. & SCHARF, K. D. 2003. Tomato heat stress protein Hsp16.1-CIII represents a member of a new class of nucleocytoplasmic small heat stress proteins in plants. *Cell Stress Chaperones*, 8, 381-94.
- SINGH, K., GROTH-VASSELLI, B. & FARNSWORTH, P. N. 1998. Interaction of DNA with bovine lens alpha-crystallin: its functional implications. *Int J Biol Macromol*, 22, 315-20.
- STAMLER, R., KAPPE, G., BOELEN, W. & SLINGSBY, C. 2005. Wrapping the alpha-crystallin domain fold in a chaperone assembly. *J Mol Biol*, 353, 68-79.
- TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*, 30, 2725-9.
- VAN DE KLUNDERT, F. A. & DE JONG, W. W. 1999. The small heat shock proteins Hsp20 and alphaB-crystallin in cultured cardiac myocytes: differences in cellular localization and solubilization after heat stress. *Eur J Cell Biol*, 78, 567-72.
- VAN MONTFORT, R., SLINGSBY, C. & VIERLING, E. 2001a. Structure and function of the small heat shock protein/alpha-crystallin family of molecular chaperones. *Adv Protein Chem*, 59, 105-56.

- VAN MONTFORT, R. L., BASHA, E., FRIEDRICH, K. L., SLINGSBY, C. & VIERLING, E. 2001b. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat Struct Biol*, 8, 1025-30.
- VICART, P., CARON, A., GUICHENEY, P., LI, Z., PREVOST, M. C., FAURE, A., CHATEAU, D., CHAPON, F., TOME, F., DUPRET, J. M., PAULIN, D. & FARDEAU, M. 1998. A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy. *Nat Genet*, 20, 92-5.
- WANG, X., OSINSKA, H., KLEVITSKY, R., GERDES, A. M., NIEMAN, M., LORENZ, J., HEWETT, T. & ROBBINS, J. 2001. Expression of R120G-alphaB-crystallin causes aberrant desmin and alphaB-crystallin aggregation and cardiomyopathy in mice. *Circ Res*, 89, 84-91.
- WHELAN, S. & GOLDMAN, N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*, 18, 691-9.
- YU, J., PACIFICO, S., LIU, G. & FINLEY, R. L., JR. 2008. DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. *BMC Genomics*, 9, 461.