# Specific Sequences in the N-terminal Domain of Human Small Heat Shock Protein HSPB6 Dictate Preferential Heterooligomerization with the Orthologue HSPB1

**Michelle Heirbaut[1], Frederik Lermyte[2], Esther M. Martin[2,3], Steven Beelen[1], Frank Sobott[2,3,4], Sergei V. Strelkov[1,5], Stephen D. Weeks[1,5]**

[1] Laboratory for Biocrystallography, Dept. of Pharmaceutical and Pharmacological Sciences, KU Leuven, Belgium
[2] Biomolecular and Analytical Mass Spectrometry Group, Dept. of Chemistry, University of Antwerp, Belgium
[3] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom
[4] School of Molecular and Cellular Biology, University of Leeds, LS2 9JT, United Kingdom

Running title: *N-terminal determinants of HSPB6 heterooliogomerization*

[5] To whom correspondence should be addressed:
Sergei V. Strelkov, Laboratory for Biocrystallography, Department of Pharmaceutical and Pharmacological Sciences, Herestraat 49 box 822, B-3000 Leuven, Belgium, Tel: +32 16330845, Fax: +32 16323469, e-mail: sergei.strelkov@kuleuven.be,
Stephen D. Weeks, Laboratory for Biocrystallography, Department of Pharmaceutical and Pharmacological Sciences, Herestraat 49 box 822, B-3000 Leuven, Belgium, Tel: +32 16377204, Fax: +32 16323469, e-mail: stephen.weeks@kuleuven.be

**SUPPLEMENTAL FIGURE LEGENDS**

**FIGURE S1. SAXS analysis of the hetero-oligomeric complexes of HSPB1 and HSPB6 10AA deletions.** For each complex 250 frames were collected upon elution from a Shodex KW-404F column. A 10 frame moving algorithm was used to improve signal-to-noise ratio before AutoRg analysis of the all collected frames, the calculated $R_g$ and forward scattering $I_0$ of each collected frame are shown. The $R_g$ is plotted as a black cross for each and scaled on the left axis, the $I_0$ is shown as a red circle and scaled on the right axis.

**FIGURE S2. MS analysis showing subunit exchange of HSPB6.Δ21-30 and HSPB6.Δ31-40.** Both truncations were mixed with $^{15}$N-labeled protein, allowed to exchange at 37°C overnight before analysis using native MS. Equivalent spectra for the wild-type HSPB6 have been previously reported (1).

**Figure S3. Chemical cross-linking of the HSPB1 and HSPB6 deletion mutants.** (A) The β7 mediated dimer interface of HSPB1. In the $AP_{II}$ type strand register the sulfur groups of the native cysteine (C137) residues are at a suitable distance for cross-linking, either by disulfide formation or by a short chemical cross-linker. (B) The two preferred 3D-conformers of bismaleimidoethane (BMOE) calculated using the Frog2 server (2). With three rotable bonds BMOE can form extended or compact conformations. In the latter case the shortest distance between the thiol reactive maleimide groups is suitable for cross-linking of HSPB1 or the HSPB6* mutants across the $AP_{II}$ dimer interface. (C) Reducing SDS-PAGE analysis of HSPB1 and HSPB6* mutants following incubation with BMOE for 15 minutes at 4°C. The constructs, either alone or mixed together as stated above the lanes, were preincubated at 42°C for 1.5 hr under reducing conditions. These samples were desalted and then a 1.1 fold molar excess of BMOE was added to each mixture. The reaction was stopped by addition of sample loading buffer containing 100 mM β-mercaptoethanol. (D) Non-reducing SDS-PAGE analysis of the HSPB1 and HSPB6* desalted mixtures following incubation at 4°C for 1 hr in the absence of either reducing agent or BMOE.

**Figure S4. Sequence properties of the human sHSP N-terminal domains.** (A and B) Consensus secondary structure (CSS) prediction of the NTDs of HSPB6 and HSPB1. The local substructure assignment, where H represents an α-helix and E a β-strand, were made using the GeneSilico meta-server (3) the results of which are based on 18 different secondary structure predictors. (C) Multiple sequence alignment (MSA) of the NTDs of the ten human sHSPs. The terminating boundary of the NTD for each sequence was determined based on a MSA of the full-length proteins, to ascertain the start of the α-crystallin domain. The truncated sequences were then realigned using Muscle (4) followed by manual editing to accommodate the large insertions present in HSPB1 and HSPB10. Residues that show conservation, at or above a 40% identity threshold, are shaded using the Clustal color scheme.

REFERENCES


1. Heirbaut, M., Lermyte, F., Martin, E. M., Beelen, S., Verschueren, T., Sobott, F., Strelkov, S. V., and Weeks, S. D. (2016) The preferential heterodimerization of human small heat shock proteins HSPB1 and HSPB6 is dictated by the N-terminal domain. *Arch. Biochem. Biophys.* **610**, 41–50
2. Miteva, M. A., Guyon, F., and Tufféry, P. (2010) Frog2: Efficient 3D conformation ensemble generator for small compounds. *Nucleic Acids Res.* **38**, W622-627
3. Kurowski, M. A., and Bujnicki, J. M. (2003) GeneSilico protein structure prediction meta-server. *Nucleic Acids Res.* **31**, 3305–3307
4. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797

Figure S1

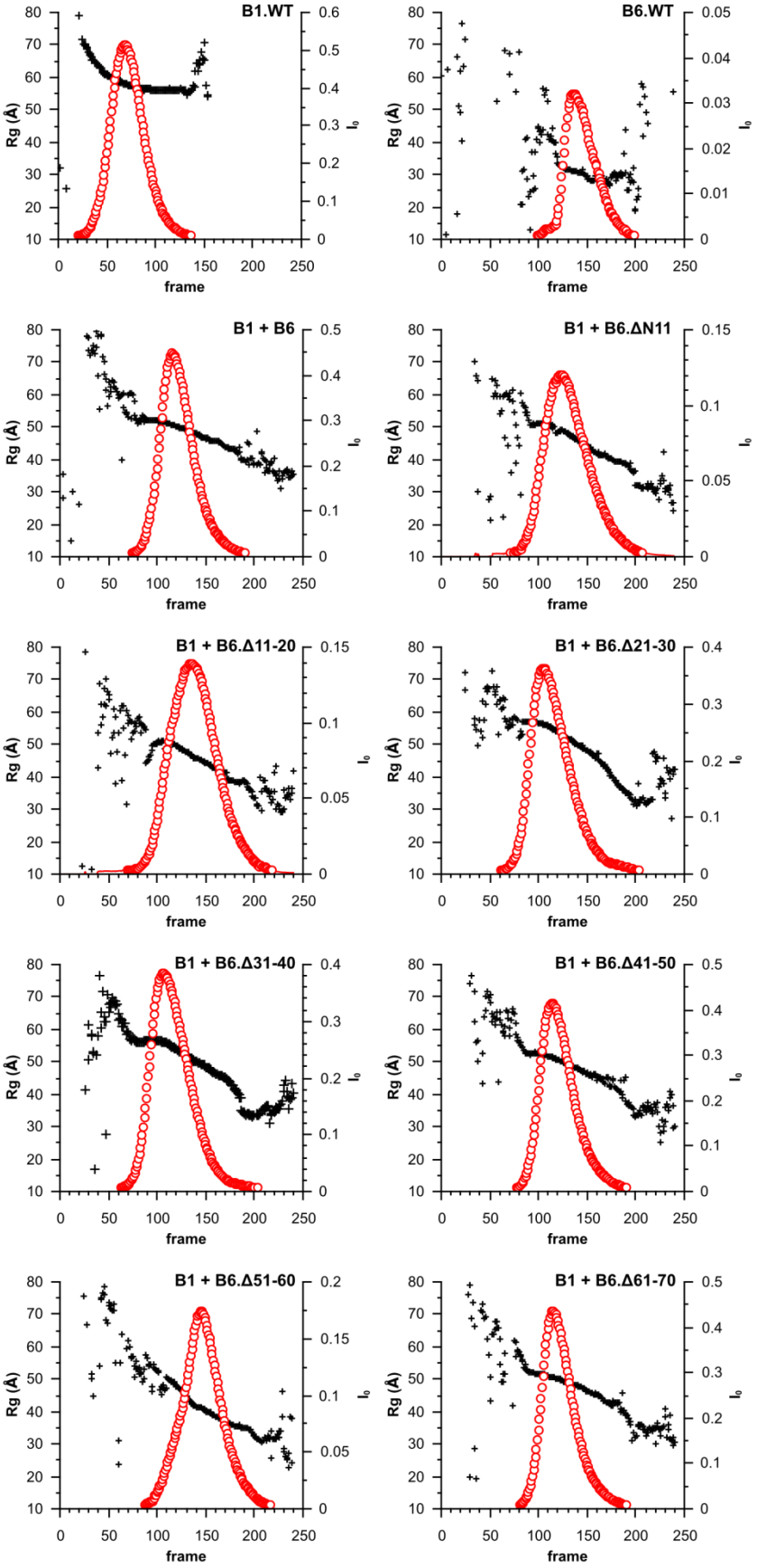
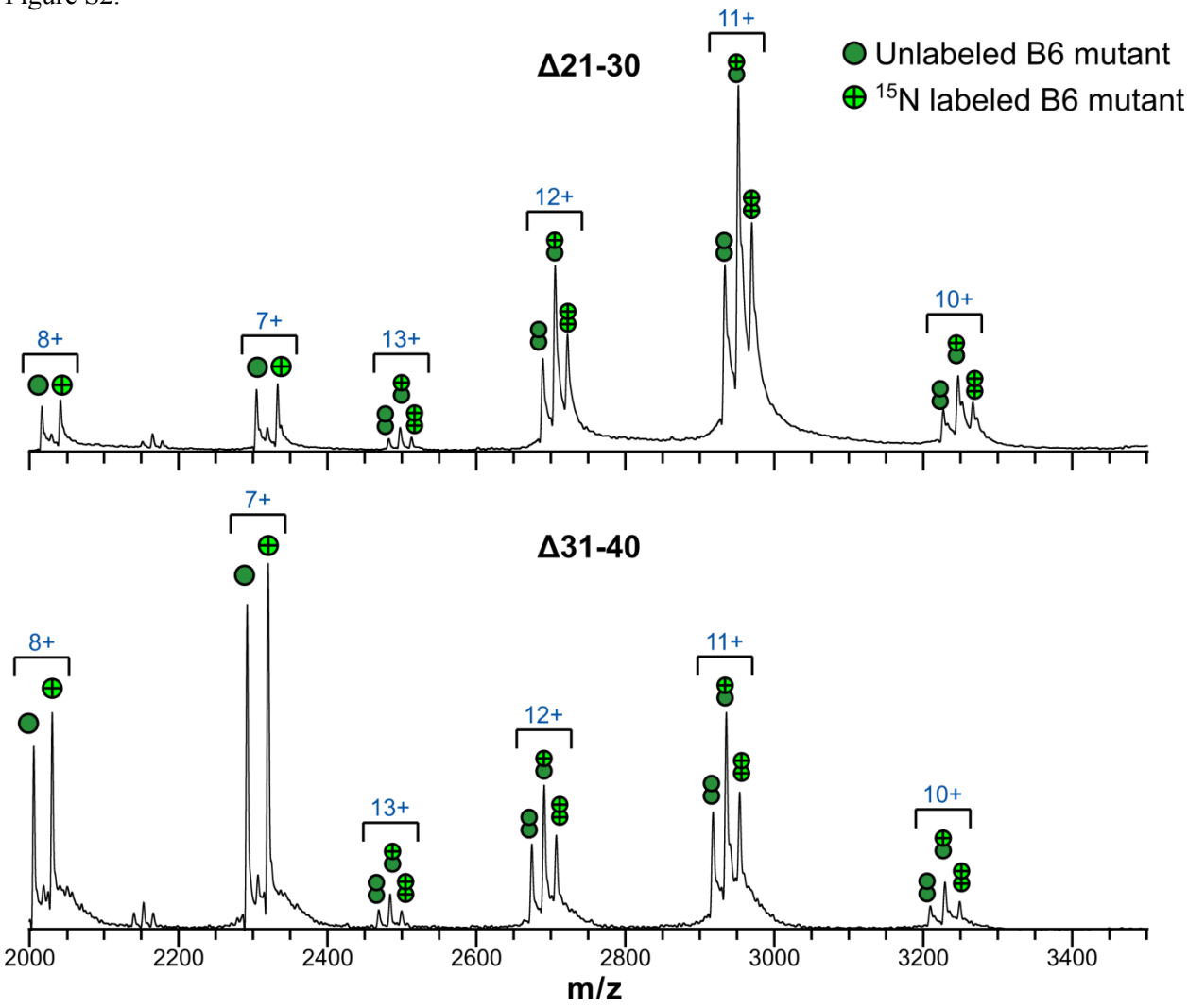

B1.WT
B6.WT
B1 + B6
B1 + B6.ΔN11
B1 + B6.Δ11-20
B1 + B6.Δ21-30
B1 + B6.Δ31-40
B1 + B6.Δ41-50
B1 + B6.Δ51-60
B1 + B6.Δ61-70
Rg (Å)
I0
frame

Figure S2.


Δ21-30
Unlabeled B6 mutant
15N labeled B6 mutant
11+
12+
10+
8+
7+
13+
Δ31-40
7+
8+
11+
12+
13+
10+
2000
2200
2400
2600
2800
3000
3200
3400
m/z

Supplemental S3.

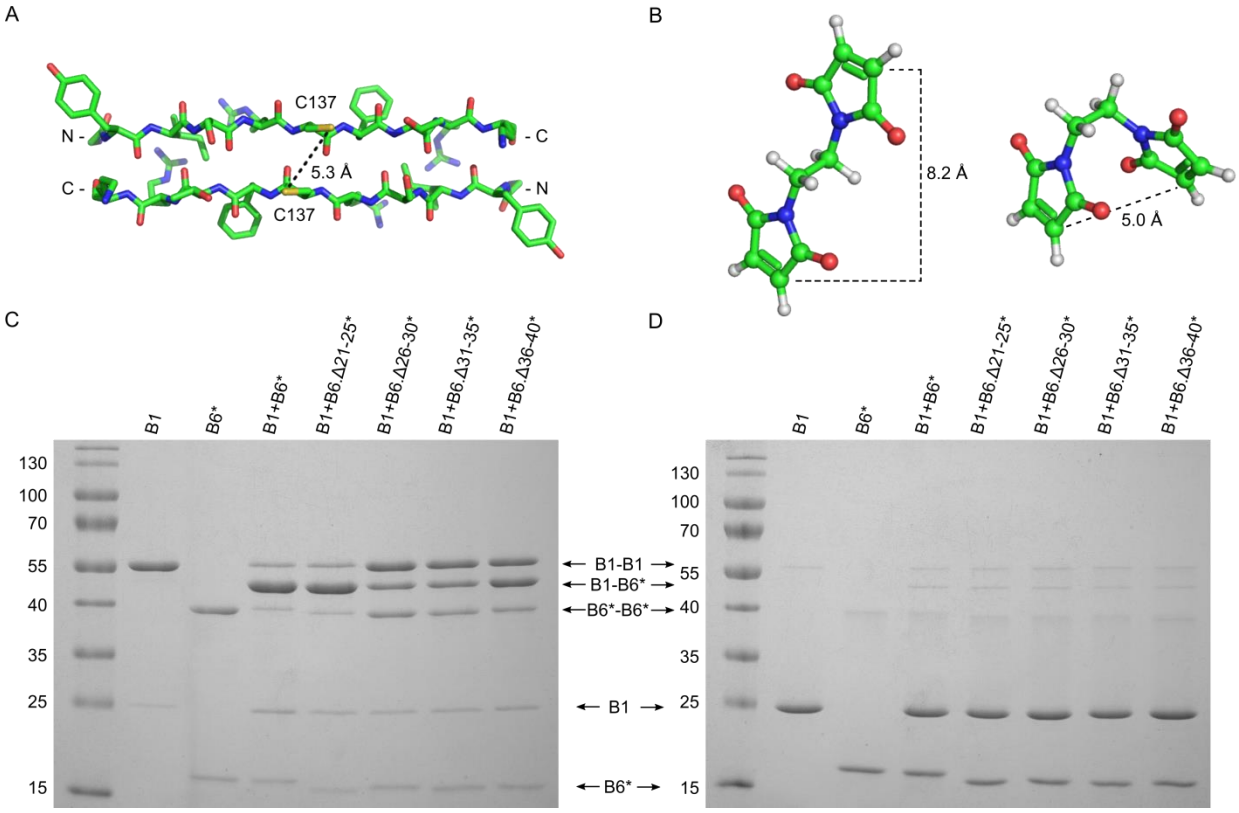
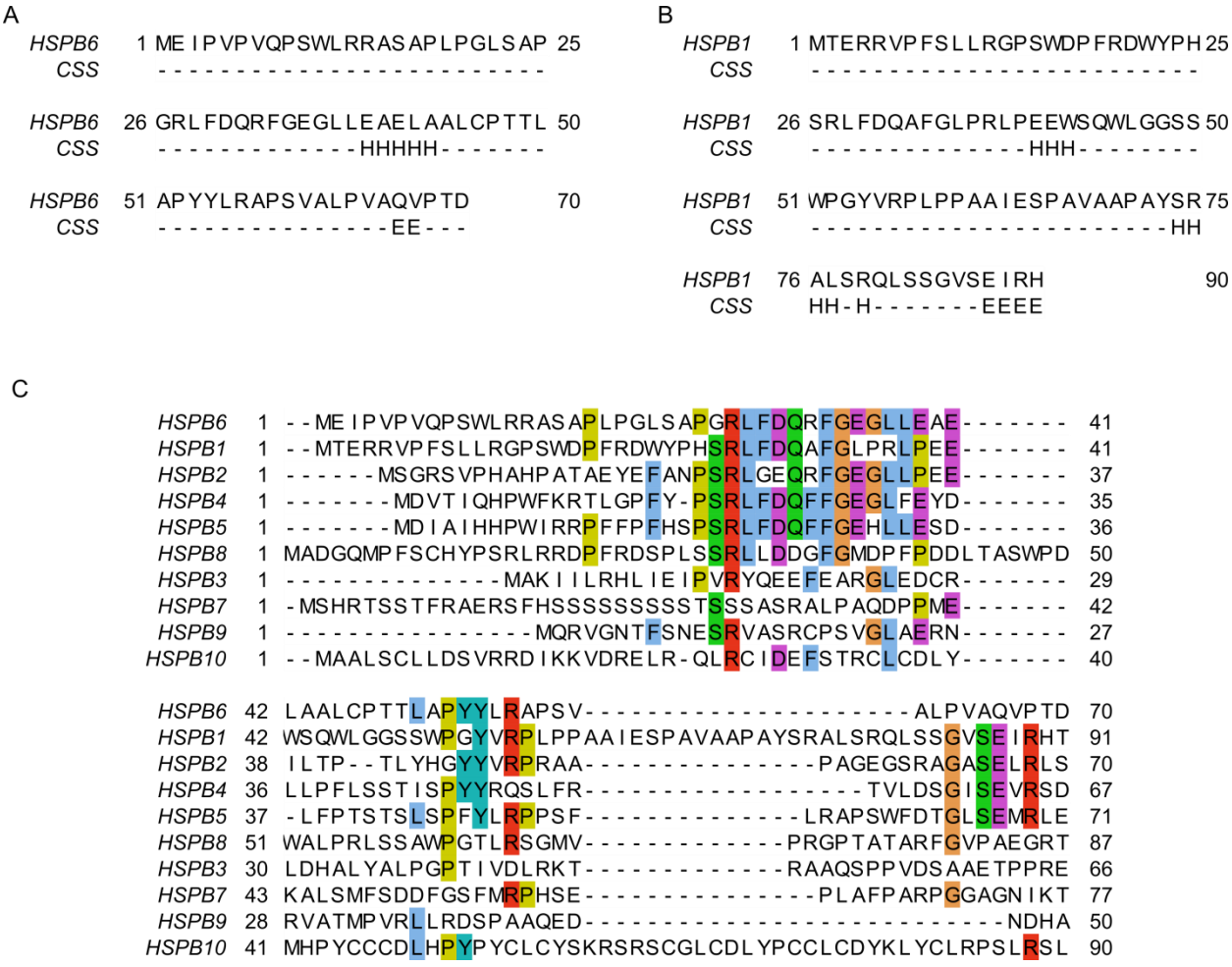


Figure S4.


A
HSPB6 1 MEIPVPVQPSWLRRASAPLPGLSAP 25
CSS - - - - - - - - - - - - - - - - - - - - - - - - -
HSPB6 26 GRLFDQRFGEGLLEAELAALCPTTL 50
CSS - - - - - - - - - - - - - HHHHH- - - - - - -
HSPB6 51 APYYLRAPSVALPVAQVPTD 70
CSS - - - - - - - - - - - - - - - EE- - -
B
HSPB1 1 MTERRVPFSLLRGPSWDPFRDWYPH 25
CSS - - - - - - - - - - - - - - - - - - - - - - - - -
HSPB1 26 SRLFDQAFGLPRLPEEWSQWLGGSS 50
CSS - - - - - - - - - - - - - - HHH- - - - - - - -
HSPB1 51 WPGYVRPLPPAAIESPAVAAPAYSR 75
CSS - - - - - - - - - - - - - - - - - - - - - - - HH
HSPB1 76 ALSRQLSSGVSEIRH 90
CSS HH-H- - - - - - - EEEE
C
HSPB6 1 - -MEIPVPVQPSWLRRASAPLPGLSAPGRLFDQRFGEGLLEAE- - - - - - - 41
HSPB1 1 - -MTERRVPFSLLRGPSWDPFRDWYPHSRLFDQAFGLPRLPEE- - - - - - - 41
HSPB2 1 - - - - - -MSGRSVPHAHPATAEYEFANPSRLGEQRFGEGLLPEE- - - - - - - 37
HSPB4 1 - - - - - - -MDVTIQHPWFKRTLGPFY-PSRLFDQFFGEGLFEYD- - - - - - - 35
HSPB5 1 - - - - - - -MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESD- - - - - - - 36
HSPB8 1 MADGQMPFSCHYPSRLRRDPFRDSPLSSRLLDDGFGMDPFPDDLTASWPD 50
HSPB3 1 - - - - - - - - - - - - - - -MAKIILRHLIEIPVRYQEEFEARGLEDCR- - - - - - - 29
HSPB7 1 -MSHRTSSTFRAERSFHSSSSSSSSSTSSSASRALPAQDPPME- - - - - - - 42
HSPB9 1 - - - - - - - - - - - - - - - - -MQRVGNTFSNESRVASRCPSVGLAERN- - - - - - - 27
HSPB10 1 - -MAALSCLLDSVRRDIKKVDRELR-QLRCIDEFSTRCLCDLY- - - - - - - 40
HSPB6 42 LAALCPTTLAPYYLRAPSV- - - - - - - - - - - - - - - - - - - - - - ALPVAQVPTD 70
HSPB1 42 WSQWLGGSSWPGYVRPLPPAAIESPAVAAPAYSRALSRQLSSGVSEIRHT 91
HSPB2 38 ILTP- - TLYHGYYVRPRAA- - - - - - - - - - - - - - - PAGEGSRAGASELRLS 70
HSPB4 36 LLPFLSSTISPYYRQSLFR- - - - - - - - - - - - - - - - - - TVLDSGISEVRSD 67
HSPB5 37 -LFPTSTSLSPFYLRPPSF- - - - - - - - - - - - - - - LRAPSWFDTGLSEMRLE 71
HSPB8 51 WALPRLSSAWPGTLRSGMV- - - - - - - - - - - - - PRGPTATARFGVPAEGRT 87
HSPB3 30 LDHALYALPGPTIVDLRKT- - - - - - - - - - - - - RAAQSPPVDSAAETPPRE 66
HSPB7 43 KALSMFSDDFGSFMRPHSE- - - - - - - - - - - - - - - PLAFPARPGGAGNIKT 77
HSPB9 28 RVATMPVRLLRDSPAAQED- - - - - - - - - - - - - - - - - - - - - - - - - - NDHA 50
HSPB10 41 MHPYCCCDLHPYPYCLCYSKRSRSCGLCDLYPCCLCDYKLYCLRPSLRSL 90