

Supplementary Information, Data S1

Method section:

Sequence recovery and alignment

The representative Hsps with known cell localization (Table S1) were used as queries, and the extensive searches were performed using BLASTP and TBLASTN in GenBank (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>) and sequenced genomes of bacteria (including alphaproteobacteria, firmicute, cyanobacteria, actinobacteria and Cytophaga-Fibrobacter-Bacteroides group), fungi, protists, plants and animals. The archaeobacterial Hsps were omitted because of their nonexistence of Hsp90s. The sequences of Hsp homolog from various species that showed high scores (E value < 0.01) in the BLAST searches were retrieved. All the selected sequences were validated to contain typical Hsp domains by BLASTP. The recovered protein sequences were aligned using CLUSTAL X with a Gonnet 250 matrix, then adjusted manually with the aid of Bioedit software.

Phylogenetic analysis

The phylogenetic analysis was completed with MEGA version 3.1. The trees were constructed using the neighbor-joining method with the Jones-Taylor-Thornton matrix-based model, and sites containing missing data or alignment gaps were removed in a pairwise fashion. A total of 1000 bootstrap replications were used to test the topology in all cases. Percentage bootstrap values are reported with each tree. The full amino acid sequences were used to construct phylogenetic trees of the large Hsps, while only the α -crystalline domains (about 80 amino acids) were selected in the

phylogenetic analysis of sHsps because of the variable terminals.