FULL PAPER *Molecular Biology*

Molecular Cloning of Horse Hsp90 cDNA and Its Comparative Analysis with Other Vertebrate Hsp90 Sequences

Kim PEPIN[1)], Fumitaka MOMOSE[1)**], Nobushige ISHIDA[2)] and Kyosuke NAGATA[1)*]

[1)]*Laboratory of Molecular Medical Engineering, Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Yokohama 226-8501 and* [2)]*Laboratory of Molecular and Cellular Biology, Equine Research Institute, JRA, 321–4 Tokami-Cho, Utsunomiya, Tochigi 320-0856, Japan*

(Received 12 June 2000/Accepted 26 September 2000)

ABSTRACT. Heat shock protein 90 (Hsp90), a molecular chaperone, is ubiquitous and involved in numerous cellular processes. To contribute to the relatively small collection of vertebrate *Hsp90* sequences in the gene data bank, we cloned and sequenced horse (*Equus caballus*) *Hsp90* α and β cDNAs. This enabled identification of horse-specific primers for development of a convenient PCR-based method that could monitor horse stress tolerance. We analyzed the sequence data comparatively and phylogenetically with other *Hsp90* cDNA sequences, and identified vertebrate-specific and isoform-specific conserved regions to facilitate future molecular investigations of Hsp90 functions. We found 4 highly conserved regions to vertebrate Hsp90 exclusively and 27 amino acids conserved among but differing between Hsp90 α and Hsp90 β sequences. Protein-based phylogenetic trees revealed high conservation between mammal species within Hsp90 α and β clusters. Comparison of nucleotide and amino acid substitution levels suggests that horse Hsp90 β has undergone strong purifying selection, while rat Hsp90 β and hamster Hsp90 α have been positively selected. Surprisingly, fish *Hsp90* α genes clearly clustered with *Hsp90* β genes, and no distinct placement of fish Hsp90 α protein was found. The *Hsp90* α isoform is apparently the result of β gene duplication. Our results highlight the importance of organism- and isoform-specific Hsp90 functional analyses in describing the role of Hsp90 in cells.

KEY WORDS: heat shock protein, Hsp90, phylogenetic analysis, sequence comparison.

J. Vet. Med. Sci. 63(2): 115–124, 2001

The heat shock protein 'Hsp90' is highly abundant in cells [5], ubiquitous and largely conserved in quite different groups of organisms [13]. Understating its functional complexity, it is classified as a cellular 'molecular chaperone'. Since it is a central component of chaperone complexes that are critical in numerous fundamental physiological processes, Hsp90 may be better described as a 'multi-chaperone system host'. Its chaperone activities range from interacting with the specific proteins/chaperones involved in cell cycle control and hormone signaling [1, 17 and 4, 8, 16; respectively], to more general interactions such as binding to metastable cellular proteins following stress [5, 10, 15, 31]. The stress-induced activities of Hsp90 help, not only to maintain competent refolding structure, but also to prevent nonspecific protein aggregation. Further illustrating its functional versatility, Hsp90 has even been found to be required for Hepatitis B replication via interacting with reverse transcriptase [14]. However, while many Hsp90-cellular protein interactions have been revealed, its specific functional domains are not well described, and organism-specific or isoform-specific functions remain uncharacterized. Considering the high cellular abundance and apparently high functional versatility of Hsp90, it is likely that the biological relevance and/or method of a given Hsp90 function differs between groups of organisms. To understand why current individuals burden themselves with a high level of Hsp90 production, group-specific functional analyses are crucial. Moreover, it is important to clarify regulatory mechanisms of Hsp90 gene expression in each organism under a variety of physiological conditions, since heat shock factors are involved in stress-induced gene expression [21]. Recently, it is reported that signalling pathways of interleukin-6 and interferon γ, multifunctional cytokines, regulate Hsp90 gene expression [28, 29]. However, the detailed expression regulation mechanisms in a vareity of organisms and tissues, are not clearly understood.

Hsp90 cDNA and amino acid sequences of a target animal of research interest are critical to designing experiments for functional analyses and understanding evolutionary relationships. Sequence information is also essential for studying gene expression regulation. To enable these pursuits using horse Hsp90, we present the molecular cloning and sequence analysis of race horse *Hsp90* cDNAs. The expression level of Hsp90 in horse lymphocytes is found elevated upon heat stress [12]. Race horses are subject to numerous stresses such as cardiovascular and muscular stresses. Hsp90 expression level has been found to increase in the presence of increased levels of catecholamines (which are released by the hypothalamic-pituitary-adrenal axis in response to physiological stress) [2 and ref. therein]. Furthermore, both Hsp90 α and β bind Ca^{+2} [23], which accumulates in muscle cells during muscle stress, and

* CORRESPONDENCE TO: NAGATA, K., Laboratory of Molecular Medical Engineering, Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Yokohama 226–8501, Japan.

** PRESENT ADDRESS: MOMOSE, S., Department of Virology, Centre for Basic Research, The Kitasato Institute, 5–9–1 Shirokane, Minato-ku, Tokyo 108–8642, Japan.

calmodulin, a Ca^{+2}-binding protein. Therefore, elevated Hsp90 levels may be good indicators of physiological stress in race horses. For the potential application of monitoring horse physiological stress via Hsp90 expression levels, we identify horse- and Hsp class-specific primers to be used in PCRs with horse cDNA. Such a monitoring system is superior to the northern blotting method which is incapable of discriminating species- and Hsp class-specific mRNA due to the lack of significant sequence and size differences.

In addition, since precise molecular mechanisms of Hsp90 stress response action are not well understood, we phylogenetically analyzed our sequence data with other vertebrate Hsp90 sequences, and compared vertebrate Hsp90 alignment data with those from a published Hsp90 amino acid sequence comparison that includes more diverse species [13]. We identify regions unique to each isoform, and conserved to vertebrates as a group, and describe the phylogenetic relationships of vertebrate *Hsp90* α and β genes. The results of the analysis are likely to be quite useful for planning studies aimed at deciphering the precise molecular mechanisms governing the role of Hsp90 in its response to particular physiological stresses and in clarifying isoform-specific and vertebrate-specific Hsp90 functions.

MATERIALS AND METHODS

Cloning and sequence determination of horse Hsp90 α *and* β *cDNAs*: A horse whole RNA fraction was prepared from 5 g of liver tissue following a standard RNA extraction protocol [24]. The isolated RNAs (10 μg) were hybridized with olido dT and used as templates for the synthesis of single-stranded complementary DNA with M-MLV reverse transcriptase (Gibco BRL). Purified cDNAs were then subjected to PCR-mediated amplification with Taq DNA polymerase (Takara). Based on nucleotide sequence analyses of vertebrate *Hsp90* cDNAs, the following oligonucleotide primers were used in the PCR reactions (where nucleotides in parentheses indicate mixtures): 5'-CA(AG)AC(AGCT)CA(AG)GACCAACC(AGCT)ATGGAGGAG-3', 5'-CC(AGCT)GAG(AGC)AAGTGCA(AGCT)CATGGA-3', as the 5' terminal primers for horse *Hsp90* α and β respectively, and 5'-(TC)TCCAT(AGCT)CG(AGCT)GA(AGCT)G(CT)(AG)TC(AGCT)TCATC-3', as a common 3' terminal primer. PCR-amplified DNA fragments were cloned into the *Eco*RV site of the Bluescript II SK+ plasmid vector. Nucleotide sequencing of cloned DNA was carried out based on the dideoxy-mediated chain termination method [25] with an ABI Prism 310 Automated Sequencer (Perkin-Elmer).

Identification of horse-specific primers: Horse-specific and human-specific primer sequences (5'-AACCCCCAGGAACGCACCCTGACGC-3' and 5'-AACCCTCAGGAACGTACCCTGACTT-3', respectively; corresponding to nucleotides 247–271) were deduced by comparing horse and human *Hsp90* sequences. The candidate primers were used in separate PCRs with horse and human cDNAs and RNAs (prepared as described above) at varying annealing temperatures. The products were electrophoresed in 1% agarose gels and visualized by ethidium bromide.

Analytical methods: Sequences were obtained from Genbank via a BLAST search with the horse *Hsp90* α and β seqeunces as queries. The included species and their Genbank accession numbers are summarized in Table 1. The 17 *Hsp90* sequences were aligned by Clustal W [30]. The areas extending beyond the shortest Hsp90 sequence (*ie.* horse Hsp90 β) were excluded from the analyses. Phylogenetic trees were constructed using the PHYLIP Ver. 3.572 package [9]. Firstly, sets of 1,000 and 500 bootstrap replicates, sampled from DNA and amino acid sequence data, respectively, were generated by SEQBOOT. Phylogenetic trees were based on the neighbor-joining distance method. The trees were assembled by inputting the bootstrapped data initially to DNADIST or PROTDIST programs and subsequently to NEIGHBOR. To all programs, each 17 sequence data set was entered randomly 10 times. All trees were unrooted. *Arabidopsis thaliana* was used as an outgroup. Final bootstrap values for both types of trees were calculated by CONSENSE.

In the alignment analyses, only the vertebrate sequences were included. The α and β sequences were aligned and assessed together and separately. The vertebrate alignments were also compared against the data of Gupta [13] which includes bacterial, fungal, invertebrate, and plant species. The endoplasmic reticulum-type Hsp90s in [13] were excluded from the comparisons.

RESULTS AND DISCUSSION

Following RT-PCRs with RNA prepared from horse liver tissues, we cloned the cDNAs corresponding to *Hsp90* α and β. Sequence determination of these cDNAs revealed the complete horse *Hsp90* α and β cDNA sequences (shown in Fig. 1) excluding the regions corresponding to the primer

Table 1. Species names and their Genbank Accession numbers for the analyzed gene sequences

Assigned Name	Gene Species Name	Accession Number
Human α	*Homo sapiens*	X15183
Horse α	*Equus caballus*	AB043677
Pig α	*Sus scrofa*	U94395
Mouse α	*Mus musculus*	J04633
Hamster α	*Cricetulus griseus*	L33676
Chicken α	*Gallus gallus*	X07265
Csalmon α	*Oncorhynchus tshawytscha*	U89945
Zfish α	*Danio rerio*	AF068773
Human β	*Homo sapiens*	Nm_007355
Horse β	*Equus caballus*	AB043676
Rat β	*Rattus sp.*	S45392
Mouse β	*Mus musculus*	M18186
Chicken β	*Gallus gallus*	X70101
Asalmon β	*Salmo salar*	AF135117
Zfish β	*Danio rerio*	AF068772
Nematode	*Brugia pahangi*	AJ005784
Arabidopsis	*Arabidopsis thaliana*	Y07613

Fig. 1. Alignment of horse and human *Hsp90* α and β sequences. Nucleotide numbers correspond to those of human *Hsp90*α sequence. Stars indicate nucleotide positions of 100% identity.

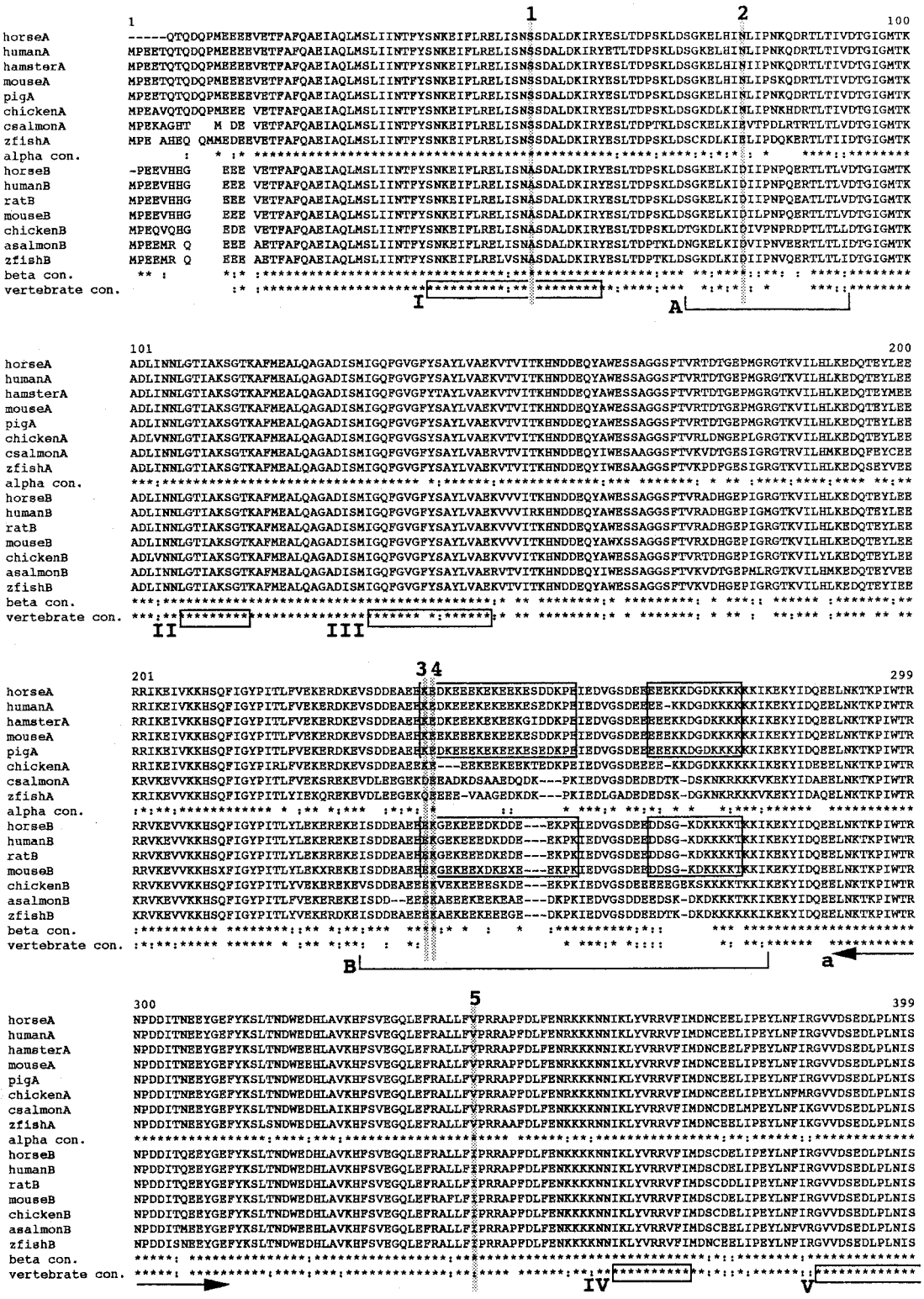
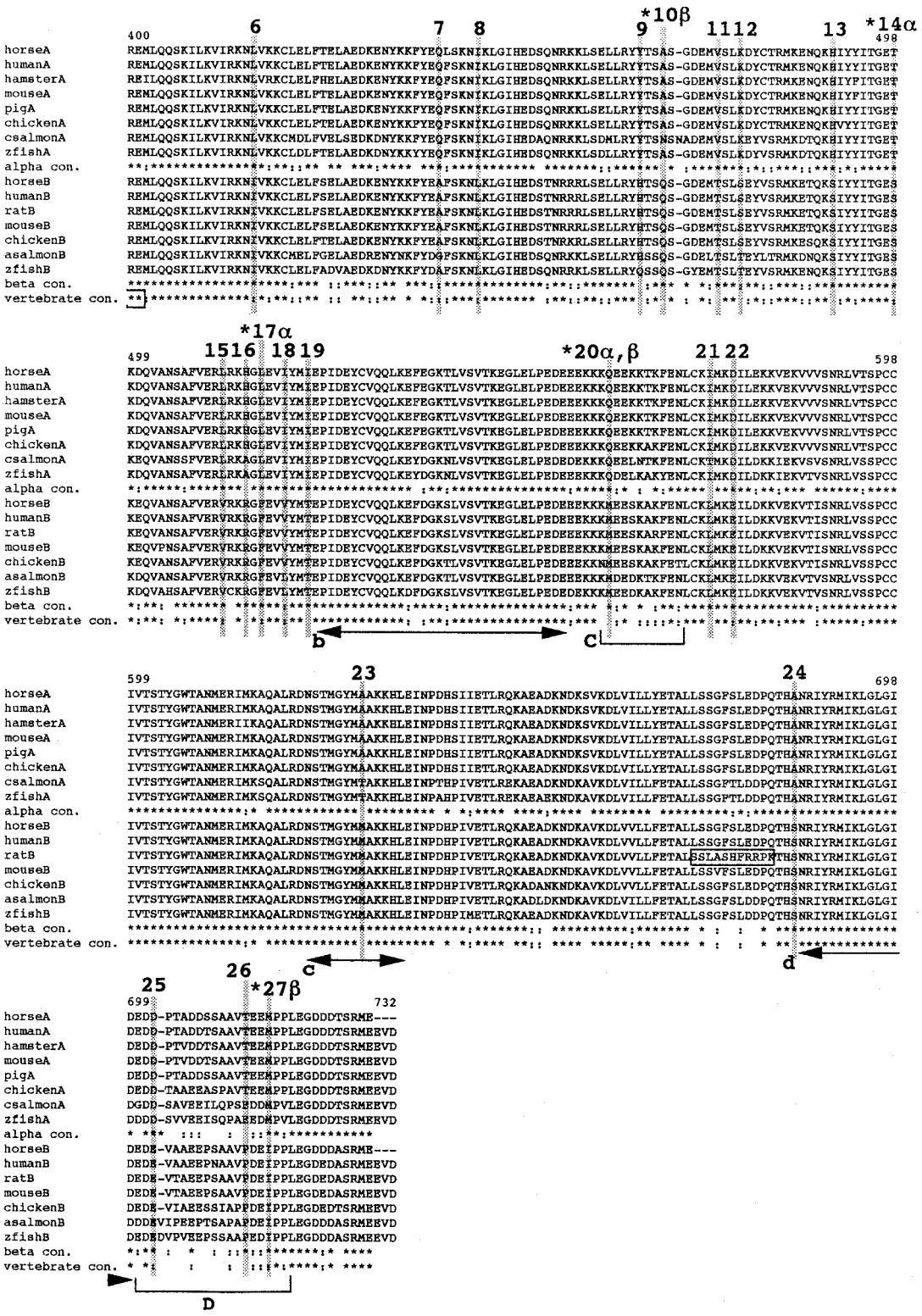


Fig. 2. Alignment of vertebrate Hsp90 α and β amino acid sequences. Amino acid positions are numbered according to human Hsp90 α and β sequences. α and β sequences are grouped, and their separate consensus evaluations are shown in the upper and lower parts of the alignments, respectively. A consensus evaluation for all vertebrate α and β sequences is given below the α and β groups. Stars indicate 100% identity and colons represent conservation of one of the following functional groups: non-polar, polar/uncharged, charged/acidic, or charged/basic. Horse sequences were deduced from the cDNA nucleotide sequences shown in Fig. 1. The boxed regions labeled by roman numerals are those corresponding to previously cited highly conserved areas as identified from Hsp90 alignments of a broad spectrum of species [13].



Determined the alignments obtained in the present analysis, large case letters name areas of high variability among vertebrate Hsp90 amino acid sequences in both data sets, and small case letters denote the areas of discrepancy in the level of conservation between the 2 data sets (*ie*. ours and 13). Numerically labeled, shaded amino acids, and boxed areas in Region B indicate positions that are conserved within α or *β* sequences but differ between them. Those positions marked with a star indicate amino acid substitutions unique to and ubiquitous within one of vertebrate Hsp90 α or *β* as determined from comparison with bacterial, fungal, plant, and invertebrate Hsp90 sequences in [13]. The box enclosing a portion of the rat *β* amino acid sequence highlights a highly divergent region of rat Hsp90*β*.

mixtures used for cloning. Nucleotide sequences were submitted to DDBJ and annotated as accession numbers AB043677 and AB043676 for horse *Hsp90 α* and horse *Hsp90 β*, respectively. As expected, horse *Hsp90* was found to be highly homologous to the human *Hsp90* (a detailed comparison of vertebrate Hsp90 is discussed below). The sequence information is key for functional analyses of horse Hsp90. Furthermore, as little is known of the horse family, accumulating sequence data from different horse strains, is important in revealing their evolutionary relationships.

In race horses, monitoring Hsp90 expression is a relatively quick and easy way to know their physiological status over time, or their tolerance to a given stress level. For this purpose, we identified a horse *Hsp90β*-specific primer (5'-AACCCCCAGGAACGCACCCTGACGC-3', corresponding to nucleotides 247–271 in Fig. 1) and semi-quantitatively analyzed the expression of horse-specific *Hsp90* mRNA in cultured cells by designing specific primers which contain the horse-specific dinucleotide sequence at their 3' termini (data not shown). In horse *Hsp90β*-specific amplification by PCR, annealing temperature was a critical factor. Non-specific primer annealing was observed for temperatures at and below 66°C, and very little product was detected at temperatures above 69°C. At an annealing temperature of 68°C, this PCR method specifically amplified the horse *Hsp90 β* derived from an mRNA preparation. Since a northern blot analysis could not discriminate horse mRNA when mixed with mRNA of other species, the PCR assay system is superior to assess horse stress levels.

To facilitate future investigations of Hsp90's functional significance to vertebrates, we conducted detailed sequence and phylogenetic analyses of vertebrate Hsp90 and compared these data to alignments of Hsp90s from more diverse species [data in 13]. The published Hsp90 amino acid alignments including bacterial, fungal, invertebrate, plant, and vertebrate Hsp90s [13], revealed 5 regions (I-V in Fig. 2) of >94% conservation. As expected, our alignments similarly show high levels of conservation (>99%) in these regions. Regions I-IV have been found to be ATP-binding sites, while Region V maps to the steroid receptor-binding domain of Hsp90 [6, 7, 11, 27; summarized in Table 2]. The high conservation of these regions highlights the importance of ATP-binding in the Hsp90-involved processes of diverse species and implies some common steroid receptor-binding function between them. ATP-binding has been found to cause structural changes in Hsp90 necessary for accommodating its interaction with target substrates (*eg.* p23, [11]). As p23 has not yet been identified in some species, the high conservation of the ATP-binding domains does not necessarily imply an important function common to vertebrates and other species. Therefore, it is important to design organism- , or closely related organism-, specific experiments when attempting to describe Hsp90's function.

The alignment comparisons also revealed variable regions common to both our alignments and those in [13] (location shown as A-D in Fig. 2, level of conservation in the 2 data sets summarized in Table 2). Regions A and C map to the N-terminal chaperone function domain and steroid receptor-binding domains respectively. Region A is thought to form exposed coils [27], while Region C is charged and in helix and coil structure [19], suggesting that both regions are involved in intermolecular interaction. Therefore, the high variability in these regions observed in our data and those of [13], may be due to organism-specific adaptation to or to co-evolution with the cellular proteins/processes involved in N-terminal protein chaperoning and steroid receptor-binding.

Interestingly, while the diverse species alignments [in 13, data not shown here] did not show a trend in the type of amino acid substitutions within the highly variable region B, most of the observed variability in Region B of our alignments arose from differences in the fish and chicken sequences relative to mammalian sequences, and differences between mammalian α and *β* sequences (compare boxed regions in Region B of Fig. 2). This region has been found to be critical to key Hsp90 functional mechanisms such as regulating N-terminal chaperone functions [26] and nucleotide access to the C-terminal domain [20], and in increasing the affinity for N-terminal binding of denatured proteins [26]. The relatively high conservation within mammalian α and *β* sequences in this area, compared to the less biased variability between more diverse species (or even lack of parts of Region B), suggests not only that mammalian Hsp90 has acquired novel chaperone activities, but that α and *β* isoforms may regulate different chaperone activities. This region is likely critical to understanding Hsp90 functions in mammalian cells. In addition, Fig. 2 shows 2 sites in Region B (numbered 3 and 4) where amino acids are conserved within but differ between vertebrate Hsp90 α and *β* isoforms. These sites may be important to defining functional difference(s) between isoforms in the Hsp90 functions enabled by Region B. Similar to Region B, Region D appeared unbiased in its variability in the diverse species alignments [in 13; data not shown here], whereas the high variability observed in the vertebrate alignments, was mainly due to differences between α and *β* isoforms. This region maps to the C-terminal Tetratricopeptide Repeat Protein (TPR)-binding domain and has been found to interact with immunophilins and hop [6]. It's acidic and coiled nature [6] indicates direct interaction with substrates. Therefore, the unbiased/high variability observed in the diverse species alignments [from 13] suggests that TPR-binding involves quite different proteins, or even none at all in some species, while the conservation among vertebrate α and *β* sequences, but variation between them, indicates a divergence in the chaperoning mechanisms of vertebrate Hsp90 α and *β* isoforms.

Next we identified regions that were highly conserved among vertebrate Hsp90s but quite variable among Hsp90s of more diverse species (Regions a-d illustrated in Fig. 2; level of conservation in the 2 data sets summarized in Table 2). Region a is located in the steroid receptor domain [6; references therein], Region b is apparently important for ATP-binding [7] and in Hsp90's response to heat [18], and

Table 2. Location and summary of the conserved and variable regions for vertebrate Hsp90 compared to corresponding regions of Hsp90 from more diverse species. In the 'Data set' column, 12 refers to the reference number and P represents our data set. The regions labeled by roman numerals are those corresponding to previously cited highly conserved areas as identified from Hsp90 alignments of a broad spectrum of species [13]. Determined from the alignments obtained in the present analysis, large case letters name the areas of highest variability among vertebrate Hsp90 amino acid sequences, and small case letters denote the areas of discrepancy in the level of conservation between the 2 data sets (*ie.* ours and [13]). All regions are illustrated on Fig. 2.

Region	Data set	Location	Length	Total Number of # of Sites	Conserved Sites	% Conservation	Structural Nature	Proposed Function
I	12	39–60	22	330	310	93.9	helix,	Substrate binding [27],
	P		22	330	322	97.6	buried [27]	ATP-binding (necessary for p23 function) [11]
II	12	107–115	9	135	127	94.1	loop/helix,	ATP-binding (necessary
	P		9	135	135	100.0	exposed [27]	for p23 function) [11] Forms adjustment mechanism critical to substrate binding processes [27]
III	12	131–146	16	240	226	94.2	helix,	Substrate binding [27],
	P		16	240	238	99.2	buried [27]	ATP-binding (necessary for p23 function) [11]
IV	12	361–370	10	150	146	97.3	N/A	ATP-binding [7]
	P		10	150	150	100.0		
V	12	387–401	15	225	213	94.7	N/A	Steroid receptor-binding
	P		15	225	225	100.0		[6; ref. therein]
A	12	72–92	20	300	145	48.3	coil,	N-terminal chaperone
	P		20	300	231	77.0	exposed [27]	function (α) [ref. In 6]
B	12	230–280	51	765	216	28.2	helix, acidic	Regulates nucleotide access
	P		51	765	527	68.9	protruding [ref. In 22]	to C-terminus [20] Increases binding affinity of non-native proteins to N-terminal domain [26] Regulates chaperone activity via interplay between peptide/ATP-binding [26]
C	12	560–570	11	165	67	40.6	helix/coil,	Steroid receptor-binding
	P		11	165	121	73.3	charged [ref. In 22]	[6; ref. therein]
D	12	700–717	19	285	64	22.5	coil,	TPR-binding such as
	P		19	285	193	67.7	acidic [6]	immunophilins/hop [6]
a	12	289–311	23	345	249	72.2	β-turn/coil,	Steroid receptor-binding
	P		23	345	334	96.8	protruding [22; ref. therein]	[6; ref. therein]
b	12	523–555	33	495	340	68.7	coil,	ATP-binding [7]
	P		33	495	487	98.4	charged [22; ref. therein]	Heat shock response [18]
c	12	622–634	13	195	92	47.2	coil/helix,	TPR-binding such as
	P		13	195	171	87.7	hydrophilic [6]	immunophilins/hop [6]
d	12	686–699	14	210	93	44.3	coil/helix,	TPR-binding such as
	P		14	210	210	100.0	acidic [6]	immunophilins/hop [6]

Regions c and d map to the TPR-binding domain [6; references therein]. As all 4 regions form structures that are likely to interact directly with substrates [6, 22], the high conservation in these areas suggests that: 1) they are involved in activities critical in and specific to vertebrate cells, and 2) the substrates that Hsp90 interacts with in regions a-d, and/or the functional mechanisms that are potentially attributed to this area are similar in α and β isoforms.

To facilitate the design of studies geared towards deciphering Hsp90 α and β functional differences, we identified 27 amino acid sites that are conserved within, but differ between α and β isoforms (numbered in Fig. 2). Of these, T_{498}, L_{516}, and Q_{560} for α; and Q_{469}, M_{561}, and I_{715} for β, are unique (determined through comparison with the alignments in [13]) amino acid substitutions thus indicating sites that are likely critical to vertebrate/isoform-specific function. The highest density of α/β amino acid discrepancies (sites 15–19 in Fig. 2) are located in a Ca^{+2}-calmodulin-binding domain [19]. Since calmodulin mediates numerous physiological effects induced by Ca^{+2} concentration, it is plausible that α and β isoforms have evolved to accommodate different calmodulin activities thus enabling vertebrate cells more efficient means of stress response or physiologi-

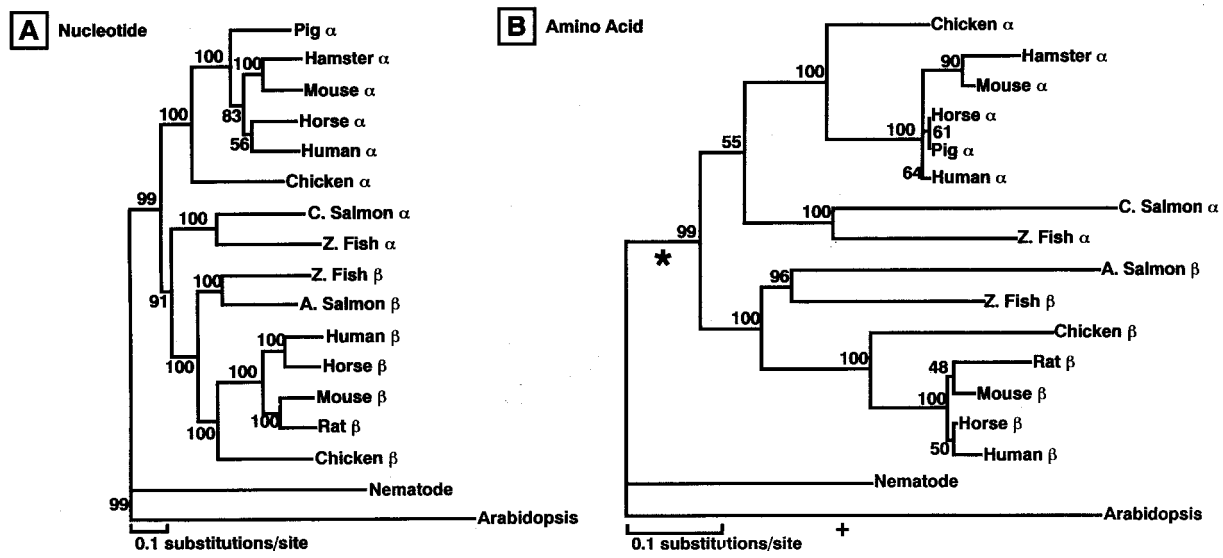


Fig. 3. Phylogenetic trees generated by the Neighbor-joining distance method. The numbers at the nodes indicate the % bootstrap support for the arrangement internal to a given node. The scale represents the number of substitutions per site. Panels A and B are based on DNA and amino acid sequence data sets of 1,000 and 500 bootstrap replicates, respectively. * and + indicate branch reductions to 25% and 75% their original values.

cal maintenance. A second area where α and *β* sequences show high frequency amino acid discrepancy (sites 6–14, Fig. 2) was between positions 415–498 in the steroid receptor domain. These positions might be good research targets to deciphering differences in the types of steroid receptors that bind α and *β* isoforms. Lastly, it has been found that while Hsp90α readily dimerizes, Hsp90*β* remains monomeric, and that the amino acid region responsible for the lack of *β* dimerization is between positions 561–685 [22]. Therefore, α/*β* differences (*ie.* sites 20–24 in Fig. 2) are likely keys in defining dimerization ability.

The neighbor-joining distance method was used for phylogenetic analyses of vertebrate *Hsp90* α and *β* sequences (including our horse sequences). In the DNA tree, chicken α and *β* genes branch outside their mammalian counterparts significantly as expected (Fig. 3A: α=100, *β*=100). Surprisingly, the horse and human (α) relationship to the rodent branch (Fig. 3A: 100) can not be distinguished (Fig. 3A: 56), while pig α significantly branches outside all mammal α genes (Fig. 3A: 83). In contrast, the protein-based tree shows no significant branching order of horse/human/pig α (Fig. 3B) relative to the strongly supported rodent branch (Fig. 3B: 90). This suggests two conditions: 1) high functional conservation of pig α to other mammal α proteins, and 2) functional diversification of rodent α from the mammal α proteins examined. Discrepancies between the DNA and protein trees were also observed for mammal *β* sequences. In the DNA trees, horse/human *β* and mouse/rat *β* branch together (Fig. 3A: 100 for both) but these relationships materialize only roughly half these values in the protein trees. The unsupported branching of horse/human/pig α and mammal *β* proteins suggests high functional similarity between these proteins, and highlights the importance of each isoform's structure. However, more mammal sequence data is needed to thoroughly evaluate the possibility of functional diversification within the α lineage.

The observed differences between nucleic acid and amino acid substitution levels (Fig. 3A and B), within and between the most closely branching genes/proteins, suggest that selection has guided the amino acid sequence changes. For example, the quite higher level of amino acid substitution shown in hamster α and rat *β* Fig. 3B), but rather similar levels of nucleotide substitution between mammal α and *β* genes (Fig. 3A), indicates that these genes have been positively selected. Recalling that hamster α and rat *β* may be quite functionally conserved to other mammal α and *β* genes, respectively, it follows that the relatively high levels of amino acid substitution indicate selection of a novel function and/or functional versatility. In support of this idea, the rat *β* sequence differs almost entirely from the other vertebrate Hsp90 sequences between positions 672–682 (see small box in Fig. 2). As point mutations D->A at 679 and E->A at 680 have been found to result in improper protein folding [23], and this region has been found to be critical to immunophilin and hop binding [6], the highly divergent rat *β* sequence suggests a different protein architecture in this region which could accommodate a novel function and/or a different role in TPR-binding, if any.

In contrast to the divergent hamster α and rat *β* sequences, the horse amino acid sequences show far less change relative to the other *β* genes, indicating that horse *β* has been subject to strong purifying selection. Remarkably, it must not be overlooked that the total number of amino acid pairwise differences between the horse/human/pig α genes is

quite low (pig-horse=2, human-horse=5, pig-human=3). This suggests that the changes are likely due to coevolution with, or adaptation to, each protein's respective target proteins. In support of this, one of the discrepancies in each the horse and human sequences were alterations unique (including the data in [13]) to each species (horse $\alpha = L_{441}$, where all other Hsp90= F; human $\alpha = T_{63}$, where all other Hsp90= S or A). The possibly different-acting selective forces highlight a discrepancy between organisms and isoforms in the biological relevance level of a given Hsp90 function.

Vertebrate *Hsp90* α and β DNA sequences displayed distinct α and β gene clusters for mammal and chicken *Hsp90* genes (Fig. 3A: α= 100, β= 100). The clustering strongly supports separate α and β lineages which are postulated to have arisen by gene duplication [13], a process thought to contribute significantly to the generation of novel genes. However, among fish species, both α and β genes clustered significantly with the β lineage (Fig. 3A: α= 91, β= 100). The high similarity of fish α genes to the β lineage suggests that the gene duplication event which may have yielded the α and β lineages (as described in [13]), occurred in a vertebrate common ancestor quite recent to fish. Since DNA sequences are subject to all types of evolutionary pressures, and amino acid sequences are predominately determined by functional selection, DNA sequences are likely to be more indicative of temporal relationships between genes. Therefore, following the logic that separate α and β isoforms of *Hsp90* are relatively recent adjustments to fish physiology, the clustering of the fish α genes with the β lineage suggests that the β genes represent the original *Hsp90* sequences, and the α genes arose via gene duplication. In support of this, if the converse were true (*ie.* fish α evolved divergently from other α genes, but convergently with the β lineage following duplication of the α gene), it is expected that the fish α genes show high levels of nucleic acid substitution relative to other α genes, and/or high amino acid similarity to the β lineage. However, those predictions are not met in the phylogenetic analyses. The DNA tree illustrates similar levels of base substitution between vertebrate α and β genes, and the branching pattern of the amino acid tree is unstable for the fish α, mammal/chicken α, and fish/mammal/chicken β branches (Fig. 3B: 55). This unresolved branching arrangement suggests some functional conservation between these proteins. Therefore, it is plausible that: 1) the fish β gene was duplicated to yield the α gene, and 2) while maintaining some functional similarity to the β protein, the fish α protein acquired novel functions that were the templates for the selection of α protein functions.

While Hsp90 is thought to play a critical role in mammalian physiology, little mammalian sequence data exists. This results in more roughly designed experiments thus hampering progress towards identification of specific Hsp90 functions. By sequencing horse Hsp90, and comparing our data to other Hsp90 sequences, we identified potentially critical amino acid sites to organism- and isoform-specific functions. Furthermore, phylogenetic analysis revealed high protein conservation, but differing levels of amino acid substitution between mammalian Hsp90s, which, given the protein's versatile nature, suggests functional diversification. This highlights the necessity of organism-specific functional studies of Hsp90 to accurately describe its biological relevance.

ACKNOWLEDGEMENTS. This work was carried out on the basis of the Collaborative Research Grant founded by JRA, and supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan (K.N.), and from Research Fellowships of the Japan Society for the promotion of Science for Young Scientists (F.M.).

REFERENCES

1. Aligue, R., Akhavan-Niak, H. and Russell, P. 1994. A role for Hsp90 in cell cycle control: Wee1 tyrosine kinase activity requires interaction with Hsp90. *EMBO J.* 13: 6099–6106.
2. Benjamin, I.J. and McMillan, D.R. 1998. Stress (heat shock) proteins: molecular chaperones in cardiovascular biology and disease. *Circ. Res.* 83: 117–132.
3. Bharadwaj, S., Ali, A. and Ovsenek, N. 1999. Multiple components of the Hsp90 chaperone complex function in regulation of heat shock factor 1 *In Vivo*. *Mol. Cell. Biol.* 19: 8033–8041.
4. Bresnick, E.H., Dalman, F.C., Sanchez, E.R. and Pratt, W.B. 1989. Evidence that the 90-kDa heat shock protein is necessary for the steroid binding conformation of the L cell glucocorticoid receptor. *J. Biol. Chem.* 264: 4992–4997.
5. Buchner, J. 1996. Supervising the fold: functional principles of molecular chaperones. *FASEB J.* 10: 10–19.
6. Carrello, A., Ingley, E., Minchin, R.F., Tsai, S. and Ratajczak, T. 1999. The common tetratricopeptide repeat acceptor site for steroid receptor-associated immunophilins and hop is located in the dimerization domain of Hsp90. *J. Biol. Chem.* 274: 2682–2689.
7. Csermely, P. and Kahn, C.R. 1991. The 90-kDa heat shock protein (hsp-90) possesses an ATP binding site and autophosphorylating activity. *J. Biol. Chem.* 266: 4943–4950.
8. Dittmar, K. D., Demady, D. R., Stancato, L. F., Krishna, P. and Pratt, W. B. 1997. Folding of the glucocorticoid receptor by the heat shock protein (Hsp) 90-based chaperone machinery: The role of p23 is to stabilize receptor-Hsp90 heterocomplexes formed Hsp90-p60-Hsp70. *J. Biol. Chem.* 272: 21213–21220.
9. Felsenstein, J. PHYLIP. Version 3.573.2000. Department of Genetics. University of Washington.
10. Freeman, B.C. and Morimoto, R.I. 1996. The human cytosolic molecular chaperones Hsp90, Hsp70 (Hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J.* 15: 2969–2979.
11. Grenert, J.P., Sullivan, W.P., Fadden, P., Haystead, T.A.J., Clark, J., Mimnaugh, E.., Krutzsch, H., Ochel, H., Schulte, T.W., Sausville, E. Neckers, L.M. and Toft, D.O. 1997. The amino-terminal domain of heat shock protein 90 (Hsp90) that binds Geldanamycin is an ATP/ADP switch domain that regulates Hsp90 conformation. *J. Biol. Chem.* 272: 23843–23850.
12. Guerriero, V.Jr. and Raynes, D.A. 1990. Synthesis of heat stress proteins in lymphocytes from liverstock. *J. Anim. Sci.* 68: 2779–2783.
13. Gupta, R.S. 1995. Phylogenetic analysis of the 90 kD Heat Shock family of protein sequences and a final examination of

the relationship among animals, plants, and fungi species. *Mol. Biol. Evol.* 12: 1063–1073.
14. Hu, J. and Seeger, C. 1996. Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* 93: 1060–1064.
15. Johnson, J.L. and Craig, E.A. 1997. Protein folding *in vivo*: Unraveling complex pathways. *Cell* 90: 201–204.
16. Kosano, H., Stensgard, B., Charlesworth, M. C., McMahon, N. and Toft, D. 1998. The assembly of progesterone receptor-Hsp90 complexes using purified proteins. *J. Biol. Chem.* 273: 32973–32979.
17. Lange, B. M.H., Bachi, A., Wilm, M. and González, C. 2000. Hsp90 is a core centrosomal component and is required at different stages of the centrosome cycle in Drosophila and vertebrates. *EMBO J.* 19: 1252–1262.
18. Louvion, J., Warth, R. and Picard, D. 1996. Two eukaryote-specific regions of Hsp82 are dispensable for its viability and signal transduction functions in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 93: 13937–13942.
19. Minami, Y, Kawasaki, H., Suzuki, K. and Yahara, I. 1993. The calmodulin-binding domain of the mouse 90-kDa heat shock protein. *J. Biol. Chem.* 268: 9604–9610.
20. Monica, G., Marcu, M.G., Chadli, A., Bouhouche, I., Catelli, M. and Neckers, L.M. 2000. The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the C-terminus of the chaperone. *J. Biol. Chem.* 275: 37181–27186.
21. Morimoto, R.I., Sarge, K.D. and Abravaya, K. 1992. Transcriptional regulation of heat shock genes: A paradigm for inducible genomic responses. *J. Biol. Chem.* 267: 21987–21990.
22. Nemoto, T., Ohara-Nemoto, Y., Ota, M., Takagi, T. and Yokoyama, K. 1995. Mechanism of dimer formation of the 90-kDa heat-shock protein. *Eur. J. Biochem.* 233: 1–8.
23. Ramsey, A.J., Russell, L..C., Whitt, S.R. and Chinkers, M. 2000. Overlapping sites of tetratricopeptide repeat protein binding and chaperone activity in heat shock protein 90. *J. Biol. Chem.* 275: 17857–17862.
24. Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Extraction and purification of RNA. pp. 7.3–7.36. *In*: Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor Laboratory Press, New York.
25. Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Sequencing techniques and strategies. pp. 13.3–13.6. *In*: Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor Laboratory Press, New York.
26. Scheibel, T., Siegmund, H.I., Jaenicke, R., Ganz, P., Lilie, H. and Buchner, J. 1999. The charged region of Hsp90 modulates the function of the N-terminal domain. *Proc. Natl. Acad. Sci. U.S.A.* 96: 1297–1302.
27. Stebbins, C.E., Russo, A.A., Schneider, C., Rosen, N., Hartl, F.U. and Pavletich, N.P. 1997. Crystal structure of an Hsp90-Geldanamycin complex: Targeting of a protein chaperone by an antitumor agent. *Cell* 89: 239–250.
28. Stephanou, A., Isenberg, D.A., Akira, S., Kishimoto, T. and Latchman, D. S. 1998. The nuclear factor interleukin-6 (NF-IL6) and signal transducer and activator of transcription-3 (STAT-3) signalling pathways co-operate to mediate the activation of the Hsp90β gene by interleukin-6 but have opposite effects on its inducibility by heat shock. *Biochem. J.* 330: 189–195.
29. Stephanou, A., Isenberg, D.A., Nakajima, K. and Latchman, D.S. 1999. Signal transducer and activator of transcription-1 and heat shock factor-1 interact and activate the transcription of the Hsp-70 and Hsp-90β gene promoters. *J. Biol. Chem.* 274: 1723–1728.
30. Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
31. Yonehara, M., Minami, Y., Kawata, Y., Nagai, J. and Yahara , I. 1996. Heat-induced chaperone activity of Hsp90. *J. Biol. Chem.* 271: 2641–2645.