

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Genomics 88 (2006) 230–240

GENOMICS

www.elsevier.com/locate/ygeno

Diversity, structure, and expression of the gene for p26, a small heat shock protein from *Artemia*[☆]

Zhijun Qiu^a, Peter Bossier^b, Xiaomei Wang^{b,c,d},
Svetla Bojikova-Fournier^a, Thomas H. MacRae^{a,*}

^a Department of Biology, Dalhousie University, Halifax, Canada NS B3H 4J1

^b Laboratory of Aquaculture and Artemia Reference Center, Faculty of Bioscience Engineering, Ghent University, Rozier 44, 9000 Ghent, Belgium

^c CLO Sea Fisheries Department, Ankerstraat 1, 8400 Oostende, Belgium

^d Department of Fishery Science, Tianjin Agriculture College, Jinjing Road, Xiqing District, Tianjin, People's Republic of China

Received 14 December 2005; accepted 14 February 2006

Available online 29 March 2006

Abstract

p26, a small heat shock protein, is thought to protect *Artemia* embryos from stress during encystment and diapause. Full-length p26 cDNAs were compared and used to determine phylogenetic relationships between several *Artemia* species. The α -crystallin domain of p26 was the most conserved region of the protein and p26 from each *Artemia* species contained characteristic amino-terminal WD/EPF and carboxy-terminal VPI motifs. Sequence conservation suggested the importance of p26 to oviparously developing *Artemia* embryos and indicated common functions for the protein during development and stress resistance, although as shown by modeling some species-specific p26 amino acid substitutions may have adaptive significance. The p26 gene obtained from *A. franciscana* exhibited a unique sHSP intron arrangement with an intron in the 5'-untranslated region. Computer-assisted analysis revealed heat shock elements and other putative *cis* regulatory sequences but their role in gene regulation is unknown. In contrast to previous results for which Northern blots were analyzed, p26 gene expression was observed in ovoviparous embryos by use of PCR-based methodology, but the p26 protein was not detected.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Small heat shock protein genes; Phylogenetic relationships; Small heat shock proteins; Embryo development; Encystment and diapause; *Artemia*

Oviparous development within various *Artemia* species, a process best characterized for *A. franciscana*, yields encysted gastrulae (cysts) that enter diapause, a resting stage at which development stops, metabolism decreases greatly, and stress resistance is high [1]. Cysts tolerate temperature extremes, desiccation, and several years of anoxia with low mortality [2]. This remarkable resistance is thought to depend on p26, a small heat shock protein (sHSP) constituting approximately 10% of cyst nonyolk protein and for which the amino acid sequence has been determined [3–5]. p26 is a molecular chaperone that promotes cell survival by preventing irreversible protein

aggregation [3,6–8] and inhibiting apoptosis [9]; however, synthesis of the protein is developmentally regulated and does not occur in response to stress [6,10]. p26 migrates from the cytoplasm into nuclei [5–8,11,12], suggesting activities in addition to chaperoning and apoptosis delay.

sHSP monomers, composed of a well-conserved α -crystallin domain of approximately 90 amino acid residues bordered by variable amino- and carboxy-terminal domains, assemble into dynamic oligomers. The α -crystallin domain mediates formation of dimers, fundamental units of oligomerization for many sHSPs. The amino-terminus influences substrate binding and oligomer dynamics, whereas the highly flexible and charged carboxy-terminus enhances solubility and also modulates oligomerization [13,14]. sHSPs represent an early line of defense against stress within cells, binding partially denatured proteins by energy-independent processes, preventing irreversible denaturation, and, in cooperation with

[☆] Sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. DQ310575, DQ310576, DQ310577, DQ310578, DQ310579, and DQ310580.

* Corresponding author. Fax: +1 902 494 3736.

E-mail address: tmacrae@dal.ca (T.H. MacRae).

other molecular chaperones, promoting either protein renaturation or protein destruction [15,16]. Many sHSP genes are induced by stress but others are expressed constitutively or in response to developmental signals and aging [17–20].

Despite the potential importance of p26 during *Artemia* embryo development and stress resistance, the sHSP has been characterized extensively only in *A. franciscana* and the p26 gene had not been examined in any *Artemia* species prior to this work. To achieve a better understanding of p26, which has broad implications for sHSP function and crustacean development, cDNAs for the protein were cloned from five *Artemia* species, including those geographically well isolated. Deduced p26 amino acid sequences were compared and the potential impact of residue substitutions was determined by protein modeling. The *A. franciscana* p26 gene was sequenced, revealing a novel sHSP intron organization and putative regulatory regions including heat shock elements. In contrast to previous results [6], p26 gene expression was detected during both oviparous and ovoviviparous development, a finding attributed to the use of sensitive PCR methodology in this study, but p26 protein was observed only in the former.

Results

p26 cDNA and protein sequences

With stop codons included, the open reading frames for p26 cDNAs from the two *A. franciscana* strains and *A. sinica* were 579 nucleotides, *A. urmiana* p26 cDNA consisted of 576 nucleotides, *A. parthenogenetica* contained 585 nucleotides, and *A. persimilis* p26 cDNA was 588 nucleotides (Fig. 1A, accession numbers in figure legend). Comparison of nucleotide sequences demonstrated varying degrees of similarity between p26 cDNAs from disparate *Artemia* species with *A. persimilis* exhibiting the most variation (Fig. 1A, Table 1). Phylogenetic trees constructed on the basis of p26 cDNA nucleotide sequences, of which an example is given (Fig. 1B), showed that the *A. franciscana* populations are closely related. *A. urmiana* and *A. parthenogenetica* were linked and more similar to *A. sinica* than to other *Artemia* species. *A. persimilis* was the most divergent and it was approximately equal distance from all other *Artemia* species.

Alignment of amino acid sequences deduced from p26 cDNAs indicated that some nucleotide substitutions led to residue changes (Fig. 2A, Table 2). The carboxy-terminal extension contained the most modifications, with 25.0% (10/40) of residues differing across species boundaries, followed by the amino-terminal region with 16.9% (10/59) and the α -crystallin domain with 6.5% (6/93). Y18 in the p26 sequence 17-WYDPF-21 of *A. franciscana* from the GSL (ARC1520), representing a widely conserved motif, was S18 in all other *Artemia* species. This substitution introduces a large bulky amino acid in place of a smaller residue, although both are potential phosphorylation sites. Other modifications include the loss or gain of several potential phosphorylation sites, the loss of a negatively charged residue at position 35 in *A. parthenogenetica* (ARC1407) p26, and the short insertions 184-STI-186 and 185-STR-187 in the carboxy-terminal exten-

sions of *A. parthenogenetica* (ARC1407) and *A. persimilis* (ARC1321), respectively. R114 in the α -crystallin domain of p26 from the *Artemia* species examined is equivalent to a highly conserved arginine found in sHSPs from other organisms, and the conserved carboxy-terminal extension V/IPI/V motif as 154-VPI-156 occurs in these *Artemia* species.

To localize amino acid substitutions spatially within the p26 tertiary structure and thus evaluate their effects more fully, residues 58–157 of each *Artemia* sHSP were modeled in three dimensions on the basis of the corresponding region in wheat Hsp16.9, a crystallized sHSP [21]. Sequence identity between p26 from each *Artemia* species and Hsp16.9 within this region was approximately 69%, and p26 consisted of β -strands 2–9 with intervening loops named for the β -strands they connect (Figs. 2B–2E). For example, L2/3 connects β -strands 2 and 3. β -strand 6 is positioned in L5/7 and is not labeled in Fig. 2B. Primary structures of p26 from *A. franciscana* GSL (ARC1520), *A. franciscana* SFB (ARC1258), and *A. sinica* (ARC1218) are identical in the compared region and they generated a single model (Fig. 2B). p26 proteins from the remaining species, although similar, have substitutions in different locations (Figs. 2C–2E). p26 from *A. urmiana* (ARC1511) contains substitutions A66D within L2/3 and T150K in a region of the carboxy-terminus not modeled (Fig. 2C). These modifications would be adjacent if the carboxy-tail folds back onto the β -sheet domain, allowing formation of a salt bridge not found in p26 from the remaining *Artemia* species. On the basis of the sequences compared in this study, substitution D90Y of p26 from *A. parthenogenetica* (ARC1407) appears in L4/5, at the interface with β -strand 5, and substitution G106V is located in β -strand 6 of L5/7, the latter loop potentially with a role in dimer formation. The D90Y modification reduces the charge in loop L4/5 and may influence protein structure, whereas the G106V substitution probably has limited impact on p26 structure. The changes in *A. parthenogenetica* p26 are well separated spatially and not likely to interact with one another (Fig. 2D). In *A. persimilis* (ARC1321), β -strand 2, potentially involved in p26 dimer formation, has adjacent amino acid modifications. The G60W alteration introduces a bulky amino acid residue, whereas S61T is structurally less intrusive and preserves a potential phosphorylation site. A third substitution in *A. persimilis* p26, G96C in β -strand 5, introduces cysteine into the protein, thus opening the possibility of disulfide bridge formation between neighboring monomers during dimer formation and higher order oligomerization.

The *A. franciscana* GSL p26 gene

To provide the basis for comparison across species boundaries the *A. franciscana* p26 gene was cloned and sequenced, work which also revealed regulatory aspects of the gene's expression. The titer of the genomic library prepared from *A. franciscana* GSL (ARC1520) was 1.2×10^7 plaque forming units/ml and 10 clones picked randomly from the library contained inserts of 15 to 23 kb. Screening approximately 4.4×10^6 plaques with labeled full-length p26 cDNA yielded four reactive clones termed 12-2, 7-2, 3-1, and 13-1

A

```

1520 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGACTGACCCATG 50
1258 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGACTGACCCATG 50
1218 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGGCGACCCATG 50
1511 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGAGCGACCCATG 50
1407 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGAGCGACCCATG 50
1321 ATGGCACTTAACCCATGGTACGGAGGATTTGGTGGTATGACTGATCCGTG 50
*****

1520 GTATGATCCATTTGGATTTGGTGGCTTCGGAGGTGGCATGGACCTTGATA 100
1258 GTCTGATCCATTTGGATTTGGTGGCTTCGGAGGTGGCATGGACCTTGATA 100
1218 GTCTGATCCGTTTGGATTTGGTGGCATCGGAGGTGGCATGGACCTCGACA 100
1511 GTCTGATCCATTTGGATTTGGTGGCTTCGGAGGTGGCATGGACCTCGACA 100
1407 GTCTGATCCATTTGGATTTGGTGGCTTCGGAGGTGGCATGGACCTCGACA 100
1321 GTCAGATCCATTTGGATTTGGTGGCTTCGGAGGTGGTATGGACCTTGACA 100
*****

1520 TTGACAGGCCCTTCGGAGAAGAATGATGAGAAGAGTCCAGATACCAGC 150
1258 TTGACAGGCCCTTCGGAGAAGAATGATGAGAAGAGTCCAGATACCAGC 150
1218 TTGACAGGCCCTTCGGAGAAGAATGATGAGAAGAGTCCAGATACCAGT 150
1511 TTGACAGGCCCTTCGGAGAAGAATGATGAGAAGAGTCCGACACCAGT 150
1407 TTACAGGCCCTTCGGAGAAGAATGATGAGAAGAGTCCGACACCAGT 150
1321 TTGATAGACCTTCAGGAGAAGAATGATGAGAAGAGTCCAGATACCAGT 150
*****

1520 AGGGCTTTAAAGGAGTTAGCTACTCCTGGGTCTTGAGGGACACAGCTGA 200
1258 AGGGCTTTAAAGGAGTTAGCTACTCCTGGGTCTTGAGGGACACAGCTGA 200
1218 AGGGCTTTAAAGGAGTACTACTCCTGGGTCTTGAGGGACACAGCTGA 200
1511 AGGGCTATAAAGAGCTAGCTACTCCTGGGTCTTGAGGGACACAGCTGA 200
1407 AGGGCTTTAAAGGAGTACTACTCCTGGGTCTTGAGGGACACAGCTGA 200
1321 CGGGCTATAAAGGAGTTACTACTCCTGGGTCTTGAGGGACACAGCTGA 200
*****

1520 TGAATTTCAAGTTCAGCTAGATGTTGGCCACTTTTACCAAACGAAATTA 250
1258 TGAATTTCAAGTTCAGCTAGATGTTGGCCACTTTTACCAAACGAAATTA 250
1218 TGAATTTCAAGTTCAGCTAGATGTTGGTCACTTTTACCAAATGAAATTA 250
1511 TGAATTTCAAGTTCAGCTAGATGTTGGTCACTTTTACCAAATGAAATTA 250
1407 TGAATTTCAAGTTCAGCTAGATGTTGGTCACTTTTACCAAATGAAATTA 250
1321 TGAATTTCAAGTTCAGCTAGATGTTGGCCACTTTTACCAAATGAGATTA 250
*****

1520 CAGTCAAGACAACCGACGATGATATCTTGTCCATGGCAAACATGACGAG 300
1258 CAGTCAAGACAACCGACGATGATATCTTGTCCATGGCAAACATGACGAG 300
1218 CAGTCAAGACCAACCGACGATGATATCTTGTCCATGGCAAACATGACGAG 300
1511 CAGTCAAGACCACTGACGATGATATCTTGTCCATGGCAAACATGACGAG 300
1407 CAGTCAAGACCACTGACTATGATATCTTGTCCATGGCAAACATGACGAG 300
1321 CAGTCAAGACCACTGATGACGATATCTTGTTCATGGCAAACATGACGAG 300
*****

1520 CGATCTGATGAATATGGACACGTCCAAAGAGAATTCGACGACGATACAG 350
1258 CGATCTGATGAATATGGACACGTCCAAAGAGAATTCGACGACGATACAG 350
1218 CGGTCCGATGAATATGGACACGTCCAAAGAGAATTCGACGACGATACAG 350
1511 CGGTCCGATGAATATGGACACGTCCAAAGAGAATTCGACGACGATACAG 350
1407 CGGTCCGATGAATATGTACACGTCCAAAGAGAATTCGACGACGATACAG 350
1321 CGGTCCGATGAATATGGACACGTCCAAAGAGAATTCGACGACGATACAG 350
*****

1520 ACTCCAGAACAATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
1258 ACTCCAGAACAATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
1218 ACTCCAGAACAATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
1511 ACTCCAGAACAATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
1407 ACTCCAGAACAATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
1321 ACTACCAGACATGTCAAACCGAATCTGTGTCATCTACTTTGTCAATCAG 400
*****

1520 ATGGTGTCTTAACTATCCATGCTCCGAAAACCTGCTTTAAGCTCACCACA 450
1258 ATGGTGTCTTAACTATCCATGCTCCGAAAACCTGCTTTAAGCTCACCACA 450
1218 ATGGTGTCTTAACTATCCATGCTCCGAAAACCTGCTTTAAGCTCACCACG 450
1511 ATGGTGTCTTAACTATCCATGCTCCGAAAACCTGCTTTGAGCTCACCACAA 450
1407 ATGGGCTTAACTATCCATGCTCCGAAAACCTGCTTTGAGCTCACCACA 450
1321 ATGGTGTCTTAACTATACATGCTCCGAAAACCTGCTTTGAGCTCACCACG 450
*****

1520 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
1258 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
1218 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
1511 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
1407 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
1321 GAACGTATCGTACCCATCACACCAGCGCCAGCTGTTGGAAGGATTGAAGG 500
*****

1520 GGAAGTACAGGTACTACTACAGGCAGTACAGCTAGTTCAACTCAGCAA 550
1258 GGAAGTACAGGTACTACTACAGGCAGTACAGCTAGTTCAACTCAGCAA 550
1218 GGCACGTCCGGTACTACTACAGGCAGTACAGCTAGTTCAACTCAGCAA 550
1511 GGCACGTACCGGTACTAC---AGGCAGTACAGCTAGTTCAACTCAGCAA 547
1407 GGCACGTACCGGTACTAC---AGGCAGTACAGCTAGTTCAACTCAGCAA 547
1321 TGAACGACGGGTCTACTGACGGTACTACAGCTAGTTCAACTCAGCAA 550
*****

1520 GAACAACAA-----GATCAGGAGGTGCAGCTTAA 579
1258 GAACAACAA-----GATCAGGAGGTGCAGCTTAA 579
1218 GATCAACAA-----GATCAGGAGGTGCAGCTTAA 579
1511 AATCAAAAA-----GATCAGGAGGTGCAGCTTAA 576
1407 GATCAACAATATCAACAAGATCAGGAGGTGCAGCTTAA 585
1321 GATCAACAAGATCAACAAGATCAGGAGGTGCAGCTTAA 588
*****

```

B

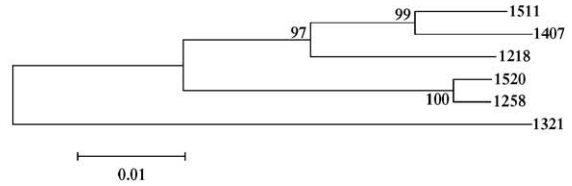


Table 1
Comparison of p26 cDNA sequences

	<i>A. franciscana</i> (GSL) (%)	<i>A. franciscana</i> (SFB) (%)	<i>A. sinica</i> (%)	<i>A. urmiana</i> (%)	<i>A. parthenogenetica</i> (%)	<i>A. persimilis</i> (%)
<i>A. franciscana</i> (GSL) (ARC1520)	100					
<i>A. franciscana</i> (SFB) (ARC1258)	99	100				
<i>A. sinica</i> (ARC1218)	94	94	100			
<i>A. urmiana</i> (ARC1511)	93	93	96	100		
<i>A. parthenogenetica</i> (ARC1407)	93	93	95	97	100	
<i>A. persimilis</i> (ARC1321)	90	90	90	90	90	100

p26 cDNA clones obtained from *Artemia* species by RT-PCR were sequenced and compared as described under Materials and methods.

with inserts of 17, 17.8, 19, and 15 kb, respectively (Fig. 3A). Overlapping DNA fragments from the clones yielded the p26 gene sequence, including structural regions and 5'- and 3'-noncoding areas (Figs. 3A and 3B), deposited in the NCBI database under Accession No. DQ310575. Comparison of *A. franciscana* p26 cDNA (Accession No. AF031367) with the genomic sequence revealed three introns, with intron 1 in the 5'-UTR and the insertion sites of introns 2 and 3 disrupting codons (Fig. 3B). Agarose gel electrophoresis of 5'-RACE products gave one DNA fragment, indicating a single transcription start site (Fig. 4), this removed from the ATG translation initiation codon in p26 cDNA and genomic sequences by 53 and 1348 bp, respectively (Figs. 3B and 4B). Computer-assisted analysis indicated that intron 1 contains many putative *cis*-acting elements of potential interest for p26 gene expression during development, including heat shock factors (HSFs) and AP-1 binding sites (Fig. 3B). The introns of 1295, 1261, and 1283 bp begin with GT and terminate in AG, consistent with the intron boundary rule. The exons, composed of 47, 113, 225, and 333 bp do not correspond to p26 domain structure (Fig. 5), nor are they consistent in position and number to intron/exon locations in other sHSP genes deposited in the NCBI database. Approximately 5 kb of noncoding sequence upstream of the p26 transcription start site was obtained and of this, approximately 1.0 kb is shown, revealing TATA and CCAAT boxes 22 and 392 bp upstream, respectively, from the transcription start site (Fig. 6). Representative transcription factor binding sites include those for HSFs and the estrogen receptor.

p26 gene expression in artemia embryos

As determined by real-time PCR, p26 mRNA first appeared in oviparously developing *Artemia* embryos 2 days postfertilization and increased until cyst release from females, whereas p26 mRNA was observed in ovoviviparous embryos at 2 and 3 days postfertilization only (Fig. 7). Oviparous embryos possessed approximately 10 times more p26 mRNA than ovoviviparous embryos at 3 days postfertilization. A strong p26 band was observed on Northern blots containing mRNA from

oviparous embryos 2 days postfertilization, while similar amounts of mRNA from ovoviviparous embryos yielded a weak band (Fig. 7, inset). p26 was detected on Western blots containing protein extracts from oviparous but not ovoviviparous embryos as reported previously [6].

Discussion

Artemia species phylogeny was inferred previously by analysis of allozymes, RAPD, and AFLP [22,23], but to our knowledge this article contains the first inference of phylogenetic relationship between *Artemia* species based on cDNA sequences derived from a nuclear protein-encoding gene. All approaches demonstrate a strong relationship between *A. urmiana* and *A. parthenogenetica*, with *A. sinica* most closely associated with these species. Moreover, the New World species *A. franciscana* and *A. persimilis* are distinct from Old World species and the genetic distance between *A. franciscana* and *A. persimilis* equals the distance between New and Old World *Artemia* species as shown before.

Different *Artemia* species contain comparable amounts of p26 [5] and as demonstrated here p26 amino acid sequences are similar from species to species with few suggestions of significant changes in protein structure and hence in function. The α -crystallin domain is modified at only 6 of 93 residues in all *Artemia* species examined and position 114 of p26 is occupied by arginine, a highly conserved sHSP residue, which, when mutated in human sHSPs, leads to reduced chaperone activity and disease [14,24]. p26 from all *Artemia* species investigated contained the conserved carboxy-terminal motif 154-VPI-156, thought to promote oligomer stability by interacting with a hydrophobic fold between β -strands 4 and 8 in the α -crystallin domain of neighboring monomers [21]. The largest modifications in primary sequence were insertions of 3 residues in the carboxy-terminal extensions of p26 from the geographically separated species *A. parthenogenetica* and *A. persimilis*. The insertions increased the hydrophilic character of the carboxy-terminal extensions, thus promoting sHSP solubility, and they may stabilize quaternary structure by increasing the

Fig. 1. p26 cDNAs from different *Artemia* species. (A) Nucleotide sequences of p26 cDNAs from *A. franciscana* (GSL) (1520) (Accession No. AF031367), *A. franciscana* (SFB) (1258) (Accession No. DQ310577), *A. sinica* (1218) (Accession No. DQ310576), *A. urmiana* (1511) (Accession No. DQ310580), *A. parthenogenetica* (1407) (Accession No. DQ310579), and *A. persimilis* (1321) (Accession No. DQ310578) were aligned by Clustal W. Nucleotide numbers are on the right. (*) Identical nucleotides; (space) nonidentical nucleotides. (B) Neighbor-joining tree constructed as described under Materials and methods. The scale represents the Kimura two-parameter distance, and bootstrap values greater than 95, indicating high confidence in the correctness of the displayed relationships, are given.

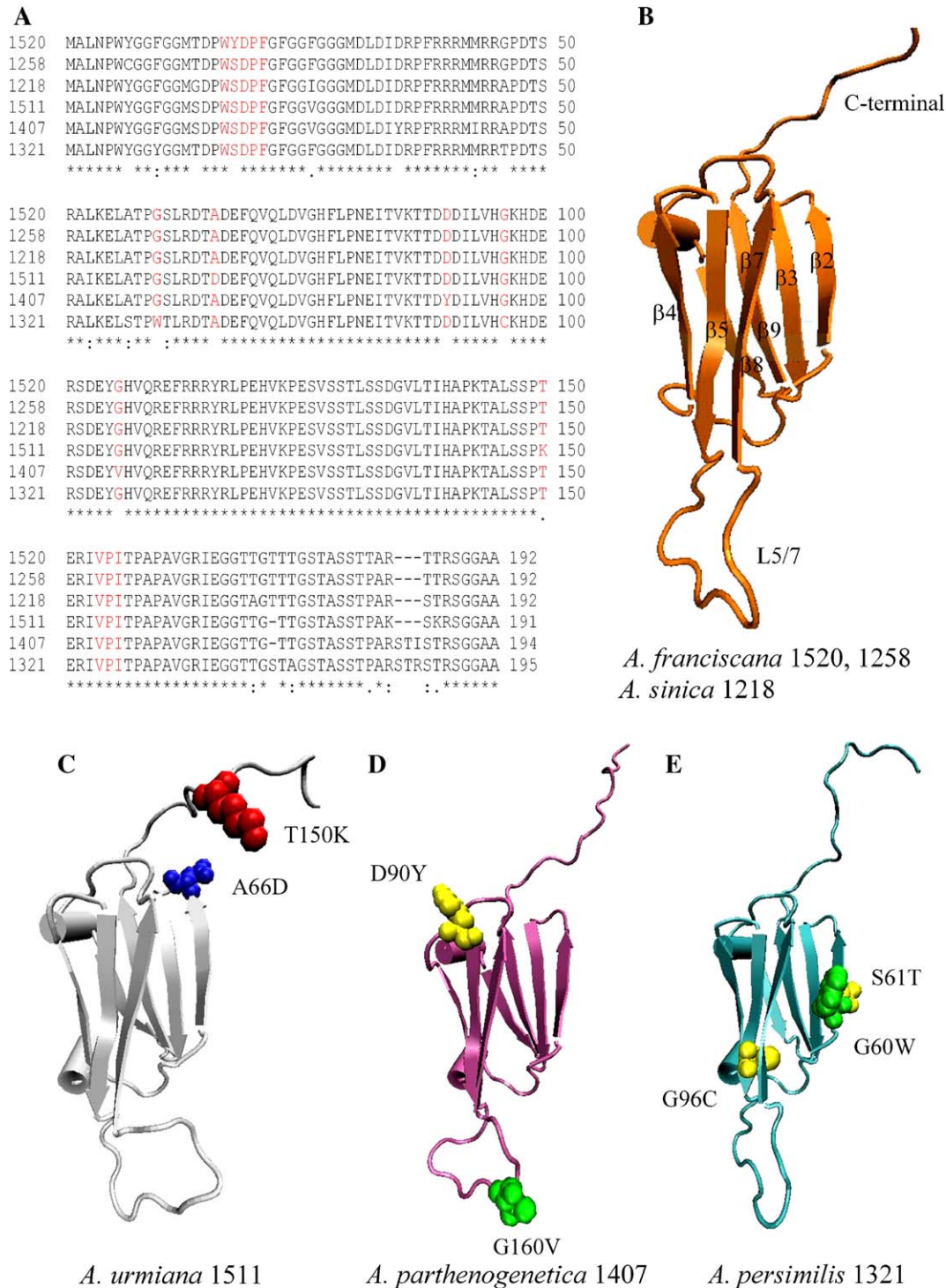


Fig. 2. p26 amino acid sequences and tertiary structures. (A) Amino acid sequences of p26 from *A. franciscana* (GSL) (1520), *A. franciscana* (SFB) (1258), *A. sinica* (1218), *A. urmiana* (1511), *A. parthenogenetica* (1407), and *A. persimilis* (1321) were aligned by Clustal W. Amino acid residue numbers are on the right. (*) Identical residues, (:) conserved substitution; (.) semiconserved substitution; (space) residue missing or nonconserved substitution. (B–E) Three-dimensional models of p26 were constructed on the basis of wheat Hsp16.9 monomer structure as described under Materials and methods. β -strand numbers and loop L5/7 are indicated in panel B. Residues are nonpolar (green), uncharged polar (yellow), acidic (blue), and basic (red).

number of residues available to interact with neighboring monomers [13,14,21]. A 3-residue insertion, Asp-Gly-Lys, occurs in the carboxy-terminal of HSP26 from natural *Drosophila* populations but its effect on structure and function are unknown [25].

Amino acid substitutions in p26 from different *Artemia* species are positioned randomly and some appear to have limited influence on the protein. On the other hand, the A66D and T150K substitutions in *A. urmiana* p26 might produce a salt bridge should the carboxy-terminal extension fold back on the

Table 2
Comparison of p26 proteins

	<i>A. franciscana</i> (GSL) (%)	<i>A. franciscana</i> (SFB) (%)	<i>A. sinica</i> (%)	<i>A. urmiana</i> (%)	<i>A. parthenogenetica</i> (%)	<i>A. persimilis</i> (%)
<i>A. franciscana</i> (GSL) (ARC1520)	100					
<i>A. franciscana</i> (SFB) (ARC1258)	98 (98)	100				
<i>A. sinica</i> (ARC1218)	96 (96)	96 (96)	100			
<i>A. urmiana</i> (ARC1511)	93 (95)	94 (96)	95 (96)	100		
<i>A. parthenogenetica</i> (ARC1407)	92 (93)	93 (94)	94 (95)	93 (94)	100	
<i>A. persimilis</i> (ARC1321)	92 (94)	93 (95)	92 (94)	92 (95)	92 (95)	100

Amino acid sequences of p26 from *Artemia* species were deduced from their corresponding cDNAs and compared as described under Materials and methods. Values not in parentheses represent amino acid identity and those in parentheses indicate similarity.

α -crystallin domain. Structure would be stabilized and the introduction of K150 into the carboxy-terminal extension also increases protein solubility. The G96C substitution in β -strand 5 of *A. persimilis* p26 could lead to disulfide bridge formation and stabilization of higher order structure. Testing these possibilities occasioned by the naturally occurring amino acid substitutions observed in this study awaits characterization of p26 either from cysts of different *Artemia* species or as bacterial expression products.

Sequencing revealed three introns in the *A. franciscana* (GSL1520) p26 gene, in agreement with previous results [4]. In comparison, plant, yeast, and many invertebrate sHSP genes lack introns. The *Drosophila* sHSP genes, with the exception of a single intron in the *l(2)efl*, are intronless [26]; however, the four *Caenorhabditis elegans* *Hsp16* genes each contain one intron coinciding in position with mammalian α -crystallin gene intron 1. The *C. elegans* *SEC-1* gene has a single intron of 56 bp that disrupts codon 93, and *Hsp12.3* has two introns, with the second approximating the mammalian α -crystallin gene intron 2 in location [18,26,27]. The p26 gene introns, two of which interrupt codons, are not positioned at protein domain interfaces, and intron 1 is in the 5'-UTR, an arrangement found in a sHSP gene from the honey bee *Apis mellifera* (Accession No. NC_007071), but no others.

sHSP genes from plants contain clusters of heat shock elements (HSEs) formed of frequently imperfect, palindromic, modular repeats of (aGAAn) and (nTTCT), often as 5'-nGAAnnTTCnnGAAn-3' [28]. Modules are considered defective if the invariant G or C and/or the two highly conserved A and T (uppercase letters) are missing [29,30]. Until now, putative *cis*-acting regulatory sites in the *Artemia* p26 gene, expressed mainly in oviparous embryos but indifferent to stress [6,10], were unidentified. HSEs composed maximally of two perfect inverted repeats were found upstream of the p26 transcription start site in this work, as were degenerative HSEs, but their roles in transcriptional regulation, as for the HSEs in intron 1, have yet to be tested. The minimum number of inverted repeats required to generate a functional HSE in *Artemia* is unknown, and this varies from one organism to another as described in the following paragraph. Thus, even though there are perfect inverted repeats within intron 1 and the 5' sequence upstream of the transcription start site of the p26 gene they may not be functional. By comparison, mammalian HSEs are generally not involved in the developmental control of sHSP genes; however, a HSE in the first intron of rat *Hsp27* may

repress transcription [31], and this could be true for HSEs in the first intron of p26. Additionally, the sunflower *Ha hsp17.6 G1* gene, which is expressed in seeds but is not heat responsive, contains a degenerative HSE distal to the TATA box involved in developmentally regulated sHSP synthesis [29,32].

Drosophila hsp22 contains three copies of the consensus heat shock element CTnGAAnTTCnAG required for stress induction, and developmental regulation is partially dependent on these sequences [33]. The *Drosophila* genes *hsp22*, *hsp23*, *hsp26*, and *hsp27* respond to ecdysterone, an arthropod, receptor-binding, steroid hormone that interacts with HERE, a β -ecdysone response element [17,34,35]. A distal estrogen receptor binding site was recognized in the p26 gene upstream sequence, this suggestive of regulation by a steroid hormone as occurs in *Drosophila* for sHSP genes. Other transcriptionally important sites in the *Drosophila hsp16* gene family are (CT)n segments, recognized by the chromatin remodeling GAGA factor [36], but these were not observed in the p26 gene. Heat induces cell-specific synthesis of *Drosophila* sHSPs [37], demonstrating the complexity of gene regulation for these chaperones. The *C. elegans hsp16* HSEs consist of three nGAAn motifs of alternating polarity with the *hsp16-1* and *hsp16-2* sequences as CtcGAAtgTTCTaGAAa and *hsp16-41* and *hsp16-48* as CtaGGAccTTCTaGAACAATTCt [18]. These genes are differentially induced by heat during development and in assorted tissues. Conversely, the *C. elegans* gene encoding sHSP SIP-1 lacks HSEs and is not induced by stress, whereas *hsp26* and *hsp43* are expressed constitutively and insensitive to stress [18,27].

As a final comparison to p26, mammalian *hsp25* and α -crystallin genes contain stress-responsive HSEs [38,39] and their expression is developmentally regulated in diverse embryonic cells and tissues [35]. Several *cis*-acting elements influence mammalian sHSP gene expression during development and stress, including the upstream transcription factor, a member of the basic helix–loop–helix zipper transcription factor family; AP-1, Pax-6, MyoD, and CREB/ATF family members; MEF2 and SRF as members of the MADS-box family; glucocorticoid, estrogen, and retinoic acid receptors; GAGA factors; and HET (HSP27-ERE-TATA-binding protein) [35,40–42]. An AP-1 binding site occurs in intron 1 of p26 and an estrogen binding site in the 5' region of the p26 gene.

p26 mRNA was detected previously only in oviparous embryos [6]; however, in this study low levels of p26 gene transcripts were identified in ovoviparous embryos. The

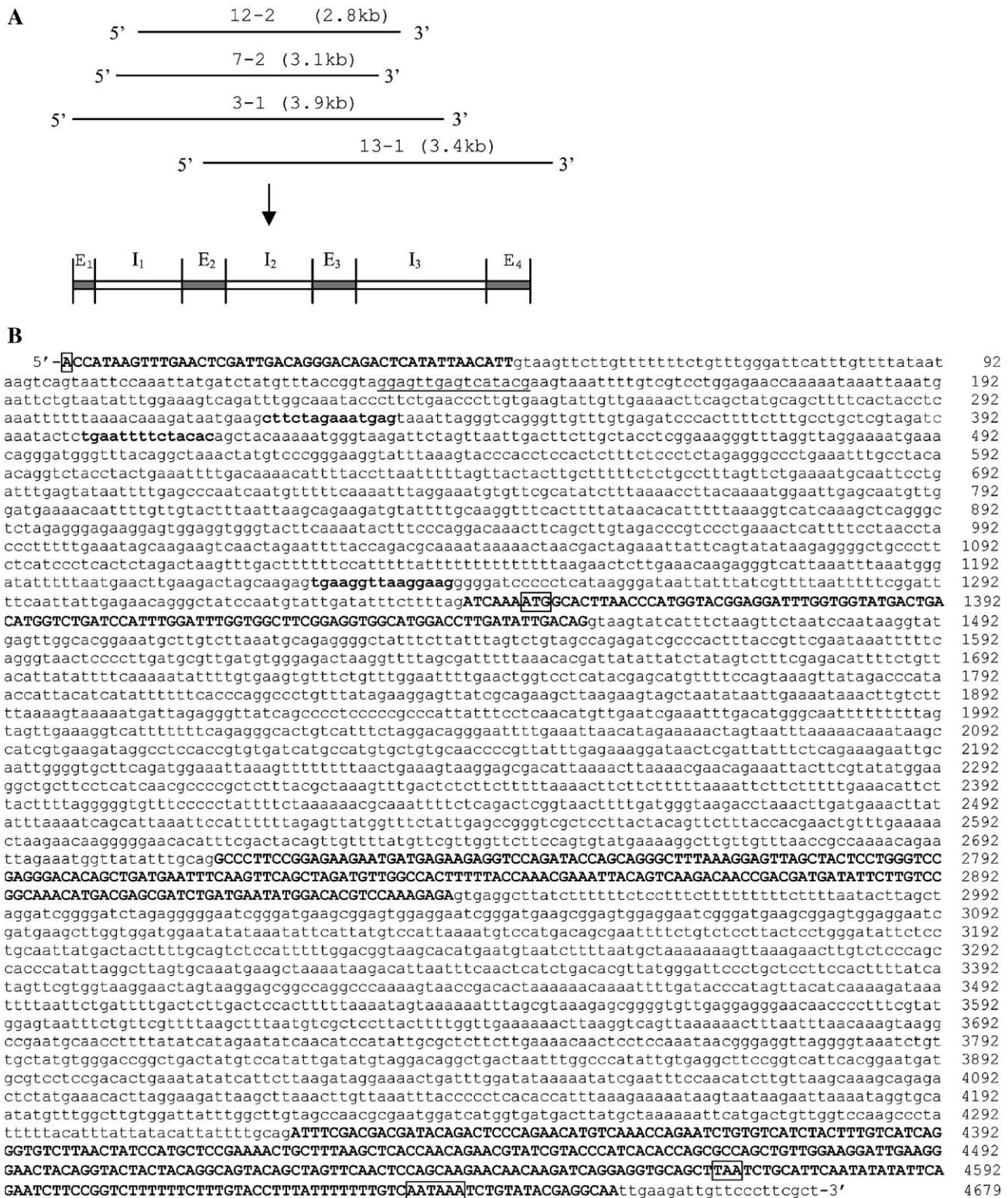


Fig. 3. Cloning and sequencing of the p26 gene from *A. franciscana* (ARC1520) GSL. (A) The p26 genomic clones 12-2, 7-2, 3-1, and 13-1 were restriction digested, electrophoresed in agarose gels, and blotted to membranes. DNA fragments on Southern blots hybridizing to labeled p26 cDNA probes, for which sizes are indicated in parentheses, were sequenced and aligned to yield the p26 gene (Accession No. DQ310575). E₁–E₄, exons; I₁–I₃, introns. (B) p26 gene sequence. Boxed A, transcription start site; boxed ATG, translation start site; gt and ag, intron boundaries; boxed TAA, stop codon; boxed AATAAA, polyadenylation signal; exons, uppercase letters in bold; introns, lowercase letters; HSEs, lowercase letters in bold; underline, AP-1 binding site; nucleotides 391–1390 are identical to those obtained by LA PCR in vitro cloning; nucleotide numbers are on the right.

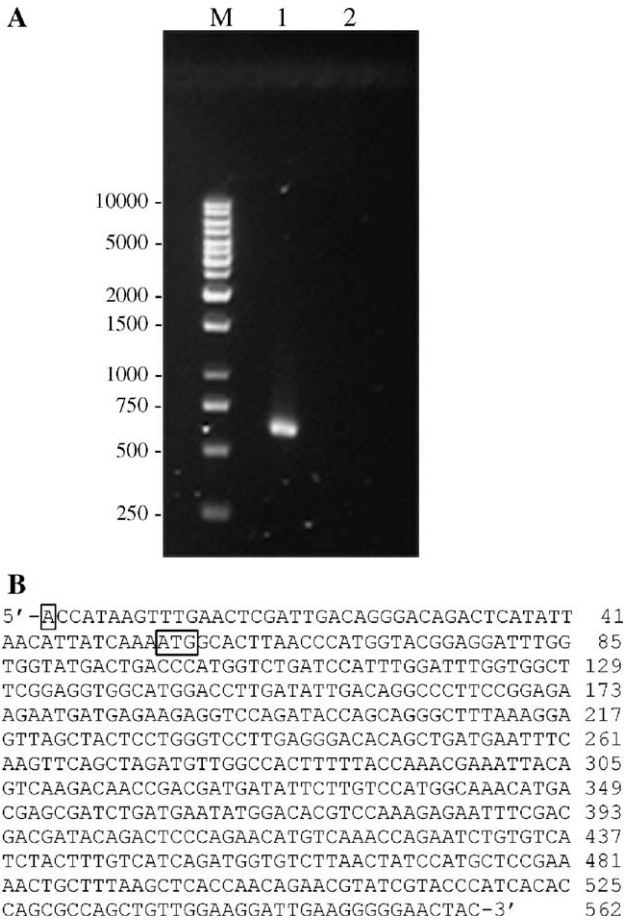


Fig. 4. Identification of the p26 gene transcription start site. (A) The products of 5'-RACE reactions were electrophoresed in agarose gels and visualized by staining with Gel Star. Lanes M, size markers, bp; 1, 5'-RACE products; 2, 5' RACE products in the absence of template DNA. (B) Sequence of DNA fragment generated by 5'-RACE. Boxed A, transcription start site; boxed ATG, translation start site. Nucleotide numbers are on the right.

detection of p26 mRNA in ovoviviparous embryos reflects the greater sensitivity of PCR compared to probing of Northern blots, and substantially more RNA was used for Northern blots than in past work. p26 was not observed in ovoviviparous embryos, indicating either a very small amount of the protein in these embryos or inhibition of p26 mRNA translation.

To summarize, p26 sequence and structure are well conserved across *Artemia* species. Arginine 114 appears in p26 from all species as do the amino-terminal WD/EPF and

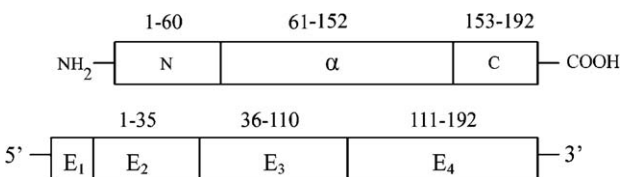


Fig. 5. Comparison of p26 domain and exon positions. Schematic representations of p26 protein and cDNA were drawn to scale and aligned. N, amino-terminal region; α , α -crystallin domain; C, carboxy-terminal extension; E₁–E₄, exons. Numbers above each schematic indicate p26 amino acid residues within protein domains and encoded by exons.

```
cctactgaaaaaaaaaacgggagatataatcagtgctag-1027
cgatgcaaaagggttttctcttttggttcaaggagggg -987
aaccataatcttctgtgaagagggttttcgaccgcaactg -947
attccaatggtagactttatctttgacctggcgccattt -907
tcgaagggatgttcaggcttttctttcagataaaat -867
ctttcgcataaaacttttaagacgaaccgaaataatag -827
ttaccattcctgagtttagtgcagatggcaattttttt -787
cactaggcatttttttcaaaatgcgtttttaactttt -747
ggttactatgggctgtggttttctctttactaggttgc -707
atctaccactgcaaccacgatcagcattttatcttact -667
atatgcaatttatatcatctccccattacagattcat -627
gtgcataatthaaggtttctgtgatgtttacgtaaatatc -587
ttgtgttataaatcaaaatgtgaattcatccccgcttt -547
taacagaactcatcgaccagattacagttaattataa -507
taaattattgaataaacctaaaagcttcagtttattaa -467
atgtgaattgggcatcaaaatgaattttcttaacaaaa -427
gtcttgtaaa[ccaat]ggaatcctttgagagtaaccaaat -387
ggcttataactaataaaataaatattcaaaagaattt -347
ctttatgatagttcttaataagtgtcttttgaacccaat -307
tgaactatataaaactgactatataacgccacaaatagc -267
aaaaattatgtgataagcttattgaaagcagggcat -227
atcaaaccttggttttgggtttgtaacacaaagaaaaat -187
aagcagctgaattttcagatataccgggattccaattct -147
ataacaacattttgatctagcctatttcacagaacc -107
gtagcttttgtaaggcctgttttagccaggtttaaac -67
tgaaaatattggtctcgccactgcaggttctctgcaaat -27
a[ata]accaagtaaaaacggttgatca[CCATAA] +7
```

Fig. 6. The noncoding region 5' of the p26 transcription start site was obtained by sequencing genomic clones. Boxed A, transcription start site; boxed tata, TATA box; boxed ccaat, CCAAT box; lowercase letters in bold, heat shock elements; underline, estrogen receptor binding site; nucleotide numbers are on the right.

carboxy-terminal VPI motifs. The amino acid substitutions observed in p26 appear to have limited effects although some may enhance oligomerization and thus affect function. Sequencing revealed three introns in the p26 gene with one in the 5'-UTR as occurs for one other sHSP gene. HSEs occur in the 5'-UTR intron and the upstream flanking region even though p26 gene expression is refractory to stress. In contrast to previous results, low levels of p26 mRNA were observed in ovoviviparously developing *Artemia* embryos, but p26 protein was not detectable. Future work will expand the analysis of p26

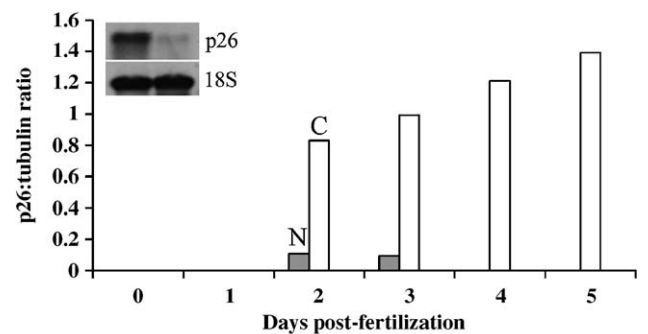


Fig. 7. Quantification of p26 mRNA during *Artemia* embryo development. p26 mRNA was quantified in *Artemia* embryos by real-time PCR as described under Materials and methods using tubulin mRNA as internal standard. Inset, p26 mRNA and 18S rRNA obtained 2 days postfertilization was electrophoresed in agarose, blotted to membranes, and hybridized to labeled probes. C, oviparously developing embryos; N, ovoviviparously developing embryos. Each lane received 5 μ g of RNA.

synthesis, structure, and expression in *Artemia* species, contributing to our understanding of sHSP function and more clearly defining the role of p26 in embryo development.

Materials and methods

Artemia culture

A. franciscana (ARC1520) cysts from the Great Salt Lake (GSL) (INVE Aquaculture, Inc., Ogden, UT, USA) were incubated at 27°C with shaking for 20 h at 200 rpm in hatch medium (422 mM NaCl, 9.4 mM KCl, 25.4 mM MgSO₄ · 7H₂O, 22.7 mM MgCl₂ · 6H₂O, 1.4 mM CaCl₂ · 2H₂O, 0.5 mM NaHCO₃) containing 0.1% (w/v) disodium tetraborate. Nauplii were transferred to seawater and maintained on either *Isochrysis galbana* (clone synonym ISO) or *Isochrysis sp* (clone synonym TISO) (Provasoli-Guillard Center for Culture of Marine Phytoplankton, West Boothbay Harbor, ME, USA). Adult females with oocyte-filled lateral pouches were placed with males in six-well culture plates and examined periodically with a dissecting microscope until fertilization occurred. Embryos recovered daily from ovisacs postfertilization were frozen immediately in liquid nitrogen.

Cloning and comparison of p26 cDNAs

mRNA was extracted from *A. franciscana* (ARC1520, GSL, UT, USA), *A. franciscana* (ARC1258, San Francisco Bay, CA, USA), *A. sinica* (ARC1218, Xiechi Lake, Yuncheng, Shanxi, China), *A. urmiana* (ARC1511, Urmia Lake, Azerbaijan, Iran), *A. parthenogenetica* (ARC1407, Karabogaz-Gol, Republic of Turkmenistan), and *A. persimilis* (ARC1321, Argentina) with the MicroFastTrack 2.0 Kit (Invitrogen, Burlington, ON, Canada). All *Artemia* species were obtained from the Artemia Reference Center, Faculty of Bioscience Engineering, Ghent University (Ghent, Belgium). RT-PCR was performed in 50- μ l reaction mixtures containing 1 \times Reaction Mix (Invitrogen), 1 μ l RT/Platinum TaqMix (Invitrogen), 0.2 μ M each primer, 3 mM Mg²⁺, and 5 μ l RNA at 1 cycle of 45°C for 30 min and 94°C for 2 min, 40 cycles of 94°C for 30 s, 55°C for 45 s, 72°C for 1.5 min, and extension at 72°C for 5 min. The forward primer was 5'-GCGCGGATCCACCATGGCACTTAACCCATG-3', containing nucleotides 1–17 of the p26 cDNA, and the reverse primer was 5'-CGCGCCTCGAGT-TAAGCTGCACCTCCTGTCT-3', with nucleotides 560–579 of p26 cDNA. PCR products were cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) and two or more p26 clones were sequenced for each species at the DNA Sequencing Facility, Centre for Applied Genomics, Hospital for Sick Children (Toronto, ON, Canada). Sequence comparisons were performed with Clustal W available at <http://www.ebi.ac.uk>.

Analysis of p26 sequence and three-dimensional structure

A neighbor-joining (NJ) tree was constructed with the software package Mega 3.1 [43] employing the Kimura two-parameter model NJ algorithm of nucleotide substitution. All sites containing alignment gaps were removed before calculation using the complete-deletion option and reliability of inferred phylogenetic trees was assessed by bootstrap analysis using 5000 pseudoreplicates.

p26 cDNA and deduced amino acid residue sequences were aligned by Clustal W (<http://www.ebi.ac.uk/clustalw/index.html>) and similarities between nucleotide and amino acid sequences were determined with NCBI Blast 2. Residues 58–157 of p26, including the α -crystallin domain and a short portion of the carboxy-terminal extension, from the five *Artemia* species examined in this study were modeled on the basis of the wheat Hsp16.9 [21] monomer structure, 1 gme_1:A, with Modeller [44] by optimizing the probability objective functions and simulated annealing. Ten models were generated for p26 from each *Artemia* species and the structure displaying the lowest objective function value represents the protein. Side-chain configurations were predicted with the graph-theory algorithm implemented in SCWRL 3.0 [45], residue environments were evaluated with Verify3D [46], and stereochemical quality was checked with ProCheck [47] and WhatIf [48]. Graphical representations were made with VMD [49].

Cloning the *A. franciscana* GSL (ARC1520) p26 gene

DNA was obtained from first-instar *Artemia* nauplii as described in [50] except ethanol precipitation was avoided prior to digestion and the DNA was dialyzed against two changes of 10 mM Tris-HCl, 1 mM EDTA, pH 7.4, for 36 h before storage at 4°C. The DNA was digested with *Sau3AI* (0.06 μ g DNA) at 37°C for 30 min, extracted twice with ethanol:chloroform:isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1), followed by ethanol precipitation. The DNA was dissolved in TE buffer and 200- μ g samples were centrifuged at 83,000g for 22 h at 20°C in 36-ml, 10–40% discontinuous gradients composed of sucrose in 1 M NaCl, 20 mM Tris-HCl, 5 mM EDTA (pH 8.0). The gradients were fractionated and samples containing DNA fragments of 14–23 kb, as determined by agarose gel electrophoresis, were pooled and incubated with Lambda Vector *Bam*HI Arms (Promega) for 3 h at room temperature and overnight at 4°C. Recombinant DNA was packaged with the Packagene Extract (Promega), library titer was determined, and 10 plaques picked randomly were digested with *Xho*I to check insert size.

Plaque DNA was transferred to duplicate nylon membranes; denatured with 0.5 M NaOH, 1.5 M NaCl for 5 min; neutralized in 1.5 M NaCl, 1.0 M Tris-HCl, pH 7.4, for 15 min; and incubated in 2 \times SSC for 10 min. The membranes were baked at 80°C for 2 h and prehybridized in Dig Easy Hyb (Roche, Mannheim, Germany) for 6 h at 42°C. Full-length p26 cDNA, labeled with digoxigenin-11-dUTP in PCR Dig Labeling Mix (Roche) using the primers 5'-GCGCGGATCCACCATGGCACTTAACCTG-3' (forward) and 5'-CGCGCCTCGAGTTAAGCTGCACCTCCTGTCT-3' (reverse), was denatured 5 min at 96°C and incubated with membranes in Dig Easy Hyb overnight at 46°C before being washed twice at room temperature for 5 min in 2 \times SSC containing 0.1% SDS (w/v), followed by two washes in 0.5 \times SSC, 0.1% SDS at 68°C for 15 min. Reactive plaques were detected as described by the manufacturer and clones hybridizing to p26 cDNA were screened four times by picking single plaques after dilution and plating on *Escherichia coli* lawns. These clones were double digested with *Bam*HI and *Xho*I, *Eco*RI and *Xho*I, *Hind*III and *Xho*I, and *Xba*I and *Xho*I; electrophoresed in agarose gels; blotted to membranes; and probed with labeled p26 cDNA to select fragments for cloning in pBluescript II SK(+) vector (Stratagene, La Jolla, CA, USA) and sequencing.

Putative *cis*-acting regulatory elements in intron 1 and the 5' upstream region of the p26 gene were identified with the MatInspector program available at www.genomatix.de.

LA PCR in vitro cloning of the p26 gene 5' flanking sequence

Five micrograms of *A. franciscana* GSL DNA was digested to completion overnight at 37°C with *Hind*III and ligated into a *Hind*III cassette (TaKaRa, Otsu, Japan) at 16°C for 30 min. After incubation at 94°C for 2 min, PCR was performed for 35 cycles at 94°C for 30 s, 62°C for 1 min, and 72°C for 2 min, followed by 72°C for 5 min. The C₁ and S₁ primers, based respectively on cassette sequence and nucleotides 86–107 of p26 cDNA, were 5'-GTACA-TATTGTCGTTAGAACGCGTAATACGACTCA-3' and 5'-CCACCTCC-GAAGCCACCAAATC-3'. One microliter of the PCR product was denatured for 1 min prior to amplification for 25 cycles at 94°C for 30 s, 61°C for 1 min, and 72°C for 2 min followed by 5 min at 72°C. The primers were C₂ (5'-CGTTAGAACGCGTAATACGACTCACTTAGGGAGA-3') and S₂ (5'-AGTCATACCACCAAATCCTCCGT-3'), based respectively on cassette sequence and nucleotides 42–63 of p26 cDNA. The PCR products were electrophoresed in 1% agarose gels, purified with the GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences, Baie d'Urfé, QC, Canada), cloned into the pGEM-T Easy vector (Promega), and sequenced.

Identification of the p26 gene transcription start site

Hydrated cysts were disrupted in TRIzol (Invitrogen) with a Dounce minihomogenizer, shaken vigorously in chloroform, and centrifuged at 12,000g for 15 min. The upper aqueous phase was incubated with isopropyl alcohol and RNA was collected by centrifugation. The pellet was washed with 75% (v/v) ethanol, dried, and dissolved in diethyl pyrocarbonate-treated water prior to recovery of poly(A)⁺ mRNA with an mRNA Purification Kit (Amersham). To

perform 5'-RACE with the SMART RACE cDNA amplification kit (BD Biosciences, Palo Alto, CA, USA), first-strand cDNA was synthesized using 1 µg of poly(A)⁺ mRNA in 10-µl reaction mixtures containing 1 µl of Smart II A oligonucleotide (5'-AAGCAGTGGTATCAACGCAGATACGCGGG-3'), 1 µl of 5'-CDS primer (5'-(T)₂₅N₁N-3'), 2 µl of 5× first-strand buffer, 1 µl of DTT, 1 µl of dNTP mix, and 1 µl of PowerScript reverse transcriptase. Amplification was performed with the adapter primers, Universal Primer A Mix long (5'-CTAATACGACTACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3') and short (5'-CTAATACGACTACTATAGGGC-3') and the primer (5'-GTAGTCCCCCTCAATCCTTCCAA-3') corresponding to nucleotides 506–530 of p26 cDNA at 25 cycles of 94°C for 30 s, 68°C for 30 s, and 72°C for 2 min, followed by 72°C for 5 min. The 5'-RACE products were electrophoresed in 1% agarose gels and cloned in the pGEM-T Easy vector (Promega) prior to sequencing.

p26 gene expression in *A. franciscana* (ARC1520) GSL

Total RNA was prepared daily postfertilization from oviparous and ovoviviparous embryos with the RNeasy Protect Mini Kit and QIAshredder (Qiagen, Mississauga, ON, Canada). Poly(A)⁺ RNA was purified with Oligotex (Qiagen) and reverse transcribed with Supermix (Sigma–Aldrich, St. Louis, MO, USA). Primers designed to span p26 intron regions were 5'-GTCCAAAGAGAATTCGACG-3' and 5'-CTGCACCTCCTGATCTTGT-3', in forward and reverse directions, respectively. Primers for α-tubulin, used as the internal standard, were 5'-CTGCATGCTGTACAGAGGAGATGT-3' and 5'-CTCCTTCAAGAGAGTCCATGCCAA-3', in forward and reverse directions, respectively. Real-time PCR was performed with IQ Supermix (Bio-Rad, Hercules, CA, USA) at 45 cycles of 94°C for 1 min, 47°C for 30 s, and 72°C for 30 s, followed by 10 min at 72°C in the Bio-Rad iCycler. The p26 mRNA quantification was done twice using independently prepared embryos at each time period examined.

For Northern blots 5 µg of total RNA prepared with TRIzol from *Artemia* embryos 2 days postfertilization was electrophoresed in 1.5% agarose gels containing formaldehyde and transferred to nylon membranes (Roche) in transfer buffer (Ambion, Austin, TX, USA) using the Rapid Downward System/Buffer Tray (Schleicher & Schuell, Keene, NH, USA). Blots were incubated in Dig Easy Hyb for 6 h at 50°C and hybridized overnight at 50°C with either full-length p26 cDNA or a 520-bp cDNA fragment encoding 18S ribosomal RNA labeled with digoxigenin-11-dUTP. Membranes were washed and hybridization was detected as for genomic clones.

Acknowledgments

This research was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant to T. H.M., EU-INCO DEV Concerted Action Project ICA4-CT-2001-10020, and Flemish–Chinese Bilateral Scientific Cooperation Project BIL 03/15. The authors thank Mark Fast, National Research Council Institute for Marine Biosciences, Halifax, for assistance with the real-time PCR.

References

- [1] T.H. MacRae, Molecular chaperones, stress resistance and development in *Artemia franciscana*, *Semin. Cell Dev. Biol.* 14 (2003) 251–258.
- [2] J.S. Clegg, S.A. Jackson, V.I. Popov, Long-term anoxia in encysted embryos of the crustacean, *Artemia franciscana*: viability, ultrastructure, and stress proteins, *Cell Tissue Res.* 301 (2000) 433–446.
- [3] P. Liang, R. Amons, T.H. MacRae, J.S. Clegg, Purification, structure and *in vitro* molecular-chaperone activity of *Artemia* p26, a small heat shock/α-crystallin protein, *Eur. J. Biochem.* 243 (1997) 225–232.
- [4] P. Liang, R. Amons, J.S. Clegg, T.H. MacRae, Molecular characterization of a small heat shock/α-crystallin protein in encysted *Artemia* embryos, *J. Biol. Chem.* 272 (1997) 19051–19058.
- [5] J.S. Clegg, J.K. Willsie, S.A. Jackson, Adaptive significance of a small heat shock/α-crystallin protein (p26) in encysted embryos of the brine shrimp, *Artemia franciscana*, *Am. Zool.* 39 (1999) 836–847.
- [6] P. Liang, T.H. MacRae, The synthesis of a small heat shock/α-crystallin protein in *Artemia* and its relationship to stress tolerance during development, *Dev. Biol.* 207 (1999) 445–456.
- [7] Y. Sun, M. Mansour, J.A. Crack, G.L. Gass, T.H. MacRae, Oligomerization, chaperone activity, and nuclear localization of p26, a small heat shock protein from *Artemia franciscana*, *J. Biol. Chem.* 279 (2004) 39999–40006.
- [8] Y. Sun, T.H. MacRae, Characterization of novel sequence motifs within N- and C-terminal extensions of p26, a small heat shock protein from *Artemia franciscana*, *FEBS J.* 272 (2005) 5143–5230.
- [9] T.S. Villeneuve, X. Ma, Y. Sun, M.M. Oulton, A.E. Oliver, T.H. MacRae, Inhibition of apoptosis by p26: implications for small heat shock protein function during *Artemia* development, *Cell Stress Chaperones* 11 (2006) 71–80.
- [10] S.A. Jackson, J.S. Clegg, Ontogeny of low molecular weight stress protein p26 during early development of the brine shrimp, *Artemia franciscana*, *Dev. Growth Differ.* 38 (1996) 153–160.
- [11] J.S. Clegg, S.A. Jackson, A.H. Warner, Extensive intracellular translocations of a major protein accompany anoxia in embryos of *Artemia franciscana*, *Exp. Cell Res.* 212 (1994) 77–83.
- [12] J.K. Willsie, J.S. Clegg, Small heat shock protein p26 associates with nuclear lamins and HSP70 in nuclei and nuclear matrix fractions from stressed cells, *J. Cell. Biochem.* 84 (2002) 601–614.
- [13] M. Haslbeck, T. Franzmann, D. Weinfurter, J. Buchner, Some like it hot: the structure and function of small heat-shock proteins, *Nat. Struct. Mol. Biol.* 12 (2005) 842–846.
- [14] Y. Sun, T.H. MacRae, Small heat shock proteins: molecular structure and chaperone function, *Cell. Mol. Life Sci.* 62 (2005) 2460–2476.
- [15] M. Haslbeck, A. Miess, T. Stromer, S. Walter, J. Buchner, Disassembling protein aggregates in the yeast cytosol: the cooperation of HSP26 with SSA1 and HSP104, *J. Biol. Chem.* 280 (2005) 23861–23868.
- [16] A.G. Cashikar, M. Duennwald, S.L. Lindquist, A chaperone pathway in protein disaggregation: HSP26 alters the nature of protein aggregates to facilitate reactivation by HSP104, *J. Biol. Chem.* 280 (2005) 23869–23875.
- [17] R.C. Ireland, E.M. Berger, Synthesis of low molecular weight heat shock peptides stimulated by ecdysterone in a cultured *Drosophila* cell line, *Proc. Natl. Acad. Sci. USA* 79 (1982) 855–859.
- [18] E.P.M. Candido, The small heat shock proteins of the nematode *Caenorhabditis elegans*: structure, regulation and biology, in: A.-P. Arrigo, W.E.G. Müller (Eds.), *Small Stress Proteins*, Springer, Berlin, 2002, pp. 61–77.
- [19] M.-A. Lopez-Matas, et al., Protein cryoprotective activity of a cytosolic small heat shock protein that accumulates constitutively in chestnut stems and is up-regulated by low and high temperatures, *Plant Physiol.* 134 (2004) 1708–1717.
- [20] G.-C. Fan, G. Chu, E.G. Kranias, Hsp20 and its cardioprotection, *Trends Cardiovasc. Med.* 15 (2005) 138–141.
- [21] R.L.M. van Montfort, E. Basha, K.L. Friedrich, C. Slingsby, E. Vierling, Crystal structure and assembly of a eukaryotic small heat shock protein, *Nat. Struct. Biol.* 8 (2001) 1025–1030.
- [22] G.V. Triantaphyllidis, et al., International Study on *Artemia*. LVII. Morphological and molecular characters suggest conspecificity of all bisexual European and North African *Artemia* populations, *Mar. Biol.* 129 (1997) 477–487.
- [23] T.J. Abatzopoulos, I. Kappas, P. Bossier, P. Sorgeloos, J.A. Beardmore, Genetic characterization of *Artemia tibetiana* (Crustacea: Anostraca), *Biol. J. Linn. Soc.* 75 (2002) 333–344.
- [24] Y. Sun, T.H. MacRae, The small heat shock proteins and their role in human disease, *FEBS J.* 272 (2005) 2613–2627.
- [25] J. Frydenberg, A.A. Hoffmann, V. Loeschcke, DNA sequence variation and latitudinal associations in *hsp23*, *hsp26* and *hsp27* from natural population of *Drosophila melanogaster*, *Mol. Ecol.* 12 (2003) 2025–2032.
- [26] G.-J. Caspers, J.A.M. Leunissen, W.W. de Jong, The expanding small heat-shock protein family, and structure predictions of the conserved “α-crystallin domain”, *J. Mol. Evol.* 40 (1995) 238–248.

- [27] B. Linder, Z. Jin, J.H. Freedman, C.S. Rubin, Molecular characterization of a novel, developmentally regulated small embryonic chaperone from *Caenorhabditis elegans*, *J. Biol. Chem.* 271 (1996) 30158–30166.
- [28] J.-C. Guan, T.-L. Jinn, C.-H. Yeh, S.-P. Feng, Y.-M. Chen, C.-Y. Lin, Characterization of the genomic structures and selective expression profiles of nine class I small heat shock protein genes clustered on two chromosomes in rice (*Oryza sativa* L.), *Plant Mol. Biol.* 56 (2004) 795–809.
- [29] R. Carranco, C. Almoguera, J. Jordano, A plant small heat shock protein gene expressed during zygotic embryogenesis but noninducible by heat stress, *J. Biol. Chem.* 272 (1997) 27470–27475.
- [30] K.-D. Scharf, M. Siddique, E. Vierling, The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing α -crystallin domains (Acid proteins), *Cell Stress Chaperones* 6 (2001) 225–237.
- [31] L.F. Cooper, K. Uoshima, Z. Guo, Transcriptional regulation involving the intronic heat shock element of the rat hsp27 gene, *Biochim. Biophys. Acta* 1490 (2000) 348–354.
- [32] R. Carranco, C. Almoguera, J. Jordano, An imperfect heat shock element and different upstream sequences are required for the seed-specific expression of a small heat shock protein gene, *Plant Physiol.* 121 (1999) 723–730.
- [33] R. Klemenz, W.J. Gehring, Sequence requirement for expression of the *Drosophila melanogaster* heat shock protein *hsp22* gene during heat shock and normal development, *Mol. Cell. Biol.* 6 (1986) 2011–2019.
- [34] J. Amin, R. Mestriil, R. Voellmy, Genes for *Drosophila* small heat shock proteins are regulated differently by ecdysterone, *Mol. Cell. Biol.* 11 (1991) 5937–5944.
- [35] S.M. Davidson, M.-T. Loones, O. Duverger, M. Morange, The developmental expression of small HSP, in: A.-P. Arrigo, W.E.G. Müller (Eds.), *Small Stress Proteins*, Springer, Berlin, 2002, pp. 103–128.
- [36] G. Wall, P.D. Varga-Weisz, R. Sandaltzopoulos, P.B. Becker, Chromatin remodeling by GAGA factor and heat shock factor at the hypersensitive *Drosophila* hsp26 promoter in vitro, *EMBO J.* 14 (1995) 1727–1736.
- [37] S. Michaud, R. Marin, J.T. Westwood, R.M. Tanguay, Cell-specific expression and heat-shock induction of Hsps during spermatogenesis in *Drosophila melanogaster*, *J. Cell Sci.* 110 (1997) 1989–1997.
- [38] K. Kato, H. Ito, Y. Inaguma, Expression and phosphorylation of mammalian small heat shock proteins, in: A.-P. Arrigo, W.E.G. Müller (Eds.), *Small Stress Proteins*, Springer, Berlin, 2002, pp. 127–150.
- [39] R. Voellmy, On mechanisms that control heat shock transcription factor activity in metazoan cells, *Cell Stress Chaperones* 9 (2004) 122–133.
- [40] R. Gopal-Srivastava, A. Cvekl, J. Piatigorsky, Involvement of retinoic acid/retinoid receptors in the regulation of murine α B-crystallin/small heat shock protein gene expression in the lens, *J. Biol. Chem.* 273 (1998) 17954–17961.
- [41] R. Gopal-Srivastava, A. Cvekl, J. Piatigorsky, Pax-6 and α B-crystallin/small heat shock protein gene regulation in the murine lens: interaction with the lens-specific regions, LSR1 and LSR2, *J. Biol. Chem.* 271 (1996) 23029–23036.
- [42] J.G. Ilagan, A. Cvekl, M. Kantorow, J. Piatigorsky, C.M. Sax, Regulation of α A-crystallin gene expression: Lens specificity achieved through the differential placement of similar transcriptional control elements in mouse and chicken, *J. Biol. Chem.* 274 (1999) 19973–19978.
- [43] S. Kumar, K. Tamura, M. Nei, MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment, *Brief Bioinform.* 5 (2004) 150–163.
- [44] A. Šali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [45] A.A. Canutescu, A.A. Shelenkov, R.L. Dunbrack Jr., A graph-theory algorithm for rapid protein side-chain prediction, *Protein Sci.* 12 (2003) 2001–2014.
- [46] R. Lüthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature* 356 (1992) 83–85.
- [47] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [48] R. Rodríguez, G. China, N. Lopez, T. Pons, G. Vriend, Homology modeling, model and software evaluation: three related resources, *CABIOS* 14 (1998) 523–528.
- [49] W. Humphrey, A. Dalke, K. Schulten, VMD—Visual molecular dynamics, *J. Mol. Graphics Modell.* 14 (1996) 33–38.
- [50] C.M. Langdon, J.C. Bagshaw, T.H. MacRae, Tubulin isoforms in the brine shrimp, *Artemia*: primary gene products and their posttranslational modification, *Eur. J. Cell Biol.* 52 (1990) 17–26.