

Cloning and Characterization of cDNAs Encoding Trehalase from Post-Dormant Embryos of the Brine Shrimp, *Artemia franciscana*

Shin Tanaka*, Fumiko Nambu and Ziro Nambu

Biology, Department of Medical Technology, School of Health Sciences,
University of Occupational and Environmental Health, Japan,
Yahatanishi-ku, Kitakyushu 807-8555, Japan

ABSTRACT—To investigate post-dormant regulation of trehalose metabolism in the brine shrimp, we cloned and characterized two trehalase cDNAs from embryos of *Artemia franciscana* using a PCR probe corresponding to a highly conserved region among trehalases. The cDNAs consisted of 2496 and 2485 nucleotides, and had almost the same open reading frame encoding 703 amino acids which showed 46.6–42.6% similarities to trehalases of animals. The calculated molecular mass of the trehalase was 79,995 Da. The deduced sequence had a cleavable signal peptide, a cell adhesion motif, four potential N-glycosylation sites, trehalase signatures and a unique, long carboxyl terminal polypeptide containing a predicted transmembrane region and a potential cAMP-dependent phosphorylation site. Phylogenetic analyses showed a large divergence among trehalases of arthropods. Northern blots revealed the presence of two mRNAs. One of them, a 2.6 kb mRNA, was abundant in the dormant cysts and prenauplii. The other 5.0 kb mRNA was newly synthesized during post-dormant development. Possible mechanisms of trehalase regulation are presented on the basis of the results shown by the Northern blots and developmental changes of trehalase activity.

INTRODUCTION

Various organisms accumulate α, α -trehalose, a non-reducing disaccharide consisting of two glucoses, against physiological and environmental changes. The accumulated trehalose has a role as an anti-stress molecule. For instance, in yeast, such anti-stress effects of trehalose against dehydration, heat shock, freezing and ethanol have been reported (reviewed by Nwaka and Holzer, 1998). These anti-stress effects of trehalose are thought to be due to its protective function on the cell membrane or proteins (Crowe *et al.*, 1987, 1988).

Embryos of *Artemia* entering dormancy also accumulate α, α -trehalose. This accumulation is not observed in non-dormant embryos (Clegg, 1962, 1965). The amount of trehalose in the encysted dormant gastrula represents 15% of its dry weight (Clegg, 1962, 1965). Under appropriate circumstances, the encysted embryos resume their subsequent development and the trehalose is hydrolyzed by a trehalase (EC 3.2.1.28).

In our previous study (Nambu *et al.*, 1997a), a soluble, alkaline trehalase with a molecular mass of 70 kDa on SDS-polyacrylamide gel electrophoresis was purified from devel-

oping embryos of *Artemia*. The *Artemia* trehalase showed a significant increase in its activity at the prenauplius stage and the activity decreased after hatching (Nambu *et al.*, 1997a). Hydrolysis of trehalose provides glucose to produce glycogen and glycerol, which are respiratory substrates in the *Artemia* cysts (Clegg, 1964; Ewing and Clegg, 1969). It is also suggested that accumulated glycerol increases the osmotic pressure in the cysts, resulting in a rupture of the shell followed by the emergence of prenauplii surrounded by hatching membrane (Clegg, 1964). Therefore, the temporal regulation of the trehalase activity is significant in post-dormant development. To investigate the mechanisms of trehalase regulation, we cloned a 228 bp cDNA fragment of the *Artemia* trehalase using a polymerase chain reaction (PCR) (Nambu *et al.*, 1997b). In the present report, we describe the cloning and characterization of *Artemia* trehalase cDNAs encoding a full open reading frame (ORF) and show the results of Northern blots and phylogenetic analyses.

MATERIALS AND METHODS

Animals

Dried cysts of *Artemia franciscana* from the Great Salt Lake in Utah, USA were obtained from Japan Pet Drugs Co. (Tokyo and Los Angeles, CA, USA). The cysts were hydrated and cultured as described by Nambu *et al.* (1997a).

* Corresponding author: Tel. +81-93-603-1611(Ext. 4539);
FAX. +81-93-691-7142.
E-mail. artemia@med.uoeh-u.ac.jp

RNA extraction

Total RNA was extracted from 0.5 g of the wet weight of the embryos and nauplii according to the method described by Maniatis *et al.* (1982) with some modifications. Each specimen was homogenized in 2.5 ml of a guanidine thiocyanate solution. The homogenate was centrifuged at 4,000 rpm for 5 min and total RNA was extracted from the supernatant. Poly (A)⁺ RNA was purified using Oligotex dT30 <Super> (Roche Japan, Tokyo, Japan).

cDNA cloning and nucleotide sequencing

Double stranded cDNA was synthesized from purified poly (A)⁺ RNA extracted from 12 hr cultured prenauplii of *A. franciscana* using oligo(dT)₁₂₋₁₈ primer. The double stranded cDNA was ligated with *EcoRI*-*NotI* adaptor and a cDNA library was constructed in λ gt10 vector (Stratagene) using gigapack®III Gold (Stratagene). The 228 bp cDNA fragment of the *Artemia* trehalase was obtained by the degenerated PCR and digoxigenin (DIG)-labelled as described by Nambu *et al.* (1997b). It was used to screen 2.0×10^5 clones. Phage plaques were lifted up and immobilized on nylon membranes (Boehringer Mannheim, Mannheim, Germany). Hybridization and detection was performed as described by Engler-Blum *et al.* (1993). Positive clones were isolated and two longer inserts termed *TreE2* and *TreE3* were fully sequenced using Amersham thermo sequenase fluorescent labelled primer cycle sequence kit (Amersham, Buckinghamshire, UK) and an ABI 373 sequencer (Perkin Elmer, Norwalk, CT, USA).

Northern blot analysis

The full length of *TreE3* was DIG-labelled according to the random priming method using a DIG DNA Labelling Kit (Boehringer Mannheim).

Total RNA was prepared as described in the section of RNA extraction, and 20 μ g of each total RNA was electrophoresed in a 1.0% agarose gel containing formaldehyde (Sambrook *et al.*, 1989) and blotted on to Hybond N⁺ membrane (Amersham). Hybridization was performed at 68°C as described by Engler-Blum *et al.* (1993) followed by washing at 65°C.

RESULTS

Isolation of trehalase cDNAs from *A. franciscana*

Screening of the cDNA library derived from the 12 hr cultured embryos of *A. franciscana* with the PCR product resulted in isolation of 12 positive clones. Two longer inserts, *TreE2* and *TreE3*, were used for full-length nucleotide sequencing and further investigation.

Comparison of *TreE2* with *TreE3*

TreE2 consisted of 2496 nucleotides while *TreE3* contained 2485 nucleotides. These had a common region consisting of 2476 nucleotides (Fig. 1). Identity between the two cDNAs was 99.7%. Seven replacements were found at the nucleotide level (Fig. 1). Both cDNAs contained the same ORF followed by a termination codon and an "AGTAAA" sequence which was similar to the polyadenylation signal, "AATAAA". An oligo (A) sequence was found 14 nucleotides downstream from the "AGTAAA" sequence in *TreE2* (Fig. 1, 2). The pro-

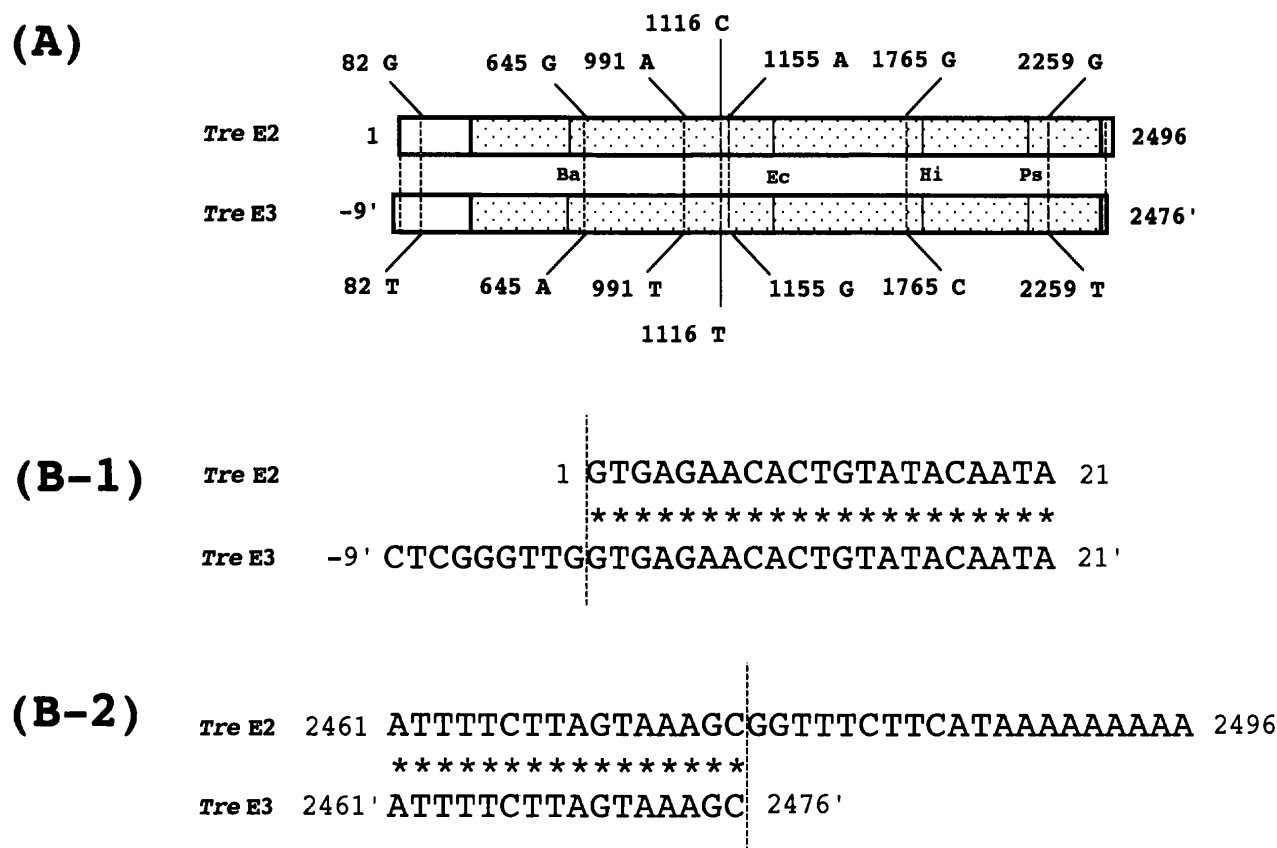


Fig. 1. Profiles of the *Artemia* trehalase cDNAs, *TreE2* and *TreE3*. Nucleotides of *TreE2* were numbered from the first nucleotide. Nucleotide numbers of *TreE3* were matched to the corresponding nucleotides of *TreE2*. The sequence identity between the two cDNAs was 99.7%. (A) Varied nucleotides are indicated by their number. Dotted areas indicate ORF regions. Both cDNAs had identical restriction sites. Ba, *Bam*HI; Ec, *Eco*RI; Hi, *Hind*III; Ps, *Pst*I. (B-1) 5'-Termini of both the cDNAs are represented. (B-2) 3'-Termini of *TreE2* and *TreE3* are aligned. *TreE2* had a oligo(A) sequence on its 3'-terminus. Asterisks indicate identical nucleotides between the two cDNAs.

```

GTGAGAACTGTATACAATACTTCTCTTGTATCAGATTGTTCCGCCGCTCAACATCTTTGTGATATTTCTATATAGAAGTGATATTCGA 90
TTCTCTTTCCAGAGATGAACATCAGGATTTCAGGAGGAAATGAAAGATTGCGCAGACTCAGAGGAAAGGTCCACCAAGAAAGGAGA 180
AGGCAACGTGCAAGTAAAAAGTGAACCTTTTGTAGTACTCTTGAAGAAATAAAAGATATTTTATCACTAAACATGTGCGCATTTTGT 270
      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *
AGTTTTTCGTCGTTTTTGGATTGCTACAACAGTTAGTTGTAGAGCGCATCTGGTCTCAGTGAAAGAGCAAAACAACCTACCGCCTGTT 360
S F F V V F S F A T T V S C R G D L V L S E R A K Q L P P V 35
TGTGCTAGCGAAATTTACTGCCATGGAACATTCTGCATACAATACAGATGGCAGGGCTGTTTCGAGACTCGAAAACCTTTGTAGATAAA 450
C A S E I Y C H G T F L H T I Q M A G L F R D S K T F V D K 65
AAGCTAAAGATAAAATCCAGAAGAAGTTTGTAGCTAGCTTTGAAGTTTGTATGAACCTCAACCGATCAAAATCCAGCCGGGACCAATTTGGCT 540
K L K I N P E E V L A S F E V L M N S T D Q N P S R D Q L A 95
GCTTTTATAAATCTACACTTCGAGCCAGAAGGGTCAGAATTTGAAGAATGGGATCCAATAGATTGGCATTCTAACCCAAGTTTGTGGAT 630
A F I N L H F E P E G S E F E E W D P I D W H S N P S F L D 125
GGTATTAGACACGAACTTGAAAAATATGGGAAATCTCTCCACGAAGCCTGGACATGTTAGGTGCAAAAATCCGTGACGACGTGAGA 720
G I R D T N L K I W G N S L H E A W T W L G R K I R D D V R 155
ATAAATCCCAATTATATTTCATGATGATATCTCCCAACCCCTTTCATCATCTCTGGTGGTCTGTTCCGTGAAACTTACTACTGGGATTC 810
I N P Q L Y S M M Y L P N P F I I E S S K E R E F Y Y W D S 185
TATTGGATTATCAAGGGCTTCTTATCTCGGGAATGCATGAGACGGTGAAGGGTATGTTGTGAACCTTTTGTGATGTTGGATACCATT 900
Y W I I K G L L I S G M H E T V K G M L L N F L L M V D T I 215
GGCCTTGTACCCAAATGGTGAAGGATATATTATGAGAAACGATCTCAGCCTCCACTTCTTACTCCAATGGTGAACCTTTACGTTAATGCT 990
G L V P N G G R I Y Y E K R S Q P P L L T P M V E L Y V N A 245
ACGGGAGACATTGAATTTTAAAGCAAAACATTCACCTGTTGAAAAGAGATGGATTTTGGCTTCAGGAGCGAACGGTCAACGTTGAT 1080
T G D I E F L K Q N I H L L E K E M D F W L Q E R T V N V D 275
GGGCACCGGCTTATCCGTTACGATGTTAAAGTTGGCGGCTCCCGACCTGAATCTTACAAGGAAGATATAGAAGCAATTCATCATACCCCT 1170
G H R L I R Y D V K V G G P R P E S Y K E D I E A I H H N P 305
GATTAGAAGCACAATTTGGATTTCATATGAACATAGCAGCTGGTGTCTGAAACAGGATGGGATTTCTCAAGCCGATGGTATTGGAACGGG 1260
D L E A Q L D F Y M N I A A G A E T G W D F S S R W Y W N G 335
GATATACAAACTAATTGTGCATGTTTGAACCCGTGACATACTACCTGTAGACCTGAATCTTTTATTGCTTGGGACTTTGATATTATG 1350
D I Q T N L S H V R T R D I L P V D L N S F I A W D F D I M 365
TCCCGCTTTGAAAGCAGTTAGGTGCGAGACAACGCATCAGTTGCTTATTCTGATTATACAGTGAATGGAAAACGTCGATTAAATGCCATT 1440
S R F E K Q L G R D N A S V V Y S D L Y S E W K T S I N A I 395
TTGTGGGATGATGAAGCTGGTAGCTGGTTTGAATTACGATTCTGCGCATCGGAGATGGAACACGAACCTTTATGTTTCTAATTTGACTCCC 1530
L W D D E A G S W F D Y D S A H R R W N T N F Y V S N L T P 425
TTATTCGTAGGATGTTACGATCCGAAACAGTTTCATCATGAAGACGTTGCCACAGAGTACTCGATTACCTTGAAAAAGTAACGCCCTTG 1620
L F V G C Y D P K T V H H E D V A T R V L D Y L E K S N A L 455
AAGTTTCTGGTGGTGTACCAACATCATTTGATGACAGCAAGCAACATGGGACTTTCTTAATGGCTGGCCCTCTCTCCAGCACATGCTG 1710
K F P G G V P T S L M Q T S Q Q W D F P N G W P P L Q H M L 485
GTTATGGGACTTGACAAGACCGGTGATCCAAAGGCCAAAGAACTGGCTTTTGACGTTGCGCAGAGATGGGTTTCAATAAATTATGAAGCG 1800
V M G L D K T G D P R A K E L A F D V A Q R W V F N N Y E A 515
TTCACGCAAGCTTACCTAATGCCATGTTTGAAGATATGACGTAACCGTCTGTTGGCTCCCGGCTGGTGGGGAGAGTACGACGTTCAA 1890
F T Q S L P N A M F E K Y D V T V V G L P G G G G E Y D V Q 545
CTAGGTTTGGCTGGACAAATGGTGTAGTAATGGACTTTTGTAGTTAAATACGGTGAGAGATTGACTTCAGTCCGTGACGCAAAACAATA 1980
L G F G W T N G V V M D F L V K Y G E R L T S V R D A K Q L 575
ATTTGGAGTCCATCATCAGTTTCAAGTTGGAGTTGGTCTGGTGAACGTTCTCGGTGCTTGGATTGGTCTTGAAGTACTCGATTCTAT 2070
I W S P S S V S V G V G A G A T F L G A W I G L A S T R F Y 605
CGCCAACTTTTCAAGTCCGACCAAAATGGGAGGCATCCATGGTGTGGCTGAGGCAATTCAAGGACTTATGCTCTGTGTTGTTCTCTG 2160
R Q L F K S D H K M G G I H G V A E A I Q G L I A C V G S L 635
CTCGTGGCTACTGCGACACTTTTGTAGTAAAGGATACGACGAAACGACCTACCTCTGGACAGACCCCTCAGTCCAGCCCAATAAGGAGA 2250
L V A T A A L F S E K D T T K R P T S G Q T P Q S S P I R R 665
ACAAAATCGTTACCCGGCTTCTTATTGTTTCAACGAGGAAGTAAGGCGACGAGAGCATCTCTTAAGGAACCGCGCAATTAGTGAAGAA 2340
T K S L P G F L L F Q R G S K G D E E H L L R N R A I S E E 695
TGTTCTGATTCTCTATAGAAGATGATGAAGGAGACTTTAGTGGATACAATTTTCTTCATATTGCTAGGAAATGCCAGAAGTTCAAAT 2430
C S Y S S I E E * * * * * 703
TAATATTCATTGTAATTTTCGAGTGTGATATTTTCTTAGTAAAGCGGTTTCTTCATAAAAAAAA 2496
      *      *

```

Fig. 2. Nucleotide and deduced amino acid sequences of *TreE2*. The nucleotide and amino acid sequences are numbered from the first nucleotide and the first methionine codon. Asterisks indicate stop codons. A region used as a probe in the screening is underlined. A sequence similar to the polyadenylation signal is heavily underlined. Cleavable N-terminal signal peptide is outlined. The cell adhesion motif is boxed. Potential N-glycosylation sites are dotted. The slashed box indicates the "trehalase signature 1" and the dotted box the "trehalase signature 2" (Henrissat and Bairoch, 1993). Both the signatures are submitted on the PROSITE database. The Thr²⁴⁶ and Val⁵⁰⁴ with circles are replaced by Ser²⁴⁶ and Leu⁵⁰⁴ in the polypeptide coded by *TreE3*. Predicted transmembrane region is printed in reverse. Potential cAMP-dependent phosphorylation site is doubly underlined.

AR	MCAFLSFFVVFVFATTVSCRGLVLSERAKQLPPVCASEIYCHGTFLLHTIQMAGLFRDSKTFVDKKLKINPEEV	75
HO	MPG-RTWELCL-L-LLL---G-LGLG-SQEALPPPCSEIYCHGELLNQVQMAKLYQDDKQFVDMPLSIAPEQVL	67
OR	MPG-STWELHL-L-LLL---G-LGLG-SEQALPPPCSEIYCHGELLNQVQMAKLYPDDKQFVDMPLSTAPDQVL	67
TE	-----MIPFLLMVAFADTVL-QVSAQSQSCDSKVYCOGKLLHVVEHSRIFNCSKTFVELKMINDEQTTL	64
BO	-----MRLFLLLVGLT-TV---IADDLPPTCIRPVYCNSTLLHYVQMAKLYPDSKTFVDQMRKDNATL	61
AR	ASHEVLNMSTQNPQRDLAAEINLHFEPEGSEFEEDWDEIDWHSNFSFLDGTITNLKIWGNLSHEAWTLGRKI	150
HO	QTFTELSRDHNHSIPREQLQAEVHEHEQAKGOELQPWTADWKDSPOFLQKISDAKLRAWAGQLHQLWKKLGKRM	142
OR	QSFAEALATYNNVTVPREQLKQFVQEHQAVGOBLESWTPGDWKESEPOFLQKISDPKLRAWAEQLHLLWKKLGKMI	142
TE	ENFDNFLRDTNHKRTRADLMKEVSDNF-KQENEFESWTPDFTDNTLLSRLEDKTIHQFADLVKIMPTLARKV	138
BO	SAFQELDRDTNHNPTEKDELQEFVVDFT-DETSLEEWKDEDHKENPELAKITRQGFREFAKALNDIMPTLARRV	135
Trehalase signature 1		
AR	RDDVRINPOLYSMMYLPNPFIIIPGGRERETYYWDSYWIIGLLISGHETVKGNLLNLFLLMVDITILVPGGRIY	225
HO	KPEVLSHPERFSLIYSEHPFIVPGGRFVEFYWDSYWMGLLLSEHAETVKGNLQNFLLDVTKYCHVPGGGRVY	217
OR	KPEVLSQPERFSLIYSQHPFIVPGGRFVEFYWDSYWMGLLLSEHAETVKGNLQNFLLDVTKYCHVPGGGRVY	217
TE	KKEVLDPYEHYSLLPVDNCFIIPGGRFTEFYWDSYWIVGLLLSDMHETVKGHLNDFLSIVEKYGFIPNGARV	213
BO	KPSVLEKPEQSSIVPMTHGFIIPGGRKEIYYWDAYWIEGLLITDITETAKGNIENLIELLYKFGHIPNGSRWY	210
AR	YEKRSQPLLTPMVELYVNATGDIFFLKQNIHLEKEMDFWLQERT--VNVDG--HRLIRYDVKVGGPRPESYKE	296
HO	YLQRSQPLLTLMDCYLTHNDTAFLOENIETLALELDFWTKNRTVSVSLECKNYLLNRYYPVYGGPRPESYSK	292
OR	YLQRSQPLLTLMDRYVAHTGDLAFLRENIETLALELDFWAENRTISVSSGNSHTLNRYHVPYGGPRPESYSK	292
TE	YLNRSQPLLTLHVSLSYVSATNDMEWLAKNIRTIDTELRFWLNRLVDVVKDCIVYKLAQYNSNSGSPRPESYYE	288
BO	YQERSQPLLAAIKLYYEKTKDIEFKRYISALEKELEYLWDLTHLIAFNKDDRVYTLRYIYIPSAQPRPESYYE	285
▽		
AR	DIEAIHHPDLEAQLDFYMNIAAGETGWDFFSSRW--YWNCDIQTNLSHVRTRDILPVDLNSFIAWDFDIMSREE	369
HO	DVELADTLPEGDRE-ALWAEKKAESGWDFFSSRW--LIGCPNPNSLSGIRTSKLVVDLNAFLCQAEELMSNFY	364
OR	DTELAHTLPEGSWE-TLWAEKKAESGWDFFSSRW--LVGSPNPDSLSGIRTSKLVVDLNAFLCQAEELLSGFY	364
TE	DVTTSVFSDERDKAELYMDLKSAAESGWDFFSSRWIVDEYGGTRGNLSALEHTRRIIPVDLNAFLCQAFQKLSBFY	363
BO	DYELAQKLDKNTDPNDIYADLKSAAESGWDFFTRWFISEGDNNGNLNLTNKNVIPVDLNAIFAGALQITANFQ	345
AR	KQLGRDNASVVSIDLYSEWKTSINAILWDDEAGSMFDYSAHRRWNTNFYVSNLTPLFVGCYDPKTVHHEDVATR	444
HO	SRLGNDSQATKYRILRSQRLAALNTVLWDEQTCAMFDYDLEKKKKNREFYPSNLTPLWAGCF-SDP-GVAD---K	434
OR	SRLGNESQATKYRNLRAQRIAAALTALWDEDKCAWFDYDLENQKKNEFYPSNLTPLWAGCF-SDP-AIAD---K	434
TE	QTLGDYPNATFWSKLVKIQHSIEMVHYNRDDCIWYDWNELSQHRRMFFPSNFAPLW--SETFDSRNAEILGEM	436
BO	AILKNPRRAHWGYMAEQWRSSIEQALWDEEDCVWHDYDILNNKPRRYFYTSNLAPLW--MNAVEKPFLLAKHGAR	433
Trehalase signature 2		
AR	VLDYLEKSNAKFPGGVPTSIMQTSQOWDFPNQWPPLOHMLVMGLDRTGDPRAKELAFDVAQRWVFNNYEAFQTQS	519
HO	ALKYLEDNRILTYQYGIPTSLQKTGQOWDFPNAWAPLQDLVIRGLAKAPLRRAQEVAFQLAQNWIRTNFDVYSQ-	508
OR	ALQYLQDSQILNHRHGIPTSLQNTGQOWDFPNAWAPLQDLVIRGLAKSPSARTQEVAFQLAQNWIRTNFDVYSQ-	508
TE	AAEYFITQNMMDYHGGIPTSLSHTGBOWDYPNAPPMQSIIVMGLDKSGSYRAKQLARELARRW-VKANLIGF-R	509
BO	VLEYLHESQALEYPGGPVSLVNSGBOWDFPNAPPEVSIIVTATQNGISEESSKLAKELAQVW-VRACKSGF-T	506
AR	LPNMFEEKYDVTVGLPGGGGEYDVLGFGWNGVVMDFVVKYGERLTSVRDAKQLIWSFSSVSVGVGAGATFLG	594
HO	-KSAMFEKYDVSNGGQPGGGGEYVQEGFGWDEGVVLMLEDRYGDRLTSGAKLAFLE-PHCLAAATLLPSLLLSLL	581
OR	-RSAMFEKYDISN-AQPGGGGEYVQEGFGWNGVVALMLEDRYGDRLSSGTQLALLE-PHCLAAALLLSFLTR--	578
TE	QTGEMFEKYNVEVPQNGGGGEYVVGSGFGWNGVVLFINQFFTT-----	555
BO	EKKQMFEEKYDALNACKYGGGGEYTVQDGFGWNGVVLFEIDRYGAVLTSVDSVDASANNQGSNEESETDSKEK--	579
AR	AWIGLASTRFYRQLFKSDHKMGGIHGVAEAIQGLIACVGSLLVATAALFSEKDTTKRPTSGQTPQSSPIRRTKSL	669
HO	PW-----	583
OR	-----	
TE	-----	
BO	-----	
AR	PGFLLFQRGSKGDEEHLNRNRAISEECSYSSIEE	703
HO	-----	
OR	-----	
TE	-----	
BO	-----	

Fig. 3. Alignment of the animal trehalases. The amino acid sequences of precursors of animal trehalases were aligned by the Clustal W program (Thompson *et al.*, 1994). Predicted N-terminal cleavage sites are indicated by arrowheads. Identical regions among the trehalases are printed in reverse. Positions of trehalase signatures are indicated by bold lines. The position of N-glycosylation site conserved among arthropod trehalases is indicated by a blank arrowhead. A unique, long C-terminal polypeptide of the *Artemia* trehalase is boxed. AR, *Artemia*; HO, human; OR, rabbit; TE, mealworm beetle; BO, silkworm.

teins deduced from *TreE2* and *TreE3* comprised 703 amino acids and showed the same molecular mass of 79,995 Da and the same isoelectric point of 5.22. Proteins encoded by these cDNAs were identical except for replacements of two amino acids (Fig. 2, *TreE2*: Thr²⁴⁶ → *TreE3*: Ser²⁴⁶ and *TreE2*: Val⁵⁰⁴ → *TreE3*: Leu⁵⁰⁴).

Characterization of *Artemia* trehalase cDNA

In a computer-assisted search of databases for SWISS-PROT and PIR, *Artemia* trehalase showed no significant similarity to other proteins than trehalases. The protein was highly similar to animal trehalases: 46.6% to human (Ishihara *et al.*, 1997), 46.1% to rabbit (Ruf *et al.*, 1990), 43.4% to mealworm beetle (Takiguchi *et al.*, 1992), and 42.6% to silkworm (Su *et al.*, 1993, 1994). The identities of the *Artemia* trehalase to trehalases of *E. coli* (Gutierrez *et al.*, 1989; Horlacher *et al.*, 1996) and neutral trehalases of fungi (Kopp *et al.*, 1993, 1994; Feldmann *et al.*, 1994; Wolfe and Lohan, 1994; Nwaka *et al.*, 1995; Amaral *et al.*, 1997; Eck *et al.*, 1997) were relatively low (24.6–32.0%). No significant similarity was observed between the *Artemia* trehalase and acid trehalases of fungi, such as *ATH1* gene product of *S. cerevisiae* (Destruelle *et al.*, 1995) and *treA* gene product of *Aspergillus nidulans* (d'Enfert and Fontaine, 1997).

The weight matrix method described by von Heijne (1986) showed the presence of a cleavable signal sequence consisting of 19 amino acids at the amino(N)-terminal end of the *Artemia* trehalase (Fig. 2, 3). A cell adhesion motif, Arg-Gly-Asp (reviewed by Ruoslahti and Pierschbacher, 1986, and D'Souza *et al.*, 1991) was found next to the N-terminal cleavable site as shown in Fig. 2. The protein had four potential N-glycosylation sites (Fig. 2). Alignment of animal trehalases showed that the location of the third potential N-glycosylation site was conserved among arthropods (Fig. 3). Trehalase signatures submitted to the PROSITE database (Henrissat and

Bairoch 1993) were found on the deduced amino acid sequences (Fig. 2, 3). A unique, long polypeptide was found at the carboxyl terminus (Fig. 2, 3, 4). The PSORT program (Nakai and Kanehisa, 1992) strongly suggested that the carboxyl(C)-terminal polypeptide contained a transmembrane region (Gly⁶²⁰-Glu⁶⁴⁵, Fig. 2, 4). The predicted transmembrane region was highly hydrophobic as was the N-terminal signal sequence (Fig. 4). A potential cAMP-dependent phosphorylation site was included in a cytoplasmic region of the C-terminal polypeptide (Fig. 2).

Phylogenetic analysis

Phylogenetic trees were drawn by the neighbor-joining method (Saitou and Nei, 1987) (Fig. 5). The results revealed that all the trehalases could be classified into three groups as follows: animal trehalases including the *Artemia* trehalase, trehalases of *E. coli*, and neutral trehalases of fungi (Fig. 5A). Genetic distances between the *Artemia* trehalase and mammalian trehalases were shorter than those between the *Artemia* trehalase and trehalases of insects (Fig. 5B).

Developmental expression of *Artemia* trehalase

Since the nucleotide sequences of *TreE2* and *TreE3* were almost identical to each other, the labelled *TreE3* was used as a probe to detect both types of mRNA.

A 2.6-kb mRNA was detected in all the samples of *A. franciscana* (Fig. 6). The 2.6-kb transcript was abundant in the dormant (0 hr) cysts and embryos at 3 hr and 6 hr of the culture (Fig. 6). The transcript then decreased and was weakly detected in the 30 hr sample. A 5.0-kb transcript was observed after 3 hr, which increased by 6–12 hr and decreased thereafter. The amount of the 5.0-kb transcript was lower than that of the 2.6-kb transcript in all the samples.

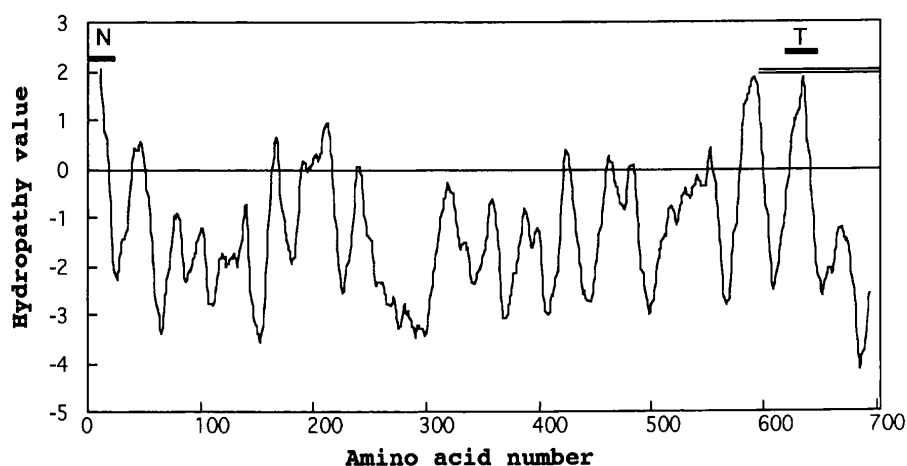


Fig. 4. Hydropathicity and hydrophobicity plot of the *Artemia* trehalase coded by *TreE2*. The hydropathy value was calculated using the method of Engelman *et al.*, (1986). Areas showing negative hydropathy values represent hydrophilic regions. Unique C-terminal region is indicated by double lines. Cleavable N-terminal signal sequence (N) and a transmembrane region (T) predicted by PSORT program are indicated by bold lines.

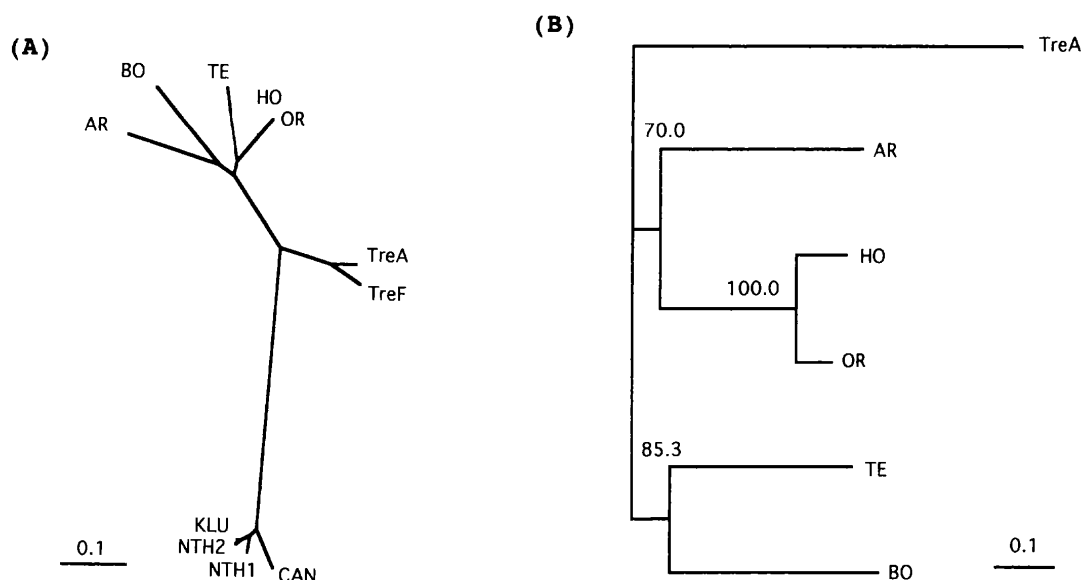


Fig. 5. Phylogenetic relationship among trehalases. A neighbor-joining tree of trehalase is represented in (A). AR, *Artemia* trehalase; HO, human trehalase; OR, rabbit trehalase; TE, mealworm beetle trehalase; BO, silkworm trehalase; TreA, *Escherichia coli* periplasmic trehalase; TreF, *E. coli* cytoplasmic trehalase; NTH1, *Saccharomyces cerevisiae* neutral trehalase; NTH2, protein product of *S. cerevisiae* NTH2 (YBR0106); CAN, *Candida albicans* neutral trehalase; KLU, *Kluyveromyces lactis* neutral trehalase. Detailed neighbor-joining tree among the animal trehalases is represented in (B) by employing TreA of *E. coli* as an outgroup. Scale bar indicates 0.1 of branch length. Values on each branching represents percent of bootstrap probability.

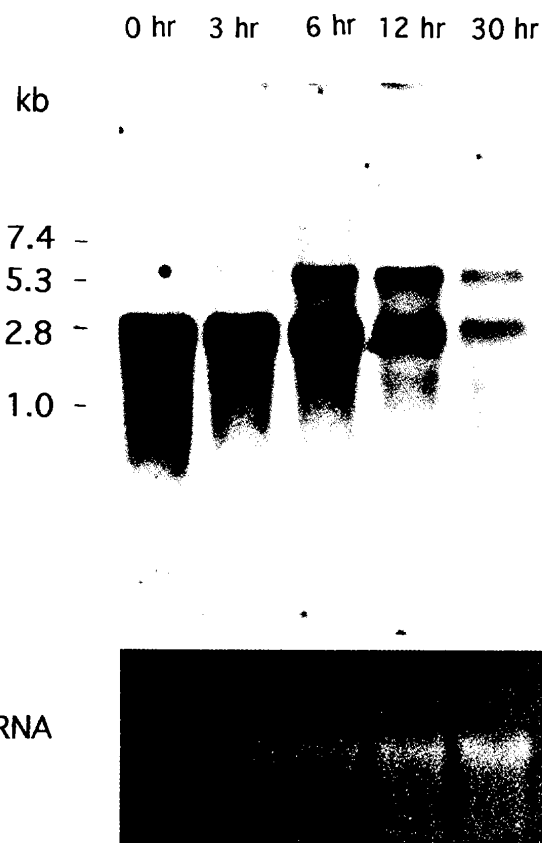


Fig. 6. Ontogenic changes of gene expression of *Artemia* trehalase. Total RNA samples were extracted from *A. franciscana* at indicated periods of post-dormant development, and 20 μ g of each RNA sample was electrophoresed, transferred to nylon membrane, and hybridized with the DIG-labelled PCR probe. Positions of molecular weight markers are indicated. 18S ribosomal RNA is detected by ethidium bromide staining and shown at the bottom.

DISCUSSION

In the present report, we described cloning and characterization of the trehalase cDNAs from crustacean *Artemia*. Most parts these two cDNAs, termed *TreE2* and *TreE3*, were the same and both contained the same ORF of 2109 nucleotides. Seven replacements were found at the nucleotide level between *TreE2* and *TreE3*, resulting in two replacements on coded 703 amino acids. The rate of change at the nucleotide level between *TreE2* and *TreE3* was 0.3%, much lower than those among alleles of Na/K-ATPase α 1 subunit of *A. franciscana* reported by García-Sáez *et al.* (1997). Therefore, *TreE2* and *TreE3* appear to be derived from point-mutated alleles of one trehalase gene rather than from two different trehalase genes. Sequencing analyses of the PCR products showed that the mRNAs were in the ratio of 1:1 (data not shown). As the dried cysts which we used in the present report were from a wild population, rather than from an established strain, variants of trehalase gene might be observed in individuals or in sub-populations.

The deduced amino acid sequence had two "trehalase signatures" (Fig. 2, 3) and showed high similarities to animal trehalases (Fig. 3), revealing itself as a trehalase.

The presence of potential N-glycosylation sites on the predicted peptide is consistent with our previous conclusion that the *Artemia* trehalase is a glycoprotein (Nambu *et al.*, 1997a).

The most significant feature of the *Artemia* trehalase is the unique, long C-terminal polypeptide. The region consisted of 107 amino acids (Ile⁵⁹⁷-Glu⁷⁰³) and had no significant similarity to any other proteins. The predicted transmembrane region (Gly⁶²⁰-Glu⁶⁴⁵) was present on the C-terminal polypep-

tide (Fig. 2, 4) and about 60 amino acids from the C-terminal region were possibly located in the cytoplasm. Membrane-bound trehalases have been reported in silkworm (Azuma and Yamashita, 1985; Takesue *et al.*, 1989) and mammals (Takesue *et al.*, 1986; Ruf *et al.*, 1990; Sasai-Takedatsu *et al.*, 1996; Oesterreicher *et al.*, 1998). The presence of the cell adhesion motif also suggests the association of *Artemia* trehalase to the cell membrane.

The predicted molecular mass of the *Artemia* trehalase without the N-terminal signal sequence was 77,936 Da and substantially larger than that of the purified soluble protein, 70 kDa (Nambu *et al.*, 1997a). The molecular mass of the soluble trehalase decreased to 66 kDa after endoglycosidase H treatment in our previous study (Nambu *et al.*, 1997a). The solubility of the *Artemia* trehalase is incompatible with the presence of the predicted transmembrane region. As similar sequences to the C-terminal polypeptide of the *Artemia* trehalase were not found among other trehalases (Fig. 3), it is probable that the C-terminal polypeptide consisting of 107 amino acids is not necessary for trehalase activity and might be removed. The molecular mass of the deduced *Artemia* trehalase without the N-terminal signal sequence and the C-terminal polypeptide was calculated to be 66,272 Da, showing a good agreement with the molecular mass of the purified *Artemia* trehalase after the endoglycosidase H treatment.

A possible cAMP-dependent phosphorylation site was found on the predicted C-terminal cytoplasmic region of the *Artemia* trehalase, although we do not have any evidences that the *Artemia* trehalase is regulated by cAMP-dependent phosphorylation. The cAMP-dependent phosphorylation site has not been found in other animal trehalases. In yeasts, cAMP-dependent phosphorylation sites in the N-terminal region and cAMP-dependent regulation of yeast trehalase have been reported (Kopp *et al.*, 1993, 1994; Amaral *et al.*, 1997; Eck *et al.*, 1997). Although the site may be removed by the post-translational processing discussed above, it is possible that the potential cAMP-dependent phosphorylation site plays a role in an immature *Artemia* trehalase.

Phylogenetic trees of trehalase are shown in Fig. 5. On the phylogenetic tree of animal trehalase, the *Artemia* trehalase was closely related to the mammalian trehalases (Fig 5B). The distributions and the physiological meanings of the arthropod trehalases were reported to be different from each other (Azuma *et al.*, 1985; Takesue *et al.*, 1989; Takiguchi *et al.*, 1992; Yaginuma *et al.*, 1996; Nambu *et al.*, 1997a). Adaptation of each trehalases to different roles and physiological environments might result in their variety in amino acid sequences.

The results of Northern blot analysis confirmed the presence of the two transcripts previously described (Fig. 6). The 2.2 kb mRNA of the *Artemia* trehalase found in the previous report (Nambu *et al.*, 1997b) was not detectable when we changed a batch of the cysts.

Three working hypotheses on the two transcripts might be proposed as follows. First, the transcripts of *Artemia* trehalase might originated from alternative splicing of one gene.

The second possibility is that the 5.0 kb transcript is a precursor and processed into the 2.6 kb transcript. The third possibility is that the two transcripts are derived from two highly homologous genes.

The abundance of the 2.6 kb transcript in the dormant cysts of *A. franciscana* suggests its importance in early period of the post-dormant development. It may be called as a cryptic RNA. It was observed that 18S rRNA gradually increased during development (Fig. 6). This means that the ratio of rRNA to total RNA is relatively low in the dormant cysts. This observation may be partly due to accumulations of cryptic RNA, such as trehalase mRNA, in the dormant cysts.

The activity of the trehalase was low in the hydrated cysts and began to increase after 6 hr in culture (Nambu *et al.*, 1997a) in spite of the abundant presence of the 2.6 kb transcript in the dormant cysts. The discrepancy between the amount of the mRNA and the trehalase activity suggests a probable presence of post-transcriptional regulation of the enzyme. Similar inconsistency between the amount of mRNA and protein products have been reported in the early period of post-dormant development of *Artemia* (Fisher *et al.*, 1986; Díaz-Guerra *et al.*, 1989).

The *Artemia* embryos have mechanisms inactivating protein synthesis during their dormancy. The presence of protein synthesis inhibitor (Warner *et al.*, 1977), a lack of template activity of some mRNAs in the dormant cysts by restricted location of mRNAs into mRNP particles (Grosfeld and Littarue, 1975), and low levels of initiation factors (Sierra *et al.*, 1974) have been reported. Moens and Kondo (1976) reported that the specific template activity of the polysomes in the cysts was low and increased linearly up to 13 hr of incubation. The discrepancy between the amount of mRNA and the trehalase activity in the embryos of *A. franciscana* might be explained by such inhibitions.

The appearance of the 5.0 kb transcript in 3 hr embryos of *A. franciscana* is a sign of initiation of post-dormant transcription of the *Artemia* trehalase. If the 5.0 kb transcript is the precursor of the 2.6 kb transcript, the appearance of the 5.0 kb transcript indicates a supplement of the 2.6 kb transcript. The amount of the 5.0 kb transcript increased in accordance with the increase in the trehalase activity (Nambu *et al.*, 1997a). Therefore, it is still possible that transcriptional regulation of the 5.0 kb mRNA directly controls the trehalase activity.

The structural and functional differences of the two transcripts and the mechanisms of the regulation of the trehalase activity are open for future investigation. Further analysis of transcriptional or translational regulations of trehalase would help us to understand the trehalose metabolism in resuming development of *Artemia*.

ACKNOWLEDGMENTS

We are deeply indebted to Drs. Shimada, Akasaka and Mitsunaga-Nakatsubo, Hiroshima University, for their helpful discussions and supports. Thanks are also offered to Drs. Satoh and Makabe and their co-workers, Kyoto University, for their advice on phyloge-

netic analysis of trehalases.

REFERENCES

- Amaral FC, Van Dijk P, Nicoli JR, Thevelein JM (1997) Molecular cloning of the neutral trehalase gene from *Kluyveromyces lactis* and the distinction between neutral and acid trehalase. *Arch Microbiol* 167: 202–208
- Azuma M, Yamashita O (1985) Immunohistochemical and biochemical localization of trehalase in the developing ovaries of the silkworm, *Bombyx mori*. *Insect Biochem* 15: 589–596
- Clegg JS (1962) Free glycerol in dormant cysts of the brine shrimp *Artemia salina*, and its disappearance during development. *Biol Bull Woods Hole* 123: 295–301
- Clegg JS (1964) The control of emergence and metabolism by external osmotic pressure and the role of free glycerol in developing cysts of *Artemia salina*. *J Exp Biol* 41: 879–892
- Clegg JS (1965) The origin of trehalose and its significance during the formation of encysted dormant embryos of *Artemia salina*. *Comp Biochem Physiol* 14: 135–143
- Crowe JH, Crowe LM, Carpenter JF, Aurell Wistrom C (1987) Stabilization of dry phospholipid bilayers and proteins by sugars. *Biochem J* 242: 1–10
- Crowe JH, Crowe LM, Carpenter JF, Rudolph AS, Aurell Wistrom C, Spargo BJ, Anchordoguy TJ (1988) Interaction of sugars with membranes. *Biochim Biophys Acta* 947: 367–384
- d'Enfert C, Fontaine T (1997) Molecular characterization of the *Aspergillus nidulans treA* gene encoding an acid trehalase required for growth on trehalose. *Mol Microbiol* 24: 203–216
- Destruelle M, Holzer H, Klionsky DJ (1995) Isolation and characterization of a novel yeast gene, *ATH1*, that is required for vacuolar acid trehalase activity. *Yeast* 11: 1015–1025
- Díaz-Guerra M, Quintanilla M, Palmero I, Sastre L, Renart J (1989) Differential expression of a gene highly homologous to *c-ras* during the development of the brine shrimp *Artemia*. *Biochem Biophys Res Comm* 162: 802–808
- D'Souza SE, Ginsberg MH, Plow EF (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *TIBS* 16: 246–250
- Eck R, Bergmann C, Ziegelbauer K, Schönfeld W, Kunkel W (1997) A neutral trehalase gene from *Candida albicans*: molecular cloning, characterization and disruption. *Microbiol* 143: 3747–3756
- Engelman DM, Steitz TA, Goldman A (1986) Identifying nonpolar transbilayer helices into amino acid sequence of membrane proteins. *Annu Rev Biophys Biophys Chem* 6: 120–121
- Engler-Blum G, Meier M, Frank J, Müller GA (1993) Reduction of background problems in nonradioactive Northern and Southern blot analyses enables higher sensitivity than ³²P-based hybridizations. *Anal Biochem* 210: 235–244
- Ewing RD, Clegg JS (1969) Lactate dehydrogenase activity and anaerobic metabolism during embryonic development in *Artemia salina*. *Comp Biochem Physiol* 31: 297–307
- Feldmann H, Aigle M, Aljinovic G *et al.* (1994) Complete DNA sequence of yeast chromosome II. *EMBO J* 13: 5795–5809
- Fisher JA, Baxter-Lowe LA, Hokin LE (1986) Regulation of Na, K-ATPase biosynthesis in developing *Artemia salina*. *J Biol Chem* 261: 515–519
- García-Sáez A, Perona R, Sastre L (1997) Polymorphism and structure of the gene coding for the $\alpha 1$ subunit of the *Artemia franciscana* Na/K-ATPase. *Biochem J* 321: 509–518
- Grosfeld H, Littauer UZ (1975) Cryptic form of mRNA in dormant *Artemia salina* cysts. *Biochem Biophys Res Comm* 67: 176–181
- Gutierrez C, Ardourel M, Bremer E, Middendorf A, Boos W, Ehmann U (1989) Analysis and DNA sequence of the osmoregulated *treA* gene encoding the periplasmic trehalase of *Escherichia coli* K12. *Mol Gen Genet* 217: 347–354
- Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 293: 781–788
- Horlacher R, Uhland K, Klein W, Ehrmann M, Boos W (1996) Characterization of a cytoplasmic trehalase of *Escherichia coli*. *J Bacteriol* 178: 6250–6257
- Ishihara R, Taketani S, Sasai-Takedatsu M, Kino M, Tokunaga R, Kobayashi Y (1997) Molecular cloning, sequencing and expression of cDNA encoding human trehalase. *Gene* 202: 69–74
- Kopp M, Müller H, and Holzer H (1993) Molecular analysis of the neutral trehalase gene from *Saccharomyces cerevisiae*. *J Biol Chem* 268: 4766–4774
- Kopp M, Nwaka S, Holzer H (1994) Corrected sequence of the yeast neutral trehalase-encoding gene (*NTH1*): biological implications. *Gene* 150: 403–404
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning: A Laboratory Manual* (1st ed.). Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York
- Moens L, Kondo M (1976) Polysome-dependent synthesis of embryonic proteins of *Artemia salina* during cell differentiation and analysis of heme-containing protein. *Dev Biol* 49: 457–469
- Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897–911
- Nambu Z, Nambu F and Tanaka S (1997a) Purification and characterization of trehalase from *Artemia* embryos and larvae. *Zool Sci* 14: 419–427
- Nambu Z, Tanaka S, Nambu F (1997b) Gene expression of trehalase during post-dormant development of the brine shrimp, *Artemia*: Comparison of the two species. *J Univ Occupational Environmental Health* 19: 255–264
- Nwaka S, Kopp M, Holzer H (1995) Expression and function of the trehalase genes *NTH1* and *YBR0106* in *Saccharomyces cerevisiae*. *J Biol Chem* 270: 10193–10198
- Nwaka S, Holzer H (1998) Molecular biology of trehalose and the trehalases in the yeast *Saccharomyces cerevisiae*. *Prog Nuc Acid Res Mol Biol* 58: 197–237
- Oesterreicher TJ, Nanthakumar NN, Winston JH, Henning SJ (1998) Rat trehalase: cDNA cloning and mRNA expression in adult rat tissues and during intestinal ontogeny. *Am J Physiol* 274: R1220–R1227
- Ruf J, Wacker H, James P, Maffia M, Seiler P, Galand G, von Kieckebusch A, Semenza G, Mantei N (1990) Rabbit small intestinal trehalase. Purification, cDNA cloning, expression, and verification of glycosylphosphatidylinositol anchoring. *J Biol Chem* 265: 15034–15039
- Ruoslahti E, Pierschbacher MD (1986) Arg-Gly-Asp: A versatile cell recognition signal. *Cell* 44: 517–518
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (3rd ed.). Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York
- Sasai-Takedatsu M, Taketani S, Nagata N, Furukawa T, Tokunaga R, Kojima T, Kobayashi Y (1996) Human trehalase: Characterization, localization, and its increase in urine by renal proximal tubular damage. *Nephron* 73: 179–185
- Sierra JM, Meier D, Ochoa S (1974) Effect of development on the translation of messenger RNA in *Artemia salina* embryos. *Proc Natl Acad Sci USA* 71: 2693–2697
- Su Z-H, Sato Y, Yamashita O (1993) Purification, cDNA cloning and Northern blot analysis of trehalase of pupal midgut of the silkworm, *Bombyx mori*. *Biochim Biophys Acta* 1173: 217–224.
- Su Z-H, Ikeda M, Sato Y, Saito H, Imai K, Isobe M, Yamashita O (1994) Molecular characterization of ovary trehalase of the silkworm, *Bombyx mori* and its transcriptional activation by diapause hormone. *Biochim Biophys Acta* 1218: 366–374
- Takesue Y, Yokota K, Nishi Y, Taguchi R, Ikezawa H (1986) Solubilization of trehalase from rabbit renal and intestinal brush-border membranes by a phosphatidylinositol-specific phospholipase C.

- FEBS Lett 201: 5–8
- Takesue Y, Yokota K, Miyajima S, Taguchi R, Ikezawa H (1989) Membrane anchors of alkaline phosphatase and trehalase associated with the plasma membrane of larval midgut epithelial cells of the silkworm, *Bombyx mori*. J Biochem (Tokyo) 105: 998–1001
- Takiguchi M, Niimi T, Su Z-H, Yaginuma T (1992) Trehalase from male accessory gland of an insect, *Tenebrio molitor*. DNA sequencing and developmental profile of the gene expression. Biochem J 288: 19–22
- Thevelein JM (1984) Regulation of trehalose mobilization in fungi. Microbiol Rev 48: 42–59
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc Acids Res 22: 4673–4680
- von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nuc Acids Res 14: 4683–4690
- Warner AH, Shridhar V, Finamore FJ (1977) Isolation and partial characterization of a protein synthesis inhibitor from brine shrimp embryos. Can J Biochem 55: 965–974
- Wolfe KH, Lohan AJE (1994) Sequence around the centromere of *Saccharomyces cerevisiae* chromosome II: Similarity of *CEN2* to *CEN4*. Yeast 10: S41–S46
- Yaginuma T, Mizuno T, Mizuno C, Ikeda M, Wada T, Hattori K, Yamashita O, Happ GM (1996) Trehalase in the spermatophore from the bean-shaped accessory gland of the male mealworm beetle, *Tenebrio molitor*: purification, kinetic properties and localization of the enzyme. J Comp Physiol B 166: 1–10

(Received September 17, 1998 / Accepted December 14, 1998)