

The occurrence of hemocyanin in Hexapoda

Christian Pick, Marco Schnerer and Thorsten Burmester

Institute of Zoology and Zoological Museum, University of Hamburg, Germany

Keywords

evolution; hemocyanin; hexamerin; insect; oxygen

Correspondence

T. Burmester, Institute of Zoology and Zoological Museum, University of Hamburg, Martin-Luther-King-Platz 3, D-20146 Hamburg, Germany
Fax: +49 40 42838 3937
Tel: +49 40 42838 3913
E-mail: thorsten.burmester@uni-hamburg.de

Database

The nucleotide sequences reported in this paper have been submitted to the EMBL/GenBank databases under the accession numbers FM242638 to FM242654

(Received 15 October 2008, revised 5 January 2009, accepted 21 January 2009)

doi:10.1111/j.1742-4658.2009.06918.x

Hemocyanins are copper-containing, respiratory proteins that have been thoroughly studied in various arthropod subphyla. Specific O₂-transport proteins have long been considered unnecessary in Hexapoda (including Insecta), which acquire O₂ via an elaborate tracheal system. However, we recently identified a functional hemocyanin in the stonefly *Perla marginata* (Plecoptera) and in the firebrat *Thermobia domestica* (Zygentoma). We used RT-PCR and RACE experiments to study the presence of hemocyanin in a broad range of ametabolous and hemimetabolous hexapod taxa. We obtained a total of 12 full-length and 5 partial cDNA sequences of hemocyanins from representatives of Collembola, Archeognatha, Dermaptera, Orthoptera, Phasmatodea, Mantodea, Isoptera and Blattaria. No hemocyanin could be identified in Protura, Diplura, Ephemeroptera, Odonata, or in the Eumetabola (Holometabola + Hemiptera). It is not currently known why hemocyanin has been lost in some taxa. Hexapod hemocyanins usually consist of two distinct subunit types. Whereas type 1 subunits may represent the central building block, type 2 subunits may be absent in some species. Phylogenetic analyses support the Pancrustacea hypothesis and show that type 1 and type 2 subunits diverged before the emergence of the Hexapoda. The copperless insect storage hexamerins evolved from hemocyanin type 1 subunits, with *Machilis germanica* (Archeognatha) hemocyanin being a possible 'intermediate'. The evolution of hemocyanin subunits follows the widely accepted phylogeny of the Hexapoda and provides strong evidence for the monophyly of the Polyneoptera (Plecoptera, Dermaptera, Orthoptera, Phasmatodea, Mantodea, Isoptera, Blattaria) and the Dictyoptera (Mantodea, Isoptera, Blattaria). The Blattaria are paraphyletic with respect to the termites.

Hemocyanins are respiratory proteins that float freely dissolved in the hemolymph of many arthropod species [1–4]. They are composed of six identical or similar subunits with molecular masses of around 75 kDa [1,3]. A subunit may bind to an O₂ molecule by means of two Cu⁺ ions, each of which is coordinated by three histidines in two distinct binding sites. Some hemocyanins assemble into large oligomers of up to 8 × 6 subunits [1]. The occurrence and properties of hemocyanins have been thoroughly studied over the last 30 years in Chelicerata and malacostracan Crustacea, but their presence in other arthropod subphyla (Onychophora, Myriapoda and Hexapoda) has been discovered only recently [5–8].

In most Hexapoda, gas exchange is mediated by the tracheal system, a network of tubules that open to the atmosphere on the cuticle and radiate to all parts of the body. O₂ is delivered through trachea and tracheoles in the gaseous phase [9] and hence respiratory proteins have long been considered unnecessary [10–12]. Nevertheless, a functional hemocyanin has been identified in the hemolymph of the stonefly *Perla marginata* [8]. This hemocyanin consists of two distinct subunit types (PmaHc1 and PmaHc2) [8] and orthologous sequences have been reported from the closely related stonefly *Perla grandis* (PgrHc1 and PgrHc2) [13]. We recently identified a hemocyanin in the hemolymph of adult firebrat *Thermobia domestica*

(Zygentoma), which also consists of two distinct subunits (TdoHc1 and TdoHc2) [14]. A hemocyanin-like protein from in the embryonic hemolymph of the grasshopper *Schistocerca americana* ('embryonic hemolymph protein', EHP) [15] resembles hemocyanin subunit 1 (Hc1), suggesting that this protein might have a respiratory function as well [8].

Arthropod hemocyanins belong to a protein superfamily that also comprises arthropod phenoloxidases, crustacean pseudohemocyanins, insect storage hexamerins and dipteran hexamerin receptors [4,16–19]. Respiratory hemocyanins most likely evolved from the phenoloxidases early in arthropod evolution. Thus the phenoloxidases, which had been identified in various crustaceans and hexapods, form the sistergroup of all other members of the arthropod hemocyanin superfamily [4,18]. Crustacean pseudohemocyanins and insect hexamerins are non-respiratory proteins that evolved independently from hemocyanins [4].

Although hexamerins might be ubiquitous in insects [14,19–21], hemocyanin appears to be missing in eumetabolous insects [22]. Here we investigate representatives from several ametabolous and hemimetabolous hexapod orders for the presence of hemocyanin, including Collembola (springtails), Diplura (diplurans), Protura (proturans), Archeognatha (bristletails), Ephemeroptera (mayflies), Odonata (dragonflies and damselflies), Orthoptera (grasshoppers and crickets), Phasmatodea (stick insects), Dermaptera (earwigs), Mantodea (mantises), Isoptera (termites) and Blattaria (cockroaches), as well as Hemiptera (true bugs).

Results

Identification of hexapod hemocyanins

We used an alignment of insect hemocyanin sequences to deduce two pairs of degenerated oligonucleotide primer, which we applied on cDNA from various hexapod species (Table 1). Products of the expected lengths were sequenced and BLAST searches were performed. We identified fragments that correspond to insect hemocyanin subunit types 1 from springtails *Sinella curviseta* (ScuHc1) and *Folsomia candida* (FcaHc1), bristletail *Machilis germanica* (MgeHc1), stick insect *Carausius morosus* (CmoHc1), grasshopper *Locusta migratoria* (LmiHc1), earwig *Chelidurella acanthopygia* (CacHc1), mantis *Hierodula membranacea* (HmeHc1), termite *Cryptotermes secundus* (CseHc1) and cockroaches *Blattella dubia* (BduHc1), *Periplaneta americana* (PamHc1) and *Shelfordella lateralis* (SlaHc1). In the other species, no hemocyanin sequence was recovered. The same two pairs of degenerated primers also resulted in fragments that correspond to insect hemocyanin subunit types 2, which were found for *Ch. acanthopygia* (CacHc2), *H. membranacea* (HmeHc2), *Cr. secundus* (CseHc2), *B. dubia* (BduHc2), *P. americana* (PamHc2) and *Sh. lateralis* (StaHc2).

Hexapod hemocyanin subunits 1

We completed the fragments of ScuHc1, MgeHc1, CmoHc1, CacHc1, HmeHc1, CseHc1, BduHc1 and PamHc1 using 5'- and 3'-RACE (Table 2). The full-

Table 1. Hexapod species used in this study.

Species	Order	Family	Developmental stage	Hc1	Hc2
<i>Sinella curviseta</i>	Collembola	Entomobryidae	Juvenile, adult	ScuHc1	–
<i>Folsomia candida</i>	Collembola	Isotomidae	Juvenile, adult	FcaHc1	–
<i>Allacma fusca</i>	Collembola	Sminthuridae	Adult	–	–
<i>Acerentomon franzi</i>	Protura	Acerentomidae	Juvenile	–	–
<i>Campodea</i> sp.	Diplura	Campodeidae	Juvenile, adult	–	–
<i>Machilis germanica</i>	Archeognatha	Machilidae	Adult	MgeHc1	–
<i>Ephemerella mucronata</i>	Ephemeroptera	Ephemerellidae	Juvenile	–	–
<i>Aeshna cyanea</i>	Odonata	Aeshnidae	Adult	–	–
<i>Locusta migratoria</i>	Orthoptera	Acrididae	Adult	LmiHc1	–
<i>Acheta domesticus</i>	Orthoptera	Gryllidae	Adult	–	–
<i>Carausius morosus</i>	Phasmatodea	Heteronemiidae	Adult	CmoHc1	–
<i>Chelidurella acanthopygia</i>	Dermaptera	Forficulidae	Adult	CacHc1	CacHc2
<i>Hierodula membranacea</i>	Mantodea	Mantidae	Juvenile	HmeHc1	HmeHc2
<i>Cryptotermes secundus</i>	Isoptera	Kalotermitidae	Juvenile, adult	CseHc1	CseHc2
<i>Blattella dubia</i>	Blattaria	Blaberidae	Juvenile, adult	BduHc1	BduHc2
<i>Periplaneta americana</i>	Blattaria	Blattidae	Juvenile, adult	PamHc1	PamHc2
<i>Shelfordella lateralis</i>	Blattaria	Blattidae	Juvenile, adult	SlaHc1	SlaHc2
<i>Graphosoma lineatum</i>	Hemiptera	Pentatomoidae	Adult	–	–

Table 2. Molecular properties of the putative hemocyanin cDNA and the deduced amino acid sequences. Sequences are given in Figs S1–S17.

Name	Accession no.	Nucleotide				Deduced amino acid sequence (aa)	Putative signal peptide (aa)	Native protein (aa)	Predicted molecular mass (kDa)
		cDNA (bp)	5'-UTR (bp)	ORF (bp)	3'-UTR (bp)				
ScuHc1	FM242638	2178	37	2016	125	672	19	653	75.59
FcaHc1	FM242650	1053	–	–	–	351	–	–	–
MgeHc1	FM242639	2208	30	2058	120	686	16	670	79.18
CmoHc1	FM242640	2310	87	2028	195	676	19	657	76.29
LmiHc1	FM242651	530	–	–	–	176	–	–	–
CacHc1	FM242641	2326	245	2007	74	669	19	650	75.43
HmeHc1	FM242642	2592	98	1980	514	660	None	660	76.97
CseHc1	FM242644	2305	25	2031	249	677	20	657	76.70
BduHc1	FM242646	2118	8	2034	76	678	19	659	77.00
PamHc1	FM242648	2844	26	2022	796	674	19	655	76.58
SlaHc1	FM242652	530	–	–	–	176	–	–	–
CacHc2	FM242654	1506	–	–	138	456	–	–	–
HmeHc2	FM242643	2454	86	1989	379	663	None	663	76.72
CseHc2	FM242645	2293	29	2052	212	684	19	665	76.10
BduHc2	FM242647	2171	16	2055	100	685	19	666	76.11
PamHc2	FM242649	2334	49	2052	233	684	19	665	76.49
SlaHc2	FM242653	527	–	–	–	175	–	–	–

length cDNA sequences comprise 2118–2844 bp and cover ORFs of 1983–2058 bp. The deduced amino acid sequences consist of 660–686 amino acids. Computer analysis suggests the presence of a typical signal peptide for transmembrane transport and export into the hemolymph [23] in all subunits except HmeHc1 (Fig. 1). Therefore, the native proteins consist of 650–670 amino acids with predicted molecular masses of 75.43–79.59 kDa. The amino acid sequences are 53.9–68.0% identical with hemocyanin subunit type 1 from *P. marginata* (PmaHc1; Table 3). The six histidine residues crucial for oxygen binding are strictly conserved in all hemocyanin proteins and a potential N-glycosylation site (NXS/T), located in PmaHc1 at Asn191, is present in all type 1 subunits (Fig. 1).

Hexapod hemocyanin subunits 2

The 5'- and 3'-ends of HmeHc2, CseHc2, BduHc2 and PamHc2 were obtained using RACE experiments (Table 2). We were able to amplify the 3'-end of CacHc2, but did not succeed with the 5'-end. The full-length cDNA sequences comprise 2171–2454 bp with ORFs of 2171–2454 bp. The deduced amino acid sequences cover 663–685 amino acids and putative signal peptides were found in all proteins except HmeHc2 (Fig. 1). Therefore, the native proteins consist of 663–666 amino acids with predicted molecular masses of 76.11–76.72 kDa. The amino acid sequences are 58.9–62.3% identical with respect to the hemocyanin

subunit type 2 of *P. marginata* (PmaHc2; Table 3). The six histidine residues crucial for oxygen binding are strictly conserved. A potential N-glycosylation site (NXS/T), found in PmaHc2 at position Asn334, is conserved in all subunit types (Hc1 and Hc2) with the exception of PmaHc1. An insertion of nine amino acids in PamHc2 starting at amino acid 435 is unique to subunit types 2. On the amino acid level, the hemocyanin subunits types 2 are 45.0–54.6% identical to the subunit types 1.

Molecular evolution of hexapod hemocyanins

A multiple alignment was constructed using the deduced amino acid sequences of the putative hemocyanin subunits and the previously identified insect hemocyanin subunit types 1 (PmaHc1, PgrHc1, SamE-HP, TdoHc1 and LsaHc1) and types 2 (PmaHc2, PgrHc2, TdoHc2 and LsaHc2). We also included selected insect hexamerins, crustacean hemocyanins, crustacean pseudohemocyanins, chelicerate hemocyanins, myriapod hemocyanins and one onychophoran hemocyanin in the final alignment (Fig. S18). The phylogenetic tree reconstructions were carried out using MRBAYES (Fig. 2) and rerun after exclusion of the incomplete sequences (i.e. FcaHc1, LsaHc1, LsaHc2, CacHc2, LmiHc1, SlaHc1 and SlaHc2) using MRBAYES and PHYML, respectively. In each case, the onychophoran hemocyanin was used to root the tree for visualization purpose.

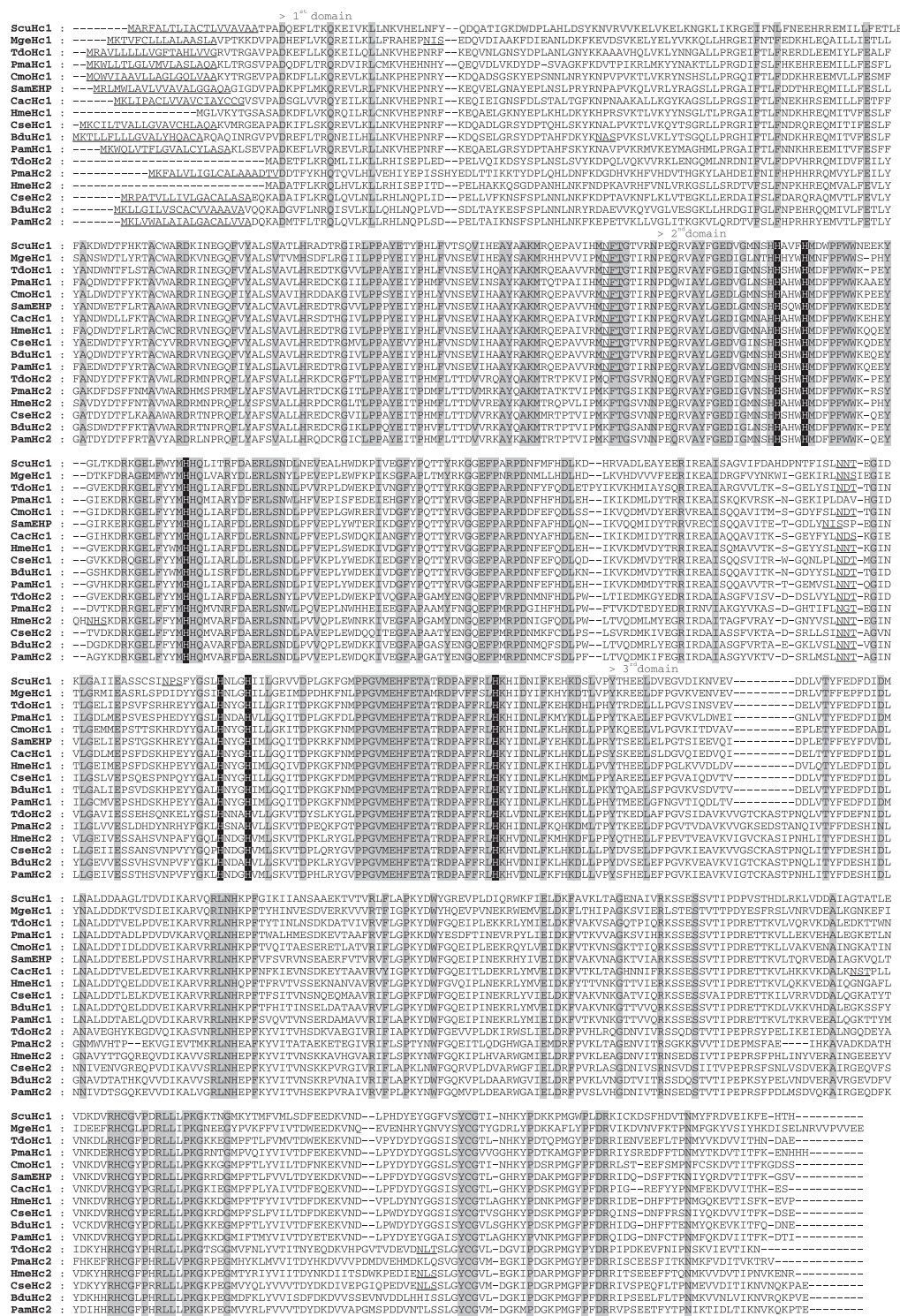


Fig. 1. Multiple alignment of hexapod hemocyanin sequences. Putative hemocyanins from *S. curviseta* (ScuHc1), *M. germanica* (MgeHc1), *C. morosus* (CmoHc1), *Ch. acanthopygia* (CacHc1), *H. membranacea* (HmeHc1 and HmeHc2), *Cr. secundus* (CseHc1 and CseHc2), *B. dubia* (BduHc1 and BduHc2) and *P. americana* (PamHc1 and PamHc2) were compared with the previously identified insect hemocyanins from *T. domestica* (TdoHc1 and TdoHc2) and *P. marginata* (PmaHc1 and PmaHc2). The copper-binding histidines are shaded in black; other strictly conserved residues are shaded in gray. Putative signal peptides and potential N-glycosylation sites (NXS/T) are underlined. The borders of the three structural domains are indicated.

Table 3. Comparison of hexapod hemocyanins. Percent identities between hemocyanins were calculated from nucleotide (above diagonal) and amino acid sequences (below). A detailed comparison of the three domains is given in Table S3.

	Scu Hc1	MgeHc1	Tdo Hc1	Pma Hc1	Cmo Hc1	Sam EHP	Cac Hc1	Hme Hc1	Cse Hc1	Bdu Hc1	Pam Hc1	Tdo Hc2	Pma Hc2	Hme Hc2	Cse Hc2	Bdu Hc2	Pam Hc2
ScuHc1	-																
MgeHc1	56.0	-															
TdoHc1	64.2	57.6	-														
PmaHc1	58.9	53.9	65.5	-													
CmoHc1	62.3	55.4	73.2	67.2	-												
SamEHP	60.6	54.3	72.9	65.3	74.8	-											
CacHc1	61.4	56.2	73.9	67.3	72.9	72.0	-										
HmeHc1	61.5	56.0	73.3	67.8	74.8	74.6	74.3	-									
CseHc1	62.6	55.0	73.8	66.7	75.1	74.2	71.5	76.1	-								
BduHc1	64.0	56.7	76.2	68.0	76.2	75.4	76.0	81.2	80.8	-							
PamHc1	64.2	54.7	73.6	67.8	75.3	74.3	74.2	80.2	79.4	82.6	-						
TdoHc2	53.6	48.9	56.2	51.3	52.8	51.9	53.0	53.2	53.0	54.1	54.0	-					
PmaHc2	49.4	45.0	49.4	48.2	48.2	47.3	46.8	47.1	48.6	48.4	47.4	59.9	-				
HmeHc2	50.9	46.0	53.1	48.7	52.2	50.0	49.4	52.2	51.9	52.4	52.1	66.0	60.3	-			
CseHc2	51.6	47.9	53.6	51.3	53.5	50.8	51.4	53.3	52.6	53.5	53.5	66.8	62.3	72.5	-		
BduHc2	52.9	48.9	54.0	50.4	54.1	50.8	51.1	53.0	54.0	54.6	53.3	68.1	61.3	74.4	73.3	-	
PamHc2	52.2	47.3	52.5	49.8	52.6	50.8	51.2	53.3	52.6	53.7	52.6	67.2	61.7	79.4	74.1	77.3	-

In all analyses, CacHc1, CmoHc1, LmiHc1, HmeHc1, CseHc1, BduHc1, PamHc1 and SlaHc1 form a well-supported monophyletic clade with the previously identified insect hemocyanin subunit types 1 (1.00 posterior probability; 100% bootstrap support) (Fig. 2). The collembolan hemocyanins ScuHc1 and FcaHc1 join this clade, albeit with lower support values (0.77 posterior probability; 66% bootstrap support); after the exclusion of incomplete sequences, however, the posterior probability was higher (0.91). MgeHc1 groups with the insect hexamerins (0.99 posterior probability; 64% bootstrap support), which form the sistergroup of the insect hemocyanin subunits 1. However, BduHc2, PamHc2, StaHc2, CseHc2, HmeHc2 and CacHc2 group with the previously identified insect hemocyanin subunit types 2 (1.00 posterior probability; 100% bootstrap support). Crustacean hemocyanins and pseudohemocyanins form a third clade with 1.00 posterior probability and 100% bootstrap support. The monophyly of crustacean and hexapod hemocyanins and hemocyanin-related proteins is highly supported (1.00 posterior probability; 100% bootstrap support). However, the relationships among the three clades of (a) crustacean proteins, (b) hexapod hemocyanin subunits 1 + hexamerins and (c) hexapod hemocyanin subunits 2 are not well resolved.

Within hemocyanin subunit types 1, the dictyopteran sequences (HmeHc1, CseHc1, BduHc1, PamHc1 and SlaHc1) are monophyletic (1.00 posterior probability; 96% bootstrap support) (Fig. 2). Within this clade, PamHc1 + SlaHc1 (Blattaria, Blattidae) and CseHc1 (Isoptera) form a monophylum (1.00 posterior probability; 82% bootstrap support), which is the sistergroup to BduHc1 (Blaberidae, Blattidae). The orthopteran subunit types 1 (SamEHP + LmiHc1) and CmoHc1 (Phasmatodea) form a well-supported common clade (1.00 posterior probability; 68% bootstrap support), which is in a sistergroup position to the dictyopteran subunits. The hemocyanins from Dermaptera (CacHc1) and Plecoptera (PgrHc1 + PmaHc1) are sistergroups (0.93 posterior probability; 47% bootstrap support). This clade is at the basal position within the Pterygota. The hemocyanins from Zygentoma (TdoHc1 + LsaHc1) form the sistergroup of the pterygote proteins. ScuHc1 + FcaHc1 (Collembola) is basal to the ectognathan subunits, whereas MgeHc1 (Archeognatha) is the sistergroup to the dicondylid hexamerins. Within the hemocyanin subunit types 2, phylogeny resembles that of subunit types 1 except that partial CacHc2 (Dermaptera) is at the basal position within the Pterygota.

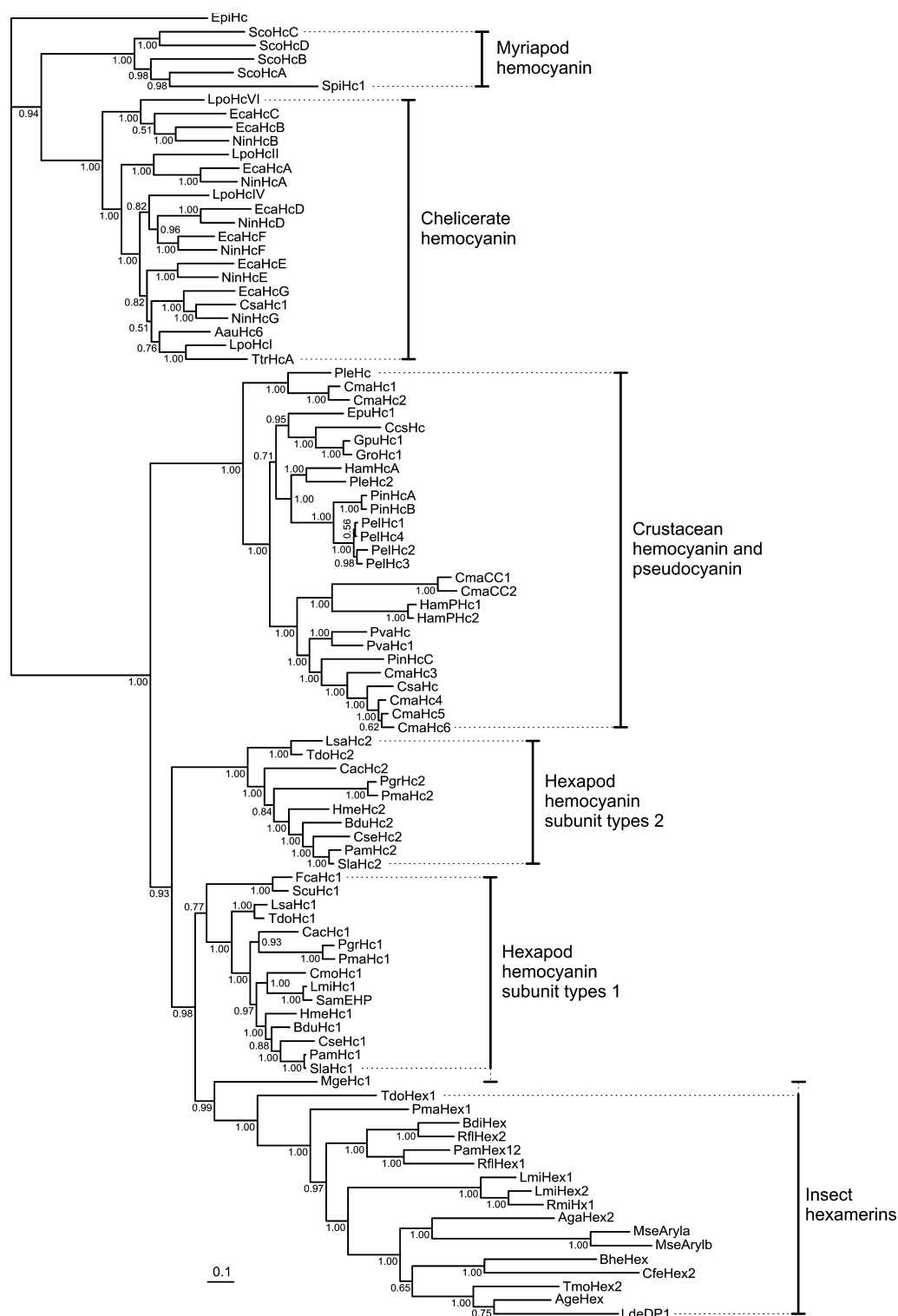


Fig. 2. Bayesian analysis of arthropod hemocyanins and hemocyanin-related proteins. A phylogenetic tree was deduced from a multiple alignment of the putative hemocyanin subunits and the previously identified insect hemocyanin subunit types 1 and 2, selected insect hexamerins, crustacean hemocyanins, crustacean pseudohemocyanins, chelicerate hemocyanins, myriapod hemocyanins and one onychophoran hemocyanin. The onychophoran hemocyanin (EpiHc1) was used to root the tree for visualization purpose. Posterior probabilities are depicted at the nodes; bar = 0.1 substitutions per site.

Discussion

Occurrence of hemocyanin in Hexapoda

Because Hexapoda usually possess a well-developed tracheal system, the presence of respiratory proteins has long been considered unnecessary in this arthropod subphylum. Only a few species that live under hypoxic conditions, represented by the aquatic larvae of the chironomid midges, some aquatic backswimmers or the larvae of the horse botfly, were regarded as exceptions [22,24]. However, a functional hemocyanin has been identified in the hemolymph of the stonefly *P. marginata* [8]. Plecoptera possess a typical tracheal system, but the presence of hemocyanin had been attributed to their semiaquatic lifecycles [8]. More recently, we also identified a putative hemocyanin in the hemolymph of the terrestrial firebrat *T. domestica* (Zygentoma), suggesting a more widespread occurrence of hemocyanin in Hexapoda [14,22]. We decided to investigate a broad range of hexapod orders for the presence of hemocyanin mRNA (Table 1). These taxa represent the majority of ametabolous and hemimetabolous hexapod orders. Embioptera (web spinners), Grylloblattodea (ice bugs), Mantophasmatodea (heel walkers) and the enigmatic Zoraptera could not be obtained for our studies.

Hemocyanins were identified in Collembola, Archeognatha, Zygentoma, Plecoptera, Dermaptera, Orthoptera, Phasmatodea, Mantodea, Isoptera and Blattaria, but not in Protura, Diplura, Ephemeroptera, Odonata and the Eumetabola (Holometabola + Hemiptera) (Fig. 3). In addition, SDS/PAGE with hemolymph samples from Ephemeroptera and Odonata does not provide any indication of the presence of hemocyanin (data not shown). The notion of the absence of hemocyanins from Holometabola is corroborated by the fact that no hemocyanin sequences could be identified in the genomes or expressed sequence tags of various holometabolous insects, such as *Drosophila melanogaster* (Diptera), *Bombyx mori* (Lepidoptera), *Apis mellifera* (Hymenoptera) or *Tribolium castaneum* (Coleoptera). Therefore, it is very likely that hemocyanins are missing in all eumetabolous insects [22]. Hemocyanins might have also been lost in the ametabolous and hemimetabolous hexapods *Allacma fusca* (Collembola), *Acerentomon franzi* (Protura), *Campodea* sp. (Diplura), *Ephemerella mucronata* (Ephemeroptera), *Aeshna cyanea* (Odonata) and *Acheta domesticus* (Orthoptera) as well. However, we cannot exclude that in these species hemocyanins are only expressed under certain environmental conditions or in some developmental stages.

Putative function of hexapod hemocyanins

Reversible binding of oxygen and hence function as a respiratory protein has been unequivocally demonstrated for *P. marginata* hemocyanin [8]. The stonefly hemocyanin binds oxygen with a half-saturation pressure (P_{50}) of ~ 8 torr and shows moderate cooperativity. In our studies, O_2 -binding kinetics could not be measured because of the small size of most specimens. However, we assume respiratory functions for all hexapod hemocyanins identified here because: (a) the six histidines crucial for oxygen binding are strictly conserved, and (b) all subunits are orthologous to the respective subunits of *P. marginata*, with the exception of HgeHc1 (see below). Other or additional functions of insect hemocyanins, such as a role as storage or immune proteins, or as functional phenoloxidase cannot be formally excluded, but are less likely.

In contrast to some hemoglobins, all known hemocyanins are not included in blood cells, but occur freely dissolved in the hemolymph. Signal peptides required for transmembrane transport [23] are present in both plecopteran subunits and the localization of hemocyanin in the hemolymph has been unequivocally demonstrated [8]. Putative signal peptides are also present in the newly identified hexapod hemocyanin subunits (except of those from *H. membranacea*; Fig. 1) and therefore a transport of the nascent polypeptide into the hemolymph is likely. Interestingly, signal peptides are absent in both subunit types from the mantis *H. membranacea* (HmeHc1 and HmeHc2), as well as in the subunit type 2 from the firebrat *T. domestica* [14]. Localization in the hemolymph has been demonstrated for the latter species, suggesting export from the cell by other means [14]. Whether this also applies to HmeHc1 and HmeHc2 must remain uncertain. There is obviously no correlation between loss of signal peptides and proteins phylogeny (Fig. 2). Therefore, the signal peptides may have been lost at least three times independently during evolution of insect hemocyanins, but the functional relevance is currently unknown.

Subunit evolution and emergence of insect hexamerins

The plecopteran hemocyanin consists of two distinct subunits (Hc1 and Hc2) that assemble into a hexamer of ~ 460 kDa in unknown stoichiometry [8]. Orthologous subunit types have been identified in the Zygentoma and hence their diversification preceded the emergence of pterygote insects [14]. Therefore, it is not surprising that both subunit types are also present in Dermaptera, Mantodea, Isoptera and Blattaria. Hemo-

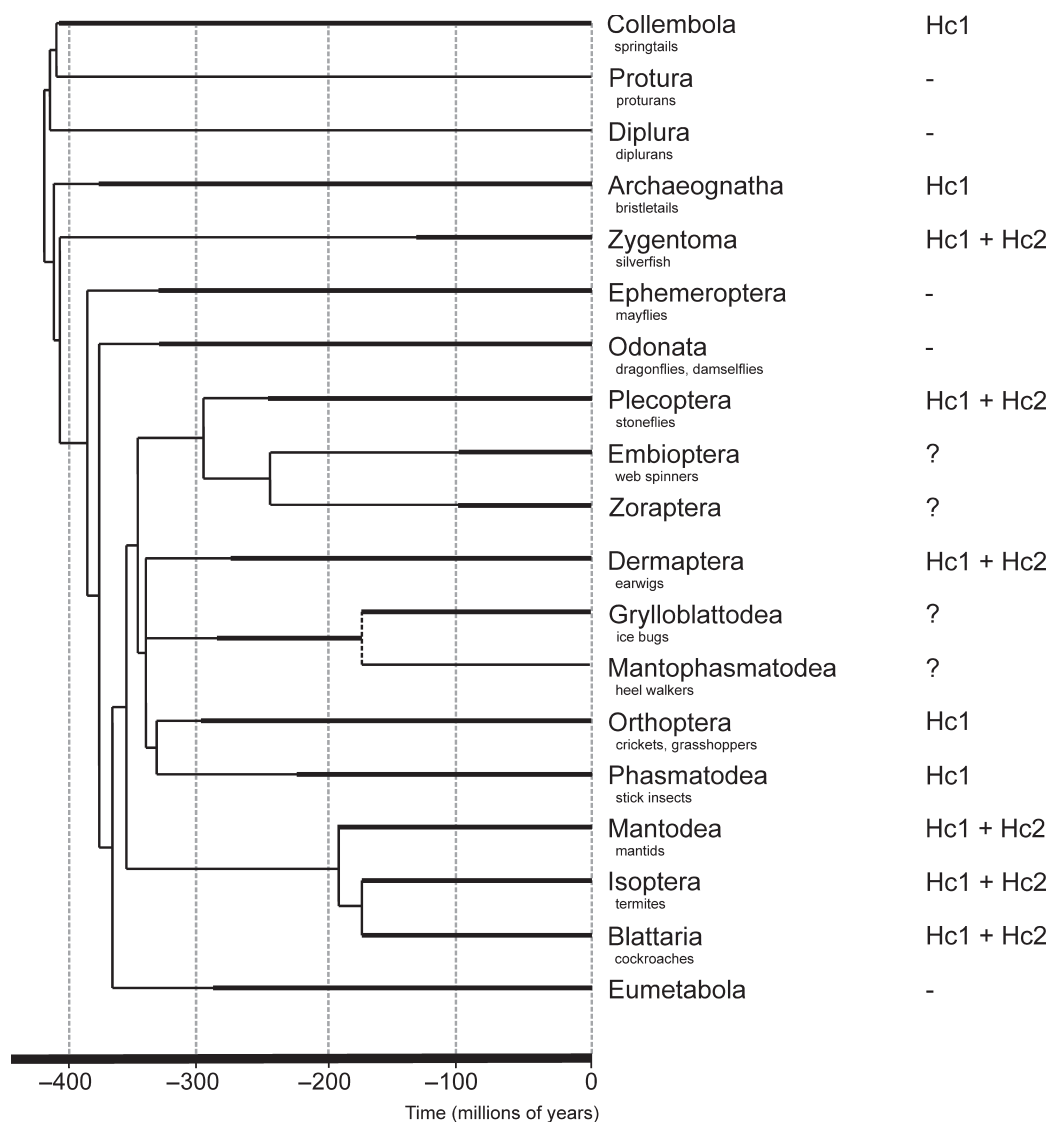


Fig. 3. Occurrence of both hemocyanin subunit types in Hexapoda. The phylogenetic tree of the hexapod orders and the times of origins were taken from Grimaldi & Engel [48].

cyanin subunit type 2, however, is apparently missing in the grasshopper *Sch. americana* [15] and appears to be absent in other orthopterids (Phasmida + Orthoptera) as well.

Because Collembola (springtails) possess a distinct subunit type 1, both subunit types must have separated before the emergence of extant hexapod orders (Figs 2 and 3). These data also imply that hemocyanin subunit type 2 might have been lost several times independently in Hexapoda. However, in none of the species did we observe only hemocyanin subunit type 2, with subunit type 1 being absent. Therefore, subunit type 1 appears to represent the central build-

ing block of a hexapod hemocyanin, whereas subunit type 2 may have modifying functions or represent a distinct hemocyanin hexamer. The presence of multiple hemocyanin subunit types may enable a more sophisticated allosteric and pH-dependent regulation of O₂ binding.

The putative hemocyanin from the bristletail *M. germanica* (MgeHc1) shows the highest amino acid identity with hemocyanin subunit type 1 from the firebrat *T. domestica* (57.6%; Table 3). Phylogenetic analyses, however, strongly suggest that MgeHc1 is basal to the hexamerins of the dicondylid insects (Fig. 2). Hemocyanins and hexamerins share many characteristics in

terms of structure but due to the loss of Cu-binding histidine residues hexamerins do not bind oxygen. Hexamerins are thought to act mainly as storage proteins for non-feeding periods [20,21]. In contrast to any known hexamerin, in MgeHc1 all six histidine residues are preserved. Therefore, MgeHc1 is in an 'intermediate' position, being structurally a hemocyanin but phylogenetically a hexamerin. It may be the descendent of a third hemocyanin subunit type that also gave rise to the insect hexamerins during evolution of the Dicondylia. This notion is reinforced by the apparent absence of hexamerins in Collembola, Diplura and Protura (data not shown).

Implications for hexapod phylogeny

Phylogenetic reconstruction among basal hexapods, e.g. the polyneopteran insects, is notoriously difficult, probably because a rapid divergence was followed by a relatively long period of subsequent evolutionary changes and hence loss of phylogenetic signal [25,26]. Hemocyanins and hexamerins have been successfully used to estimate evolutionary patterns among arthropods [18,27]. In fact, hemocyanin evolution is correlated with the evolution of arthropod taxa and we obtained strong support for well established taxa such as the Pancrustacea (Hexapoda + Crustacea), Hexapoda, Insecta, Dicondylia and Pterygota (Fig. 2).

Within the pterygote insects, the Polyneoptera (Plecoptera, Embioptera, Dermaptera, Grylloblattodea, Mantophasmatodea, Orthoptera, Phasmatodea and Dictyoptera) is a widely accepted monophylum based on an expansion in the anal region of the hind wing. Within this clade relationships are unclear and the placement of Plecoptera (stoneflies) and Dermaptera (earwigs) in particular is much disputed [28–32]. Molecular phylogenetic analyses of hemocyanins (Fig. 2) and other sequences [25,33–35] suggest a close relationship between Plecoptera and Dermaptera. However, at present there is no morphological evidence to support this topology [28–32].

Dictyoptera (Mantodea, Isoptera and Blattaria) is also a well-supported monophylum based on distinctive structures in the reproductive system, but the relationship among the three orders has remained unresolved [28–30,32]. In our analyses, the dictyopteran hemocyanin subunits also form a monophyletic clade. The hemocyanins subunits from *H. membrancea* (Mantodea) form the sistergroup of those from Isoptera + Blattaria. Hennig [28] further mentioned that Blattaria might be paraphyletic with respect to the Isoptera and recent studies suggest that termites actually evolved from wood-feeding cockroaches of the

genus *Cryptocercus* [25,36,37]. Indeed, the blattarian hemocyanin subunits are paraphyletic in our analyses: the subunits from the cockroach *B. dubia* (Blaberidae) are sistergroup of those from the termite *Cr. secundus* and the cockroach *P. americana* (Blattidae). In summary, our analyses have shown that hemocyanins are in fact excellent markers for reliable reconstruction of hexapod phylogeny.

Conclusions

Here we have demonstrated that hemocyanins are widely present in representatives of most ametabolous and hemimetabolous hexapod orders. All species used in our studies possess a typical tracheal system, with the exception of *S. curviseta* (Collembola), *F. candida* (Collembola) and *A. franzi* (Protura), in which cutaneous respiration might be sufficient due to their small body size [38,39]. Therefore, the presence or absence of hemocyanin in certain hexapod taxa cannot be readily related to a tracheal gas-exchange system. At present, the specific additional function of hemocyanin in Hexapoda must remain uncertain. There is little doubt that this respiratory protein is involved in O₂ transport, at least under certain environmental conditions or during some developmental stages. The Eumetabola, as well as certain ametabolous and hemimetabolous taxa (Protura, Diplura, Ephemeroptera and Odonata), have lost hemocyanin. One must assume that some currently unknown physiological or morphological modifications during the evolution of these taxa have rendered this type of respiratory protein unnecessary. The loss of hemocyanin might be one reason why hemoglobins are used as respiratory proteins in holometabolous species that are adapted to hypoxic environments [22,24].

Material and methods

Identification and molecular cloning of hemocyanin sequences

Total RNA was extracted from various hexapod species (Table 1) employing either the urea procedure [40] or the RNeasy Mini Kit (Qiagen, Hilden, Germany). An additional DNase digestion was performed using the RNase-Free DNase Set (Qiagen) according to the manufacturer's instructions. First-strand cDNA syntheses and subsequent PCR were carried out by using SuperScript II reverse transcriptase and AccuPrime Taq DNA Polymerase (Invitrogen, Karlsruhe, Germany) according to the manufacturer's instructions. For control of the efficiency of the cDNA synthesis, β -actin was amplified using the following

degenerated oligonucleotide primers: 5'-TGGCAYCAYACNTTYTAYAA-3' and 5'-GCDATNCCNGGRTACATNGT-3'. For the amplification of partial hemocyanin sequences, two pairs of degenerated oligonucleotide primers were designed according to conserved amino acid sequences of insect hemocyanins: 5'-ATGGAYTTYCCNTTYTGGTGGAA-3' and 5'-GTNGCGGTYTCRAARTGYTCCAT-3' to amplify a fragment of ~ 550 bp and 5'-GAGGGNSAGTTCGTNTACGC-3' and 5'-GAANGGYTTGTGGTTNAGRCG-3' to amplify a fragment of ~ 1050 bp. PCR fragments of the expected size were cloned into the pGem-T Easy/JM109 system (Promega, Mannheim, Germany) and 12–24 independent clones per species were sequenced by a commercial service (Genterprise, Mainz, Germany). 5'- and 3'-RACE experiments were carried out by RNA ligase-mediated rapid amplification method employing the GeneRacer Kit with SuperScript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions. Sets of genespecific primers were constructed according to the partial sequences (Table S1). The cDNA fragments were cloned into the pGem-T Easy/JM109 system (Promega) and three independent clones were sequenced as described above.

Sequence and molecular phylogenetic analyses

Partial sequences were assembled with GENEDOC 2.7 [41]. The tools provided with the ExPASy Molecular Biology Server of the Swiss Institute of Bioinformatics (<http://www.expasy.org>) were used for the analyses of DNA and amino acid sequences. Signal peptides were predicted using SIGNALP 1.1 [42]. The putative hemocyanin subunits identified in this study and the previously identified insect hemocyanins from *P. marginata* (PmaHc1 and PmaHc2), *P. grandis* (PgrHc1 and PgrHc2), *Sch. americana* ('embryonic hemolymph protein', SamEHP), *T. domestica* (TdoHc1 and TdoHc2) and *L. saccharina* (LsaHc1 and LsaHc2) were used to construct a multiple sequence alignment with MAFFT using the L-INS-i method and the BLOSUM 62 matrix [43]. We also included 17 selected insect hexamerins, 23 crustacean hemocyanins, 4 crustacean pseudohemocyanins, 20 chelicerate hemocyanins, 5 myriapod hemocyanins and 1 onychophoran hemocyanin in the final alignment, which was manually adjusted with the aid of GENEDOC. A list of sequences used in this study is provided in Table S2. Bayesian phylogenetic analysis was performed using MRBAYES 3.1 [44], using the WAG [45] model and assuming a gamma distribution of substitution rates. Prior probabilities for all trees were equal. Metropolis-coupled Markov chain Monte Carlo sampling was performed with one cold and three heated chains that were run for 1 000 000 generations. Starting trees were random, trees were sampled every 100th generation and posterior probabilities were estimated on the final 8000 trees (burnin = 2000). Bayesian phylogenetic analysis was rerun after partial sequences were excluded and additionally a maximum likelihood analysis was

performed using PHYML 2.4.3 [46,47] with the WAG [45] evolutionary model. The reliability of the branching pattern was assessed by bootstrap analysis with 100 replications.

Acknowledgements

This work has been supported by a grant of the Deutsche Forschungsgemeinschaft (Bu956/9). We thank J. Korb, K. Meusemann, M. Marx, B. Misof and B. Walz for providing hexapod species and M. Machola for her help with the experiments.

References

- 1 Markl J & Decker H (1992) Molecular structure of the arthropod hemocyanins. *Adv Comp Environ Physiol* **13**, 325–376.
- 2 van Holde KE & Miller KI (1995) Hemocyanins. *Adv Protein Chem* **47**, 1–81.
- 3 van Holde KE, Miller KI & Decker H (2001) Hemocyanins and invertebrate evolution. *J Biol Chem* **276**, 563–566.
- 4 Burmester T (2002) Origin and evolution of arthropod hemocyanins and related proteins. *J Comp Physiol B* **172**, 95–117.
- 5 Jaenicke E, Decker H, Gebauer W, Markl J & Burmester T (1999) Identification, structure, and properties of hemocyanins from Diplopod myriapoda. *J Biol Chem* **274**, 29071–29074.
- 6 Kusche K, Ruhberg H & Burmester T (2002) A hemocyanin from the Onychophora and the emergence of respiratory proteins. *Proc Natl Acad Sci USA* **99**, 10545–10548.
- 7 Kusche K, Hembach A, Hagner-Holler S, Gebauer W & Burmester T (2003) Complete subunit sequences, structure and evolution of the 6 × 6-mer hemocyanin from the common house centipede, *Scutigera coleoptrata*. *Eur J Biochem* **270**, 2860–2868.
- 8 Hagner-Holler S, Schoen A, Erker W, Marden JH, Rupprecht R, Decker H & Burmester T (2004) A respiratory hemocyanin from an insect. *Proc Natl Acad Sci USA* **101**, 871–874.
- 9 Whitten JM (1972) Comparative anatomy of the tracheal system. *Annu Rev Entomol* **17**, 373–402.
- 10 Mangum CP (1985) Oxygen transport in invertebrates. *Am J Physiol* **248**, 505–514.
- 11 Law JH & Wells MA (1989) Insects as biochemical models. *J Biol Chem* **264**, 16335–16338.
- 12 Willmer P, Stone G & Johnston I (2000) *Environmental Physiology of Animals*. Blackwell, Oxford.
- 13 Fochetti R, Belardinelli M, Guerra L, Buonocore F, Fausto AM & Caporale C (2006) Cloning and structural analysis of a haemocyanin from the stonefly *Perla grandis*. *Protein J* **25**, 443–454.

- 14 Pick C, Hagner-Holler S & Burmester T (2008) Molecular characterization of hemocyanin and hexamerin from the firebrat *Thermobia domestica* (Zygentoma). *Insect Biochem Mol Biol* **38**, 977–983.
- 15 Sánchez D, Ganfornina MD, Gutiérrez G & Bastani MJ (1998) Molecular characterization and phylogenetic relationship of a protein with potential oxygen-binding capabilities in the grasshopper embryo. A hemocyanin in insects? *Mol Biol Evol* **15**, 415–426.
- 16 Beintema JJ, Stam WT, Hazes B & Smidt MP (1994) Evolution of arthropod hemocyanins and insect storage proteins (hexamerins). *Mol Biol Evol* **11**, 493–503.
- 17 Burmester T & Scheller K (1999) Ligands and receptors: common theme in insect storage protein transport. *Naturwissenschaften* **86**, 468–474.
- 18 Burmester T (2001) Molecular evolution of the arthropod hemocyanin superfamily. *Mol Biol Evol* **18**, 184–195.
- 19 Hagner-Holler S, Pick C, Girgenrath S, Marden JH & Burmester T (2007) Diversity of stonefly hexamerins and implication for the evolution of insect storage proteins. *Insect Biochem Mol Biol* **37**, 1064–1074.
- 20 Telfer WH & Kunkel JG (1991) The function and evolution of insect storage hexamers. *Annu Rev Entomol* **36**, 205–228.
- 21 Burmester T (1999) Evolution and function of the insect hexamerins. *Eur J Entomol* **96**, 213–225.
- 22 Burmester T & Hankeln T (2007) The respiratory proteins of insects. *J Insect Physiol* **53**, 285–294.
- 23 von Heijne G (1986) A simple method for predicting signal peptide cleavage sequences. *Nucleic Acids Res* **14**, 4683–4690.
- 24 Weber RE & Vinogradov SN (2001) Nonvertebrate hemoglobins: functions and molecular adaptations. *Physiol Rev* **81**, 569–628.
- 25 Kjer MK, Carle FL, Litman J & Ware J (2006) A molecular phylogeny of Hexapoda. *Arthropod Syst Phylogeny* **64**, 35–44.
- 26 Whitfield JB & Kjer MK (2008) Ancient rapid radiations of insects: challenges for phylogenetic analysis. *Annu Rev Entomol* **53**, 449–472.
- 27 Burmester T, Massey HC Jr, Zakharkin SO & Benes H (1998) The evolution of hexamerins and the phylogeny of insects. *J Mol Evol* **47**, 93–108.
- 28 Hennig W (1969) *Stammesgeschichte der Insekten*. W. Kramer, Frankfurt.
- 29 Boudreaux HB (1979) *Arthropod Phylogeny with Special Reference Process to Insects*. Wiley, New York, NY.
- 30 Kristensen NP (1981) Phylogeny of insect orders. *Annu Rev Entomol* **26**, 135–157.
- 31 Kukalova-Peck J (1991) Fossil history and evolution of hexapod structures. In *The Insects of Australia* (Naumann ID, ed.), pp. 141–179. Melbourne University Press, Melbourne.
- 32 Wheeler WC, Whiting M, Wheeler QD & Carpenter JM (2001) The phylogeny of the extant hexapod orders. *Cladistics* **17**, 113–169.
- 33 Flook PK & Rowell CHF (1998) Inferences about orthopteroid phylogeny and molecular evolution from small subunit nuclear ribosomal DNA sequences. *Insect Biochem Mol Biol* **7**, 163–178.
- 34 Whiting MF, Bradler S & Maxwell T (2003) Loss and recovery of wings in stick insects. *Nature* **421**, 264–267.
- 35 Terry MD & Whiting MF (2005) Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics* **21**, 240–257.
- 36 Lo N, Tokuda G, Watanabe H, Rose H, Slaytor M, Maekawa K, Bandi C & Noda H (2000) Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. *Curr Biol* **10**, 801–804.
- 37 Klass KD (2001) Morphological evidence on *Blattaria* phylogeny: ‘phylogenetic histories and stories’ (Insecta, Dictyoptera). *Dtsch Entomol Z* **48**, 223–265.
- 38 Davies WM (1927) On the tracheal system of *Collembola*, with special reference to that of *Sminthurus viridis*, Lubbock. *Q J Microsc Sci* **71**, 15–30.
- 39 Williams CB (1913) A summary of the present knowledge of the Protura. *Entomologist* **46**, 225–232.
- 40 Holmes DS & Bonner J (1973) Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. *Biochem* **12**, 2330–2338.
- 41 Nicholas KB & Nicholas HB Jr (1997) *GeneDoc: Analysis and Visualization of Genetic Variation*. <http://www.psc.edu/biomed/genedoc/>.
- 42 Nielsen H, Engelbrecht J, Brunak S & von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1–6.
- 43 Katoh K, Kuma K-i, Toh H & Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**, 511–518.
- 44 Huelsenbeck JP & Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
- 45 Whelan S & Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* **18**, 691–699.
- 46 Guindon S & Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696–704.
- 47 Guindon S, Lethiec F, Duroux P & Gascuel O (2005) PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res* **33**, 557–559.
- 48 Grimaldi D & Engel MS (2005) *Evolution of the Insects*. Cambridge University Press, New York, NY.

Supporting information

The following supplementary material is available:

Fig. S1. cDNA and deduced amino acid sequence of *S. curviseta* hemocyanin subunit 1 (ScuHc1, FM242638).

Fig. S2. Partial cDNA and deduced amino acid sequence of *F. candida* hemocyanin subunit 1 (FcaHc1, FM242650).

Fig. S3. cDNA and deduced amino acid sequence of *M. germanica* hemocyanin subunit 1 (MgeHc1, FM242639).

Fig. S4. Partial cDNA and deduced amino acid sequence of *L. migratoria* hemocyanin subunit 1 (LmiHc1, FM242651).

Fig. S5. cDNA and deduced amino acid sequence of *C. morosus* hemocyanin subunit 1 (CmoHc1, FM242640).

Fig. S6. cDNA and deduced amino acid sequence of *Ch. acanthopygia* hemocyanin subunit 1 (CacHc1, FM242641).

Fig. S7. Partial cDNA and deduced amino acid sequence of *Ch. acanthopygia* hemocyanin subunit 1 (CacHc2, FM242654).

Fig. S8. cDNA and deduced amino acid sequence of *H. membranacea* hemocyanin subunit 1 (HmeHc1, FM242642).

Fig. S9. cDNA and deduced amino acid sequence of *H. membranacea* hemocyanin subunit 2 (HmeHc2, FM242643).

Fig. S10. cDNA and deduced amino acid sequence of *Cr. secundus* hemocyanin subunit 1 (CseHc1, FM242644).

Fig. S11. cDNA and deduced amino acid sequence of *Cr. secundus* hemocyanin subunit 2 (CseHc2, FM242645).

Fig. S12. cDNA and deduced amino acid sequence of *B. dubia* hemocyanin subunit 1 (BduHc1, FM242646).

Fig. S13. cDNA and deduced amino acid sequence of *B. dubia* hemocyanin subunit 2 (BduHc2, FM242647).

Fig. S14. cDNA and deduced amino acid sequence of *P. americana* hemocyanin subunit 1 (PamHc1, FM242648).

Fig. S15. cDNA and deduced amino acid sequence of *P. americana* hemocyanin subunit 2 (PamHc2, FM242649).

Fig. S16. Partial cDNA and deduced amino acid sequence of *Sh. lateralis* hemocyanin subunit 1 (SlaHc1, FM242652).

Fig. S17. Partial cDNA and deduced amino acid sequence of *Sh. lateralis* hemocyanin subunit 2 (StaHc2, FM242653).

Fig. S18. Multiple alignment used for phylogenetic analysis.

Table S1. Genspecific primer used for 5'- and 3'-RACE.

Table S2. List of sequences used for phylogenetic analysis.

Table S3. Comparison of hexapod hemocyanin domains.

This supplementary material can be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Table S1. Genspecific primer used for 5' and 3' RACE.

Genespecific primer for 5' and 3' RACE

ScuHc1 5' primer	5' - GGATGACTTGGCTGGTGACGAA - 3'
ScuHc1 5' nested primer	5' - CAGCACGGTGGAGGGTGGCAA - 3'
ScuHc1 3' primer	5' - CTGCCTTCTCCGCCTCCACA - 3'
ScuHc1 3' nested primer	5' - GCTCAACGCTCTGAATGACGCT - 3'
MgeHc1 5' primer	5' - GGCTTAAACCATTCAGGGGCTCAAC - 3'
MgeHc1 5' nested primer	5' - GCGTTGCTCGGGATTGCGGATGGT - 3'
MgeHc1 3' primer	5' - CCAGGTGTAATGGAGCACTTTGAGACT - 3'
MgeHc1 3' nested primer	5' - CATTACAAGGACACATTGGTGCCCTACA - 3'
CmoHc1 5' primer	5' - TCGTCGCTCGTAGTCCACCAT - 3'
CmoHc1 5' nested primer	5' - CTCCCTCCATCCCAGCGGCTCCA - 3'
CmoHc1 3' primer	5' - CCTGGGGGAGATGATGGAGCCGT - 3'
CmoHc1 3' nested primer	5' - GACACATCCTGCTCGGCAAGATCA - 3'
CacHc1 5' primer	5' - CCAACTCGGTAGGTGGTTTGAGGAT - 3'
CacHc1 5' nested primer	5' - CTGGTCCCATGAAAGAGGTTCAACA - 3'
CacHc1 3' primer	5' - CCACCTACCGAGTTGGTGGTGA - 3'
CacHc1 3' nested primer	5' - GACATGATCGACTACACTCGCAGA - 3'
CacHc2 5' primer	5' - GCAAGAGGTTCCACATGAGGGAGA - 3'
CacHc2 5' nested primer	5' - CGTTTGAGAGTCGTTGCGCATC - 3'
CacHc2 3' primer	5' - GGACAACTGCACAATGATGCTCAC - 3'
CacHc2 3' nested primer	5' - CAACTCAGATATGGCGTACCACCT - 3'
HmeHc1 5' primer	5' - GCGGTATGTTGTCTGTGGGTAGAATC - 3'
HmeHc1 5' nested primer	5' - GGCAAGTTCGTTGGAAAGCGTT - 3'
HmeHc1 3' primer	5' - TCCAAGCACCCAGAATACTACGGAG - 3'
HmeHc1 3' nested primer	5' - CCACATCATGTTGGGACAAATCGTC - 3'
HmeHc2 5' primer	5' - GAGGTTGCACCACAGGAAGGTTGT - 3'
HmeHc2 5' nested primer	5' - GGCGTCGAATCGAGCAACCATCTGA - 3'
HmeHc2 3' primer	5' - CGTCTCGCTGAACAACCGAGGGTA - 3'
HmeHc2 3' nested primer	5' - GTCAGCTCCACAACGACGGACA - 3'
CseHc1 5' primer	5' - GTTGTCTGTGGGTAGAAGCCATCCT - 3'
CseHc1 5' nested primer	5' - GGAGTTGACGAAGAGGTGTGGTA - 3'
CseHc1 3' primer	5' - CCTGGCGTCATGGAGCACTT - 3'
CseHc1 3' nested primer	5' - GAACGCCCTGGATGACACCCTAGA - 3'
CseHc2 5' primer	5' - CCTGATGCGACCCCTCGACGAT - 3'
CseHc2 5' nested primer	5' - GAGAGGGAAGGCAAGTCACAGAACT - 3'
CseHc2 3' primer	5' - CGTCAACCTGCTAGGGGAGATCAT - 3'
CseHc2 3' nested primer	5' - CGTGATGCTGTCCAAGGTGACGGAT - 3'
BduHc1 5' primer	5' - GCAGGGAATTCCGCCACCTACTCT - 3'
BduHc1 5' nested primer	5' - GGCTCCACAAATGGAAGGTCGTT - 3'
BduHc1 3' primer	5' - GCCATCTCACAGCAAGCCGTA - 3'
BduHc1 3' nested primer	5' - CGCTCTGATAGAACCATCAGTGGACT - 3'
BduHc2 5' primer	5' - GGCAAGTCCGAGAACTGCATGT - 3'
BduHc2 5' nested primer	5' - GGCTGCACTACGGGCAGATCA - 3'
BduHc2 3' primer	5' - CGAGGGACGCATTAGGGATGCTAT - 3'
BduHc2 3' nested primer	5' - GGAAGTGGTGGAATCGTCAGTTCACA - 3'
PamHc1 5' primer	5' - AACTCCCCACCCACACGGTATGTTGT - 3'
PamHc1 5' nested primer	5' - TCCACGATGGGAAGGTCGTTGGA - 3'
PamHc1 3' primer	5' - GGAACCTTCACACGATTCCAAGCA - 3'
PamHc1 3' nested primer	5' - GGGTCAGATTACCGATCCCAAAGGA - 3'
PamHc2 5' primer	5' - CCTTCCACGATCTTTTGGTCCCATTCC - 3'
PamHc2 5' nested primer	5' - GTAGAACAGTTCTCCCTTGCGGTCTT - 3'
PamHc2 3' primer	5' - CTGGCATAAACCTTCTGGGCGAGAT - 3'
PamHc2 3' nested primer	5' - CGGTCACGTGATGTTGTCCAAAGTGA - 3'

Table S2. List of sequences used for phylogenetic analysis.

Abbreviation	Accession No.	Protein	Species
EpiHc	CAD12808	hemocyanin	<i>Epiperipatus</i> sp.
ScoHcC	Q8T115	hemocyanin SU C	<i>Scutigera coleoptrata</i>
ScoHcD	Q8T116	hemocyanin SU D	<i>Scutigera coleoptrata</i>
ScoHcB	Q8IFJ8	hemocyanin SU B	<i>Scutigera coleoptrata</i>
ScoHcA	Q95P08	hemocyanin SU A	<i>Scutigera coleoptrata</i>
SpiHc1	CAC33894	hemocyanin SU 1	<i>Spirostreptus</i> sp.
LpoHcVI	CAJ91100	hemocyanin SU VI	<i>Limulus polyphemus</i>
EcaHcC	Q9NFL6	hemocyanin SU C	<i>Eurypelma californicum</i>
EcaHcB	Q9NFH9	hemocyanin SU B	<i>Eurypelma californicum</i>
NinHcB	CAD68053	hemocyanin SU B	<i>Nephilia inaurata</i>
LpoHcII	CAJ91097	hemocyanin SU II	<i>Limulus polyphemus</i>
EcaHcA	P14750	hemocyanin SU A	<i>Eurypelma californicum</i>
NinHcA	CAD68052	hemocyanin SU A	<i>Nephilia inaurata</i>
LpoHcIV	CAJ91099	hemocyanin SU IV	<i>Limulus polyphemus</i>
EcaHcD	P02241	hemocyanin SU D	<i>Eurypelma californicum</i>
NinHcD	CAD68054	hemocyanin SU D	<i>Nephilia inaurata</i>
EcaHcF	Q9NFL5	hemocyanin SU F	<i>Eurypelma californicum</i>
NinHcF	CAD68056	hemocyanin SU F	<i>Nephilia inaurata</i>
EcaHcE	P02242	hemocyanin SU E	<i>Eurypelma californicum</i>
NinHcE	CAD68055	hemocyanin SU E	<i>Nephilia inaurata</i>
EcaHcG	Q9NFL4	hemocyanin SU G	<i>Eurypelma californicum</i>
CsaHc1	CAC44749	hemocyanin SU 1	<i>Cupiennius salei</i>
NinHcG	CAD68057	hemocyanin SU G	<i>Nephilia inaurata</i>
AuaHc6	P80476	hemocyanin SU 6	<i>Androctonus australis</i>
LpoHcI	1LL1_A	hemocyanin SU I	<i>Limulus polyphemus</i>
TtrHcA	AAZ40335	hemocyanin SU A	<i>Tachyplesus tridentatus</i>
HamHcA	AJ272095	hemocyanin SU A	<i>Homarus americanus</i>
PinHcB	P10787	hemocyanin SU b	<i>Panulirus interruptus</i>
PinHcA	P04254	hemocyanin SU a	<i>Panulirus interruptus</i>
PinHcC	S21221	hemocyanin C	<i>Panulirus interruptus</i>
PvaHc1	AJ250830	hemocyanin SU 1	<i>Penaeus vannamei</i>
PvaHc	X82502	hemocyanin	<i>Penaeus vannamei</i>
PelHc1	AJ344361	hemocyanin SU 1	<i>Palinurus elephas</i>
PelHc2	CAC69244	hemocyanin SU 2	<i>Palinurus elephas</i>
PelHc3	CAC69245	hemocyanin SU 3	<i>Palinurus elephas</i>
PelHc4	CAD56697	hemocyanin SU 4	<i>Palinurus elephas</i>
EpuHc1	EST	hemocyanin SU 1	<i>Eurydice pulchra</i>
CcsHc	ABB59715	hemocyanin	<i>Cyamus scammoni</i>
GpuHc1	CAI78901	hemocyanin SU 1	<i>Gammarus roeseli</i>
GroHc1	CAI78901	hemocyanin SU 1	<i>Gammarus roeseli</i>
PleHc	AF522504	hemocyanin	<i>Pacifastacus leniusculus</i>
PleHc2	AAO47336	hemocyanin SU 2	<i>Pacifastacus leniusculus</i>
CmaHc1	AAW57889	hemocyanin SU 1	<i>Cancer magister</i>
CmaHc2	AAW57890	hemocyanin SU 2	<i>Cancer magister</i>
CmaHc3	AAW57891	hemocyanin SU 3	<i>Cancer magister</i>
CmaHc4	AAW57892	hemocyanin SU 4	<i>Cancer magister</i>
CmaHc5	AAW57893	hemocyanin SU 5	<i>Cancer magister</i>
CmaHc6	U48881	hemocyanin SU 6	<i>Cancer magister</i>
CsaHc	AF249297	hemocyanin	<i>Callinectes sapidus</i>
CmaCC1	AF091261	cryptocyanin 1	<i>Cancer magister</i>
CmaCC2	ABB59714	cryptocyanin 2	<i>Cancer magister</i>
HamPHc1	AJ132141	pseudo-hemocyanin 1	<i>Homarus americanus</i>
HamPHc2	Q6KF81	pseudo-hemocyanin 2	<i>Homarus americanus</i>
SamEHP	AF038569	embryonic hemolymph protein	<i>Schistocerca americana</i>
PgrHc1	DQ118369	hemocyanin SU 1	<i>Perla grandis</i>
PgrHc2	DQ118370	hemocyanin SU 2	<i>Perla grandis</i>
PmaHc1	AJ555403	hemocyanin SU 1	<i>Perla marginata</i>
PmaHc2	AJ555404	hemocyanin SU 2	<i>Perla marginata</i>

TdoHc1	FM165288	hemocyanin SU 1	<i>Thermobia domestica</i>
TdoHc2	FM165289	hemocyanin SU 2	<i>Thermobia domestica</i>
LsaHc1	FM165291	hemocyanin SU 1	<i>Lepisma saccharina</i>
LsaHc2	FM165292	hemocyanin SU 2	<i>Lepisma saccharina</i>
ScuHc1	FM242638	hemocyanin SU 1	<i>Sinella curviseta</i>
FcaHc1	FM242650	hemocyanin SU 1	<i>Folsomia candida</i>
CacHc1	FM242641	hemocyanin SU 1	<i>Chelidurella</i>
			<i>acanthopygia</i>
CacHc2	FM242654	hemocyanin SU 2	<i>Chelidurella</i>
			<i>acanthopygia</i>
CmoHc1	FM242640	hemocyanin SU 1	<i>Carausius morosus</i>
LmiHc1	FM242651	hemocyanin SU 1	<i>Locusta migratoria</i>
HmeHc1	FM242642	hemocyanin SU 1	<i>Hierodula membranacea</i>
HmeHc2	FM242643	hemocyanin SU 2	<i>Hierodula membranacea</i>
BduHc1	FM242646	hemocyanin SU 1	<i>Blaptica dubia</i>
BduHc2	FM242647	hemocyanin SU 2	<i>Blaptica dubia</i>
CseHc1	FM242644	hemocyanin SU 1	<i>Cryptotermes secundus</i>
CseHc2	FM242645	hemocyanin SU 2	<i>Cryptotermes secundus</i>
PamHc1	FM242648	hemocyanin SU 1	<i>Periplaneta americana</i>
PamHc2	FM242649	hemocyanin SU 2	<i>Periplaneta americana</i>
SlaHc1	FM242652	hemocyanin SU 1	<i>Shelfordella lateralis</i>
SlaHc2	FM242653	hemocyanin SU 2	<i>Shelfordella lateralis</i>
MgeHc1	FM242639	hemocyanin SU 1	<i>Machilis germanica</i>
TdoHex1	FM165290	hexamerin 1	<i>Thermobia domestica</i>
PmaHx1	AM690365	hexamerin 1	<i>Perla marginata</i>
LmiHex1	EST	hexamerin 1	<i>Locusta migratoria</i>
LmiHex2	EST	hexamerin 2	<i>Locusta migratoria</i>
RmiHx1	AY923852	hexamerin 1	<i>Romalea microptera</i>
BdiHex	U31328	hexamerin	<i>Blaberus discoidalis</i>
PamHex12	L40818	hexamerin (allergen clone C12)	<i>Periplaneta americana</i>
RflHex1	AY572858	hexamerin 1	<i>Reticulitermes flavipes</i>
RflHex2	AY572859	hexamerin 2	<i>Reticulitermes flavipes</i>
AgeHex2	AY293287	arylphorin-like hexamerin 2	<i>Apriona germari</i>
TmoHex2	AF395329	hexamerin 2	<i>Tenebrio molitor</i>
LdeDP1	X76080	diapause protein 1	<i>Leptinotarsa</i>
			<i>decemlineata</i>
CfeHex2	AJ251271	hexamerin 2	<i>Camponotus festinatus</i>
BheHex	I25974	Insecticidal toxin (hexamerin)	<i>Bracon hebetor</i>
MseAryla	M28396	arylphorin alpha	<i>Manduca sexta</i>
MseArylb	M28397	arylphorin beta	<i>Manduca sexta</i>
AgaHex2	XM_321800	hexamerin 2	<i>Anopheles gambiae</i>

Table S3. Comparison of hexapod hemocyanin domains. Percent identities between hemocyanin domains 1 (top), 2 (middle) and 3 (bottom) were calculated from amino acid sequences.

	Scu Hc1	Mge Hc1	Tdo Hc1	Pma Hc1	Cmo Hc1	Sam EHP	Cac Hc1	Hme Hc1	Cse Hc1	Bdu Hc1	Pam Hc1	Tdo Hc2	Pma Hc2	Hme Hc2	Cse Hc2	Bdu Hc2
Domain 1																
MgeHc1	59.3															
TdoHc1	66.3	57.6														
PmaHc1	62.0	54.1	61.6													
CmoHc1	63.4	58.7	70.9	63.7												
SamEHP	61.6	58.1	75.0	64.0	74.4											
CacHc1	63.4	57.0	69.2	61.0	69.2	68.0										
HmeHc1	65.1	58.7	72.7	66.1	73.8	77.9	69.2									
CseHc1	65.7	57.0	72.7	59.6	73.8	73.3	65.1	77.9								
BduHc1	68.6	61.6	75.0	63.2	72.7	74.4	69.8	80.8	84.9							
PamHc1	65.1	56.4	73.3	66.1	73.8	74.4	70.3	80.8	80.8	84.9						
TdoHc2	49.4	41.3	47.7	47.4	44.8	48.8	43.6	48.3	47.7	49.4	48.3					
PmaHc2	45.3	40.1	41.3	44.4	40.1	39.5	38.4	41.9	41.3	43.0	40.1	58.7				
HmeHc2	44.8	42.4	44.2	47.1	45.3	46.5	44.8	45.3	46.5	48.3	46.5	68.0	59.9			
CseHc2	44.8	41.3	47.7	48.0	45.3	48.8	44.2	46.5	48.8	48.3	45.9	66.3	58.7	70.3		
BduHc2	50.0	43.6	46.5	47.4	48.3	49.4	44.8	46.5	48.8	50.6	47.1	66.3	58.7	65.7	72.1	
PamHc2	48.3	43.0	45.3	48.8	45.9	46.5	45.9	47.1	48.8	49.4	47.7	67.4	61.6	73.3	77.3	73.8
Domain 2																
MgeHc1	62.5															
TdoHc1	68.9	63.1														
PmaHc1	61.3	59.6	69.1													
CmoHc1	67.1	61.0	79.4	71.1												
SamEHP	63.1	57.0	77.6	68.0	81.8											
CacHc1	64.4	61.0	80.3	74.7	81.8	80.4										
HmeHc1	67.1	60.1	82.1	73.3	83.6	79.6	85.3									
CseHc1	65.3	60.1	77.6	74.2	80.4	77.8	79.6	83.1								
BduHc1	68.9	59.2	81.2	72.0	85.3	83.1	85.8	89.3	84.0							
PamHc1	69.3	59.6	79.4	73.3	83.1	80.9	84.0	88.4	84.0	88.9						
TdoHc2	64.3	58.9	70.1	62.5	65.2	62.5	66.5	64.3	59.0	65.6	66.5					
PmaHc2	60.3	55.8	61.2	59.8	61.2	57.1	59.4	58.9	62.5	58.9	61.6	68.8				
HmeHc2	64.9	59.0	66.2	59.6	66.4	60.5	63.7	65.5	63.7	64.1	65.9	74.6	71.0			
CseHc2	62.9	56.3	65.6	57.1	65.2	61.2	62.5	64.7	63.1	64.3	65.9	71.9	68.3	77.7		
BduHc2	63.8	58.9	66.5	59.8	67.0	59.8	63.8	63.8	64.3	64.3	66.5	76.8	70.5	84.4	80.4	
PamHc2	64.7	57.1	65.6	58.0	66.5	62.1	63.8	66.5	63.8	65.6	67.4	75.9	70.1	82.9	86.6	85.7

[illegible]

CCAGAGCTCTTTGTGAAGTAAACCAGCTCGCGTTATC
ATGGCACGGTTCGCACTCACTTTGATTGCCTGCACCCTTGTAGTCGCAGTTGCAGCTACC
M A R F A L T L I A C T L V V A V A A T
CCCGCTGACCAAGAATTTCTACCAAGCAGAAGGAAATTGTCAAAGTCTTAACAAAGTC
P A D Q E F L T K Q K E I V K L L N K V
CACGAACTGAACTTTTATCAAGATCAAGCTACCATTGGAAAGGACTGGGACCCATTAGCT
H E L N F Y Q D Q A T I G K D W D P L A
CACTTGGATAGCTACAAGAATGTCCGAGTTGTAAAGGAATTGGTCAAGGAACTTAAGAAT
H L D S Y K N V R V V K E L V K E L K N
GGAAAGCTTATCAAGCGTGGAGAAATCTTCAACCTGTTCAATGAGGAGCACCGAAGGGAG
G K L I K R G E I F N L F N E E H R R E
ATGATCTTGCTCTTTGAGACTCTCTTTGCCAAGGATTGGGATACCTTCCACAAGACT
M I L L F E T L F F A K D W D T F H K T
GCCTGTTGGGCTCGTGACAAGATCAATGAGGGTCAATTCGTCTACGCCCTCAGCGTTGCC
A C W A R D K I N E G Q F V Y A L S V A
ACCCTCCACCGTGTGATACTCGCGGAATCAGGTTGCCACCTGCATACGAACTTATCCC
T L H R A D T R G I R L P P A Y E T Y P
CATCTTTTCGTCACCAGCCAAGTCATCCATGAGGCTTATGCTGCCAAGATGAGGCAAGAA
H L F V T S Q V I H E A Y A A K M R Q E
CCCGCTGTTATCCACATGAACTTCACTGGTACCGTGCAGCAACCCGAACAACGCGTCGCC
P A V I H M **N F T** G T V R N P E Q R V A
TACTTTGGAGAGGATGTTGGTATGAACAGCCATCACGCAGTGTTCCATATGGACTGGCCC
Y F G E D V G M N S H A V F M D W P
TTCTGGTGGAAACGAAGAGAAGTACGGTTTGACCAAGGACCGTAAGGGAGAAGTCTTCTGG
F W W N E E K Y G L T K D R K G E L F W
TACATGCATCACCAGCTTATTACACGTTTTGATGCTGAGCGTCTGTCCAACGACTTGAAC
Y M H Q L I T R F D A E R L S N D L N
GAAGTTGAGGCACTTCACTGGGACAAGCCATCGTCGAGGGATTCTACCCACAAACCACC
E V E A L H W D K P I V E G F Y P Q T T
TACAGGAAGGGAGGAGAATTCCTGCCAGGCCTGATAACTTCATGTTCCATGACCTTAAG
Y R K G G E F P A R P D N F M F H D L K
GACCACCGTGTGCTGATTTGGAAGCCTACGAAGAACGCATCCGCGAAGCCATCTCTGCC
D H R V A D L E A Y E E R I R E A I S A
GGTGTCACTTTTGTGCTCATGATCCCAACACTTTCATCTCCTTGAATAACACTGAAGGT
G V I F D A H D P N T F I S L **N N T** E G
ATCGACAAGTTGGGTGCCATCATTGAAGCCAGTTCCTGCTCCATCAACCCAGCTTCTAT
I D K L G A I I E A S S C S I **N P S** F Y
GGATCACTCCACAATCTTGGCCACATCATACTGGGACGCGTTGTCGATCCCCTCGGCAAA
G S L N L G I I L G R V V D P L G K
TTCGGTATGCCCCCAGGTGTTATGGAGCATTTCGAAACTGCCACCCGTGACCCTGCCTTC
F G M P P G V M E H F E T A T R D P A F
TTCCGCCCTCCACAAGCATATTGACAACATCTTCAAGGAACACAAGGACAGTCTCGTCCCA
F R L K H I D N I F K E H K D S L V P
TACACCCATGAAGAATTGGATGTTGAAGGAGTTGACATCAAAAATGTTGAGGTTGACGAC
Y T H E E L A D V E G V D I K N V E V D D
CTTGTCACCTACTTTGAAGACTTTGACATTGATATGCTCAACGCTCTGGATGACGCTGCT
L V T Y F E D F D I D M L N A L D D A A
GGTCTTACCGATGTTGACATCAAGGCCCGTGTCCAACGTCTTAACCACAAGCCTTTTGGA
G L T D V D I K A R V Q R L N H K P F G
ATTAAGATCATTGCCAACAGCGCTGCTGAGAAGACTGTGACTGTCAGACTTTTCTTGCT
I K I I A N S A A E K T V T V R L F L A
CCCAAATACGACTGGTATGGTCGTGAGGTTCTCTTGACATCCAACGCTGGAAGTTTATT
P K Y D W Y G R E V P L D I Q R W K F I
GAACTCGACAAATTCGCCGTCAAGTTGACTGCCGGAGAAAACGCCATTGTCCGCAAGTCC
E L D K F A V K L T A G E N A I V R K S
TCCGAGTCTTCAGTTACCATTCCAGACCCAGTCTCCACCCATGATCTTCGCAAGCTCGTT
S E S S V T I P D P V S T H D L R K L V
GATGATGCTATCGCCGGCACCGCTACCTTGGAAGTTGACAAGGATGTCAGACACTGTGGA
D D A I A G T A T L E V D K D V R H C G
GTTCCAGACAGGCTTTTGCTTCCCAAGGGCAAGACCAACGGAATGAAGTACACTATGTTCC
V P D R L L L P K G K T N G M K Y T M F

GTCATGCTCAGTGATTTTGAAGAAGACAAGGTCAATGATCTTCCTCACGACTATGAATAC
 V M L S D F E E D K V N D L P H D Y E Y
 GGTGGTTTCGTCTCATACTGCGGAACATTAACCACAAGTACCCAGACAAGAAGCCCATG
 G G F V S Y C G T I N H K Y P D K K P M
 GGTGGGCCCTTGGACAGGAAGATCTGCAAGGACAGCTTCCATGACGTCACCAACATGTAC
 G W P L D R K I C K D S F H D V T N M Y
 TTCCGTGATGTTGAAATCAAATTCGAGCACACTCATTAAAGTCTTTCAACACATTAGGCA
 F R D V E I K F E H T H ***
 TTTTCAAATTTTGTCTTTAAAAATCAGAGATTCAAAAACATTTCTTACAATAAACTCTTGA
 AAGTTCATCGAAGCCAATTAAAAAAAAAAAAAAAAAAAAA [2178 bp]

Fig. S1. cDNA and deduced amino acid sequence of *S. curviseta* hemocyanin subunit 1 (ScuHc1, FM242638). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

GAGGGCCAGTTCGTCTACGCCCTGAGCGTCGCAACCCTTCACCGTGCCGACACTCGCGGA
 E G Q F V Y A L S V A T L H R A D T R G
 GTCAGGTTGCCCCCTGCGTATGAGACATACCCTCATCTTTTCGTCACCAGCCAAGTCATC
 V R L P P A Y E T Y P H L F V T S Q V I
 CATGAGGCTTACGCTGCGAAGATGAGACAGGAGCCTGCAGTCATCCACATGAACTTCACT
 H E A Y A A K M R Q E P A V I H M **N F T**
 GGAACCTGCCGCAACCCTGAACAGCGTGTGGCTTATTTTGGAGAGGACATTGGCATGAAC
 G T C R N P E Q R V A Y F G E D I G M N
 AACCACCACGAGTCTTCCACATGGACTGGCCTTTCTGGTGGAAACGAAGAGAAGTATGGA
 N H A V F M D W P F W W N E E K Y G
 CTTCAACAAGGACCGCAAGGGAGAACTTTTCTGGTACATGCATCACCAGCTTATCACTCGG
 L H K D R K G E L F W Y M H Q L I T R
 TTTGATGCTGAACGTTTGTCCAATGACCTGAACGAAGTGGAGCCACTTAAATGGGACAAG
 F D A E R L S N D L N E V E P L K W D K
 CCCATCGTCGACGGATTTTACCCCCAAACCACGTACAGGAAGGGAGGAGAATTCCCTGCC
 P I V D G F Y P Q T T Y R K G G E F P A
 AGGCCTGACAACTTCAAGTTCCAAGACCTTAAGGATCACCGTGTGCTGATTTGGAAGCC
 R P D N F K F Q D L K D H R V A D L E A
 TACGAAGAACGTATCCTTGAAGCTATTGCTGCCGATTACGTCATCGCTGCTGATGATCAC
 Y E E R I L E A I A A D Y V I A A D D H
 CACACCGTCACCTCCTTGAACAACACTGAGGGTATCGACAAGTTGGGTGCCATCATTGAA
 H T V T S L **N N T** E G I D K L G A I I E
 GCCAGTTCCTGCTCCGTCAATCCCCACTATTACGGATCCCTCCACAACCTCGGCCACATT
 A S S C S V N P H Y Y G S L N L G I
 ATGCTTGGACGTGTCGTTGATCCCCTCGGCAAGTTCGGTATGCCACCAGGTGTCATGGAA
 M L G R V V D P L G K F G M P P G V M E
 CATTTGAGACCGCTACCAGGGACCCTGCCTTCTTCCGCCTCCACAAGCACATTGACGAA
 H F E T A T R D P A F F R L K H I D E
 ATCTTGTTCAAGAAACACAAGGACAGTCTTACCCCATACACCCACGAGGAACTCGATGTT
 I L F K K H K D S L T P Y T H E E L D V
 GAGGGAGTAGACATTAAGGATGTCGAGGTTGATGATCTTGAAACCTACTTTGAAGAATAC
 E G V D I K D V E V D D L E T Y F E E Y
 GACATCGACATGCTCAATGCTCTTGATGACGCCGAGGGACTTCCCGATGTTGAGATCAAG
 D I D M L N A L D D A E G L P D V E I K
 GCCCGTGTCCAACGTCTAAACCACAAACCCTTC **[1053 bp]**
 A R V Q R L N H K P F

Fig. S2. Partial cDNA and deduced amino acid sequence of *F. candida* hemocyanin subunit 1 (FcaHc1, FM242650). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined.

GACTTTGTGTTTCGACGTTGGGTCCCTCTCAAC

ATGAAGACTGTTTTCTGCCTTCTCCTGGCTCTTGCAGCATCCTTGGCGGTGCCAACCAAG

M K T V F C L L L A L A A S L A V P T K

AAGGATGTACCCGCTGATCATGAATTCCTGGTCAAGCAACGTGAGATTTTGAAGCTTTTG

K D V P A D H E F L V K Q R E I L K L L

TTCCGTGCCCATGAACCCAACATTTCTGAAGACCAAGTGATATAGCAGCGAAGTTCGAC

F R A H E P **N I S** E D Q V D I A A K F D

ATTGAAGCTAATTTGGACAAATTTAAGGAAGTAGACAGCGTAAAAAGAGTTGTACGAGCTG

I E A N L D K F K E V D S V K E L Y E L

TACGTCAAGAAGCAGCTTCTCCATAGAGGAGAAATATTTAATACCTTCGAAGACAAACAT

Y V K K Q L L H R G E I F N T F E D K H

TTGGAACAGGCTATTCTTCTTATCGAAACACTGTTGTCTGCAAATAGTTGGGATACTCTG

L E Q A I L L I E T L L S A N S W D T L

TATAGAACAGCATGCTGGGCTCGTGACAAGGTCAATGAGGGACAGTTTGTCTACGCACCTC

Y R T A C W A R D K V N E G Q F V Y A L

AGTGTAACGGTCATGCACAGCGATTTCTCCGTGGAATCATATTGCCTCCACCTTATGAA

S V T V M H S D F L R G I I L P P P Y E

ATCTACCCACACCTGTTTGTCAACAGCGAGGTTATCCATGAAGCCTACAGCGCCAAATG

I Y P H L F V N S E V I H E A Y S A K M

CGTCATCATCCTGTTGTGATTCCAATGAACTTCACCGGAACCATCCGCAATCCCAGCAA

R H H P V V I P M **N F T** G T I R N P E Q

CGCGTTGCCTATTTTCGGAGAAGACATTGGTCTGAACACCCATCATCACTACTGGCACATG

R V A Y F G E D I G L N T H H Y W M

AATTTCCCTTCTGGTGGAGTCTCATTATGACACAAAATTCGACCGTGCAGGAGAAATG

N F P F W W S P H Y D T K F D R A G E M

TTCTGGTATATGCACCACCAACTGGTAGCTCGTTACGACTTGAGCGTCTGTCTAACTAC

F W Y M H Q L V A R Y D L E R L S N Y

CTTCCAGAAGTTGAGCCCCTGGAATGGTTTAAGCCAATCAAGTCTGGATTGCCCCCTCTG

L P E V E P L E W F K P I K S G F A P L

ACGATGTACCGCAAAGGTGGTGAATTCCTCAACCGCCAGACAACATGCTTCTGCACGAT

T M Y R K G G E F P N R P D N M L L H D

CTTCATGACCTAAAAGTGCATGATGTTGTTGTTTTGAGGAGAGAATCAAAGAAGCAATT

L H D L K V H D V V V F E E R I K E A I

GATCGTGGATTTGTGTACAACAAGTGGATCGGAGAAAAAATTCGTTTGAACAACAGCATT

D R G F V Y N K W I G E K I R L **N N S** I

GAAGGCATTGAGACCCTTGCCGTATGATTGAGGCCTCCCGATTATCACCAGACATCGAC

E G I E T L G R M I E A S R L S P D I D

TATTACGGCTCATAACAACCTTGCCACATCCTTTGGGTGAAATTATGGACCCTGAC

Y Y G S I N L G I L L G E I M D P D

CACAAGTTCAATCTTCTCCAGGTGTAATGGAGCACTTTGAGACTGCAATGAGAGACCCA

H K F N L P P G V M E H F E T A M R D P

GTATTCTTCAGCCTTCACAAGCATATTGATTATATCTTCAAGCATTACAAGGACACATTG

V F F S L K H I D Y I F K H Y K D T L

GTGCCCTACAAGAGAGAAGAGCTGGATTTCCAGGAGTGAAAGTAGAAAATGTTGAGGTT

V P Y K R E E L D F P G V K V E N V E V

GACAGACTCGTTACCTTCTTTGAACCTTCGACATTGATCTTTATAATGCACTTGATGAC

D R L V T F F E P F D I D L Y N A L D D

GACAAGACCGTTTCCGACATTGAAATCAAGGCCCGGTACAACGCATCAACCACAAACCT

D K T V S D I E I K A R V Q R I N H K P

TTCACCTACCACATCAATGTAGAGAGCGATGTGCGAGAGAAAAGTTGTCGTGCGCACCTTC

F T Y H I N V E S D V E R K V V V R T F

ATTGGACCCAAGTACGACTGGTACCATCAAGAGGTGCCTGTTAACGAGAAGCGATGGGAG

I G P K Y D W Y H Q E V P V N E K R W E

ATGTTTGAAGTGGACAAATTCCTTACTCATATTCCTGCAGGCAAAATCCGTTCATTGAAAGA

M V E L D K F L T H I P A G K S V I E R

CTGTGCACTGAATCCACTGTGACCACACCTGACTACGAATCCTTCCGTTCACTTGTAAC

L S T T E S T V T T P D Y E S F R S L V N

CGCGTTGATGAAGCTCTCAAGGGCAACAAGGAATTCATTATTGATGAGGAATCCGTCAT

R V D E A A L K G N K E F I I D E E F R H

TGTGGCCTGCCTGATCGTCTTCTCATCCCCAAGGGCAACGAGGAAGTTATCCCGTCAAG

C G L P D R L L I P K G N E E G Y P V K

TTCTTTGTTATAGTCACCGACTGGGAAGAGGACAAGGTCAATCAGGAGGTGAGAACCA

```

F F V I V T D W E E D K V N Q E V E N H
CGTTACGGAAACGTCTACAGCTACTGTGGCACATACGGAGACCGTCTCTACCCTGACAAG
R Y G N V Y S Y C G T Y G D R L Y P D K
AAGGCCTTCCTTTATCCTTTTCGACCGTGTCAAGGACGTCAATGTTTTCAAACACCC
K A F L Y P F D R V I K D V N V F K T P
AACATGTTTCGGCAAATACGTCAGCATCTACCACAAGGATATCAGTGAACCTAACAGAGTC
N M F G K Y V S I Y H K D I S E L N R V
GTTCTGTTGTTGAGGAATAAAACATGAATGTATAAAAGATTTGACTTCTTGTGGCAATG
V P V V E E ***
TAATTTTCTGAGTATTTCTTGGGAAATTTATTTGTTATTAAATAACTTCCTTGTAATAA
AAAAAAAAAAAAAAAAAAAA [2208 bp]

```

Fig. S3. cDNA and deduced amino acid sequence of *M. germanica* hemocyanin subunit 1 (MgeHc1, FM242639). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATGGACTTCCCATTTTGGTGGAAGGAGGAGGAGTACGGCATCCCCAAGGAGCGCAAGGGC
 M D F P F W W K E E E Y G I P K E R K G
 GAGCTCTTCTACTACATGCACCACCAGCTGATCGCGCGCTTCGACCTCGAGCGCCTGTCC
 E L F Y Y M H Q L I A R F D L E R L S
 AACGACCTGCCGTTCTGTCGAGCCGCTGTACTGGAACGAGCGCATCAAGGACGGCTTCTAC
 N D L P F V E P L Y W N E R I K D G F Y
 CCGCAGACCACCTACCGCGTCGGAGGGGAGTTCCCCGCCAGGCCCGACAACCTTCGCCTTC
 P Q T T Y R V G G E F P A R P D N F A F
 CACGACCTGCAGAACATCAAGGTGCAGCAGATGATCGACTACACCAGGCGCGTCCGAGAG
 H D L Q N I K V Q Q M I D Y T R R V R E
 GCCATCTCCCAGCAGGCAGTCATCACCGAAACCGGCGACCTGTACAACATTAGCAGCCCG
 A I S Q Q A V I T E T G D L Y **N I S** S P
 GAGGGCATCAACGTGCTGGGAGAGCTGGTTCGAGCCGTCCACCCGCTCCAAGCACCGCGAG
 E G I N V L G E L V E P S T R S K H R E
 TACTACGGCGCCCTGCACAACCTACGGACACATCATGCTGGGCCAGATCACCGACCCCAAG
 Y Y G A L N Y G I M L G Q I T D P K
 CGCAAGTTCAACATGCCTCCCGGTGTGATGGAACATTTTCGAGACCGCCAC **[530 bp]**
 R K F N M P P G V M E H F E T A

Fig. S4. Partial cDNA and deduced amino acid sequence of *L. migratoria* hemocyanin subunit 1 (LmiHc1, FM242651). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined.

TATGCCAGGGGCAAGGCCCGCCAGGC
GCAACGCAGCGATGGTCTTCCAGCTTCTGCAGCAGGGGAGCCACACTCGAAACAGCGAAC
ATGCACTGGGTGATCGCGCGGTGCTGTTGGCCGGCCTGGGCCAGCTGGTCGCGGCCAAG
M Q W V I A A V L L A G L G Q L V A A K
TACACCAGGGGCGAGGTGCCAGCCGACAAGGACTTCCTGTTGAAGCAGCGCGAGGTGCTG
Y T R G E V P A D K D F L L K Q R E V L
AAGCTACTCAACAAACCCACGAGCCCAACAGGTACAAAGACCAGGCCGACTCCGGCAGC
K L L N K P H E P N R Y K D Q A D S G S
AAGTACGAGCCGTCCAACAACCTCAACAGATACAAGAACCCGGTGCCTGTGAAGACGCTG
K Y E P S N N L N R Y K N P V P V K T L
GTGAAGCGCTACAACCTCGCACTCGCTGCTGCCTCGCGGCCAGATCTTCACACTGTTGAC
V K R Y N S H S L L P R G Q I F T L F D
GACAAGCACCGCGAGGAGATGGTGTGCTGTTTCGAGTCGATGTTCTACGCCCTCGACTGG
D K H R E E M V L L F E S M F Y A L D W
GACACCTTCTACAAGACGGCGTGCTGGGCCCCGCGACAGGATCAACGAGGGACAGTTCATC
D T F Y K T A C W A R D R I N E G Q F I
TACGCGCTGAGCGTGGCGGTGATCCACGGGACGACGCCAAGGGCATCGTGCTGCCGCCG
Y A L S V A V I H R D D A K G I V L P P
TCGTATGAGATTTACCCGCACCTGTACGTCAACAGCGAGGTCATACACGCCCGCTACAAG
S Y E I Y P H L Y V N S E V I H A A Y K
GCCAAGATGAGGCAGGAGCCCGCGTCGTGCGCATGAACTTCACAGGAACCATCCGCAAC
A K M R Q E P A V V R M **N F T** G T I R N
CCAGAGCAGCGTGTAGCCTACCTAGGCGAAGACCTGGGCATGAACTCTCACCATGCGCAC
P E Q R V A Y L G E D L G M N S H A H
TGGCACATGGACTTCCCTTCTGGTGGAAGCCTGAAGAGTACGGCATCGACAAGGACAGG
W M D F P F W W K P E E Y G I D K D R
AAGGGAGAGCTCTTCTACTACATGCACCACCAGCTCATCGCCAGATTGACCTGGAGCGG
K G E L F Y Y M H Q L I A R F D L E R
CTGTGCAACGACCTTCCCCCGTGGAGCCGCTGGGATGGAGGGAGCGCATCGTGACGGG
L S N D L P P V E P L G W R E R I V D G
TTCTACCCACAGACGACGTACCGCGTGGGAGGGGAGTTCCCCGCCAGGCCCGACGACTTC
F Y P Q T T Y R V G G E F P A R P D D F
GAGTTCCAGGACCTGAGCTCCATCAAGGTGCAGGACATGGTGGACTACGAGCGACGAGTG
E F Q D L S S I K V Q D M V D Y E R R V
CGCGAGGCGATCTCCCAGCAGGCCGTATCACCATGTGCGGGCGACTACTTCTCCCTGAAC
R E A I S Q Q A V I T M S G D Y F S L **N**
GACACGACGGGCATCAACACCCTGGGGGAGATGATGGAGCCGTCCACCACGTCCAAGCAC
D T T G I N T L G E M M E P S T T S K H
CGCGACTACTACGGCGCGTGCACAACTACGGACACATCCTGCTCGGCAAGATCACCGAC
R D Y Y G A L N Y G I L L G K I T D
CCCAAGGGCAAGTTCAACATGCCGCCAGGAGTGATGGAGCACTTCGAGACGGCGACCAGG
P K G K N M P P G V M E H F E T A T R
GACCCGGCTTTCTTCAGACTGCACAAGCACATAGACAACCTCTTCAAAGTGCACAAGGAC
D P A F F R L K H I D N L F K L H K D
CTGCTGCCACCTTACAAGCAGGAAGAGCTGGTGTGCTGCCGGGCGTGAAGATCACCGACGTG
L L P P Y K Q E E L V L P G V K I T D V
GCGGTGGAGCCGCTGGAGACGTACTTCGAGGACTTCGACGCGGACCTGCTGAACGCGCTG
A V E P L E T Y F E D F D A D L L N A L
GACGACACCATCGACCTTGACGACGTGGAGATCAAGGCGCGCGTGCGCCGCTCAACCAC
D D T I D L D D V E I K A R V R R L N H
AAGCCCTTCACCGTGCAGATCACCGCCGAGAGCGAGCGTGAGACGCTGGCCACCGTGCGC
K P F T V Q I T A E S E R E T L A T V R
ATCTTCCTCGGCCCAAGTACGACTGGTTCGGCCAGGAGATCCCCCTCGAGGAGAAGCGC
I F L G P K Y D W F G Q E I P L E E K R
CAGTACCTCGTCGAGATAGACAAGTTCTGTCATAAGGTGAATTCGGGCAAGACGACGATC
Q Y L V E I D K F V T K V N S G K T T I
CAGAGGAAGTCCAGCGAGTTCGAGTGTGACCATCCCTGACCGCGAGACGACCAAGCTGCTG
Q R K S S E S S V T I P D R E T T K L L
GTCGCCAAGGTGGAGAACGCCATCAACGGTAAAGCCACCATCAACGTCAACAAGGACGTG
V A K V E N A I N G K A T I N V N K D V
CGTCACTGTGGCTACCCTGACCGCCTGCTGCTGCCTAAGGGCAAGAAGGGAGGCATGCC
R H C G Y P D R L L L P K G K K G G M P

TTCACACTCTACGTCATCCTCACAGACTTCAATAAGGAGAAGGTGAACGACCTGCCCTAC
 F T L Y V I L T D F N K E K V N D L P Y
 GACTACGACTACGGCGGCTCGCTCAGCTACTGCGGCACCATCAACCACAAGTACCCGGAC
 D Y D Y G G S L S Y C G T I N H K Y P D
 AGCCGTCCCATGGGCTTCCCCTTCGACCGCCGCCTGTCCACCGAGGAGTTCAGCATGCCC
 S R P M G F P F D R R L S T E E F S M P
 AACTTCTGCTCCAAGGACGTCACCATCACCTTCAAGGGCGACAGCAGCTAGGGCTGGAAC
 N F C S K D V T I T F K G D S S ***
 ATAAGTGGCGGCTGTTGGCGTCACTGAGGCAGGGCGCGGGCGACCAGACAGTAGGGTGTG
 TTCATCGAACCTTGGTTGAACTCGTCACTAGAATAAACGACAGTGAATGCTCTCACACAA
 ACCTATGACCAACATTTAGCAAGCTCCTATAAAAGCCAGATTTAAAAAAAAAAAAAAAAAAAA
 AAA [2310 bp]

Fig. S5. cDNA and deduced amino acid sequence of *C. morosus* hemocyanin subunit 1 (CmoHc1, FM242640). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

AATTG

AGGTCAAGCTATCAAAGTTAGTAGTTGATAAAAAAAAAATCCCTCAGCAGGGGCAACAGCG
GCCAGCGGCAGAGCATGACACATGTGTGTGGGGGAGTTGGTGgGGAGAAAAAAGATAGGA
AGATGTGTCAATGAAACACCCGGCACAGATTCACACGGCACTTTTTCCACTGGCTATATA
AGATCGGGTCGAAACCGTGGTCAGCAGAGTCTTCGGCTAGTAGTCGATTCTCCTTTAACA
ATGAAGTTGATACCAGCTTGTCTGGTAGTAGCTGTTTGCATTGCTTACTGTTGTGGCGTA
M K L I P A C L V V A V C I A Y C C G V
AGTGTCCCAGCTGATTCAAGCCTTGTGTGAGGCAGTATGAAATCTTAAACTCCTAAAT
S V P A D S G L V V R Q Y E I L K L L N
AAAGTTCATGAACCCAATCGAATCAAAGAACAAATTGAAATTGGAAACTCCTTCGATCTT
K V H E P N R I K E Q I E I G N S F D L
TCTACTGCCCTGACTGGATTCAAGAACCCGAATGCCGCCAAGGCTCTTTTGAAGGGCTAC
S T A L T G F K N P N A A K A L L K G Y
AAGGCAGGGTCTCTTCTGCCTCGTGGATCCATATTCACTCTCTTCAATGAAAAACATCGA
K A G S L L P R G S I F T L F N E K H R
AGTGAATGATTCTGCTTTTCGAAACCTTTTCTACGCCAATGACTGGGATCTGTTGTTT
S E M I L L F E T F F Y A N D W D L L F
AAGACTGCCTGCTGGGCTAGGGACCGTATCAACGAAGGTCAATTCGTATACGCCCTTTG
K T A C W A R D R I N E G Q F V Y A L S
GTCGCCGTTTTGCACCGAAGTGATACTCACGGCATTGTTCTTCCACCTCCTTATGAAATC
V A V L H R S D T H G I V L P P P Y E I
TACCCCACTTGTGTGTTAACAGCGAAGTAATCCATGCTGCCTACAAAGCCAAGATGCGA
Y P H L F V N S E V I H A A Y K A K M R
CAAGAACCAGCCATTGTTAGGATGAACCTTACTGGAATATCAAAAACCTGAACAACGA
Q E P A I V R M **N F T** G T I K N P E Q R
GTTGCCTATCTGGAGAAGATATCGGAATGAATGCTCACCATGCTCATTGGCATATGGAC
V A Y L G E D I G M N A H A H W M D
TTCCCATTTCTGGTGAAAGAACATGAATATGGTATTCATAAAGACCGAAAAGGAGAACTT
F P F W W K E H E Y G I H K D R K G E L
TTCTACTACATGCATCATCAATTGATTGCAAGATTTGATCTTGAACGACTCTCGAACAAT
F Y Y M H Q L I A R F D L E R L S N N
TTGCCTTTTGTGAACTCTTTTCATGGGACCAGAAAATCGCAAATGGTTTCTATCCTCAA
L P F V E P L S W D Q K I A N G F Y P Q
ACCACCTACCGAGTTGGTGGTGAATTCCTGCCAGGCCAGATAACTATGCCTTCCACGAT
T T Y R V G G E F P A R P D N Y A F H D
TTGAAAAACATTAATAAAGACATGATCGACTACACTCGCAGAATCCGTGAAGCCATT
L E N I K I K D M I D Y T R R I R E A I
GCACAACAGGCCGTAATCACGAAATCCGGTGAATATTTTTACTTAAATGATTCAAAAGGC
A Q Q A V I T K S G E Y F Y L **N D S** K G
ATTGAAGTTCTTGGTGACCTTATGGAACCATCTTTCGATTGCGAAACACCCAGAATACTAT
I E V L G D L M E P S F D S K H P E Y Y
GGAGCTCTTCATAACTATGGACACATTTTGTCTGGCCAAATCACAGATCCAAAAGGAAAA
G A L N Y G I L G Q I T D P K G K
TTCAATATGCCACCTGGTGTGATGGAACACTTCGAGACTGCTACCCGTGATCCTGCCTTC
F N M P P G V M E H F E T A T R D P A F
TTCAGGTTGCATAAATATATCGACAATCTCTTCAAATTACACAAAGATCTGCTACCTTCT
F R L K Y I D N L F K L H K D L L P S
TACAGCAAAGAAGAGCTGAGCCTTGATGGTGTACAGATCGAGGATGTACAAATCGATGAA
Y S K E E L S L D G V Q I E D V Q I D E
TTAACTACTTACTTTGAAGACTTCGACATCGACCTTTTGAATGCTCTTGACGACACTGTC
L T T Y F E D F D I D L L N A L D D T V
GAATTGGAAGATGTTGAAATCAAGGCCCGTGTCCGTCGACTCAATCACAAACCCTTCAAC
E L E D V E I K A R V R R L N H K P F N
TTCAAGATTGAAGTCAATTCCGACAAAGAATACACCGCAGCTGTCCGAGTATACATTGGA
F K I E V N S D K E Y T A A V R V Y I G
CCGAAATATGATTGGTTTGGACAAGAAATCACTTTGGACGAGAAAAGATTATACATGGTT
P K Y D W F G Q E I T L D E K R L Y M V
GAAATTGATAAATTTGTTACTAAATTGACTGCTGGACACAACAACATTTTTAGAAAATCA
E I D K F V T K L T A G H N N I F R K S
AGTGAATCGAGTGTGACTATTCTGTAGAGAAACTACCAAAGTTCTTCATAAGAAAGTC
S E S V T I P D R E T T K V L H K K V
AAAGATGCTCTTAAAAATAGCACTCCACTCTTAGTTAACAAGGATGTCCGACACTGCGGA
K D A L K **N S T** P L L V N K D V R H C G

TACCCAGGTCGCCTTCTCTTGCCAAAGGGTAAAATCGAGGGAATGCCCTTCCCATTGTAT
Y P G R L L L P K G K I E G M P F P L Y
GCAATCGTGACCGATTTTGAACAAGAAAAAGTCAACGATCTTCCATTGACTATGACTAC
A I V T D F E Q E K V N D L P F D Y D Y
GGTGGTTCAATTAGCTACTGTGGGACTCTTGGACACAAATATCCCGACTCGAAACCCATG
G G S I S Y C G T L G H K Y P D S K P M
GGATATCCATTTGATCGTCCTATTGGCCGGGAATTCTATTATCCCAACATGTTTGAAAAA
G Y P F D R P I G R E F Y Y P N M F E K
GATGTCGTGATTACTCATAAAGAATCTTAAAGAGATTATATGTGCTGTAGTTGGAAAATA
D V V I T H K E S ***
AAAGTATCAAACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA [2326 bp]

Fig. S6. cDNA and deduced amino acid sequence of *C. acanthopygia* hemocyanin subunit 1 (CacHc1, FM242641). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATGAATGCCACCATGCCACTGGCATATGGACTTCCCTTTTTGGTGGAAACCAAGCTAC
 M N A H A H W M D F P F W W K P S Y
 AATAACTCAAAAGATCGTAAAGGAGAACTGTTTTCTACATGCATACCAATTGGTGGCA
N N S K D R K G E L F F Y M H Q L V A
 CGTTTCGATGCCGAACGACTCTCAAACGGTCTCCCTCATGTGGAACCTCTTGCCTGGGAC
 R F D A E R L S N G L P H V E P L A W D
 GAAAAGATCGTTGAAGGATTCGCTCCTGCTGCCATGTATCATAACGGCCAAGAATTCCCA
 E K I V E G F A P A A M Y H N G Q E F P
 ATGCGTCCTAATGGAGTCAGATTCCAAGACTTACCATGGATTACTATTCATGATCTTGTT
 M R P N G V R F Q D L P W I T I H D L V
 CTCTATGAAGGCCGAATTCGAGATGCTATTGCTGCTGGATTTATTAACACCTTCAGAT
 L Y E G R I R D A I A A G F I K T P S D
 GAAGTTCTCTATTTAAACAATACTCAAGGTATTGATCTTCTCGGAGCAATTATCGAATCA
 E V L Y L **N N T** Q G I D L L G A I I E S
 TCAGCCTACAGTCCAAATCCCGTTACTATGGACAACGACCAATGATGCTCACGTAGTT
 S A Y S P N P A Y Y G Q L N D A V V
 CTCTCTCGAGTCACTGATCCACAACCTCAGATATGGCGTACCACCTGGTGTAAATGGAACAT
 L S R V T D P Q L R Y G V P P G V M E H
 TTTGAAACAGCTACAAGAGATCCTGCCTTTTTTCAGACTTCACAAACATATGGATAATTTG
 F E T A T R D P A F F R L K H M D N L
 TTTAACTTCACAAAGATCATCTTCCAGTCTACACTCACCAAGAGTTGGAATTCCCCGGC
 F K L H K D H L P V Y T H Q E L E F P G
 GTCAGCGTTGAGGCAGTTAAGGTGGTTAGTACTTGCAAAGGAAGCAAACCAACAATATG
 V S V E A V K V V S T C K G S K P N N M
 TACACCTACTTCGACGAATCTAACATCGATTTGGGAAATGCTGTCCAGGGCTCTGCTGAT
 Y T Y F D E S N I D L G N A V Q G S A D
 GTTGAATTAAGCTCAAGTTTCTCGTTTGAGCCACGAGCCATACAAATATATTTTGACT
 V E I K A Q V S R L S H E P Y K Y I L T
 ATTCATGCAGAACATGACATAGAAGGATACGTGAGAATTTATTTGACTCCGAAATATGAC
 I H A E H D I E G Y V R I Y L T P K Y D
 TGGTTTGGCCAAAGCGTTCCGCTTGATACTGCTCGTTGGTCCATGTTTGAAGTTGATAGA
 W F G Q S V P L D T A R W S M F E L D R
 TTCCCAGTCAAACCTAAGACCCGGTGATAATATTGTTAAGAGAGATTCAGATGAATCTTCT
 F P V K L R P G D N I V K R D S D E S S
 GTAACCATTTCTGAACCTCACAGCTTCCCTAATTTAATTGCTGAAGTCGAACAGGCCATT
 V T I S E P H S F P N L I A E V E Q A I
 AACGGAACGGGAGATTACGAAATCGACAAGTACAATCGCCATTGCGGTTTCCCTCATCGC
N G T G D Y E I D K Y N R H C G F P H R
 CTTCTGTTGCCTAAAGGAAGTGAATACTCTCTCTACGTTGTCGTTTCC
 L L L P K G T S E G M Q Y S L Y V V V S
 GATTATAAATCAGATAATATTGTTCCAACATGACAATGGAAGACACCGCGCGTCTCTCT
 D Y K S D N I V P T M T M E D T A R L S
 TCTATCGCTTATTGTGGAGTGGTCGATGGAAAGATTCCAGATGGAAAACCAATGGGCTAT
 S I A Y C G V V D G K I P D G K P M G Y
 CCATTTGACAGAAAAATCGTTTCCGAGGAAGAGTTTTTGAAGAAAATATGAAAGTTATTG
 P F D R K I V S E E E F L K K I ***
 ATGTTACAATAAACCATGTATCCAATGAGCAATTTTGAATGTTTAATTGTAAACCCAAT
 TAGAATTTGACTTGATGAAATAGAGAATATTCTTATGACCAAATAAAAAAAAAAAAAAAAAA
 AAAAAA [1506 bp]

Fig. S7. Partial cDNA and deduced amino acid sequence of *C. acanthopygia* hemocyanin subunit 1 (CacHc2, FM242654). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATAATCGGAAGTGCCTAGTCCAGTCGGCACCCGCTGGC
AATAGACATATTTTCCTGTGTATGTGCGAAACAGTTGAAGTAATTTCCAAACAGATGGCAGG
ATGGGCTTGGTAAAGTACACTGGTTTCAGCTTCAGCGGACAAGGATTTCTTGTAAGCAG
M G L V K Y T G S A S A D K D F L V K Q
AGACAGATTCTACGGCTATGCAACAAAGTGCACGAACCCAACAGATTCAAGGAACAGGCC
R Q I L R L C N K V H E P N R F K E Q A
GAACTTGGGAAGAACTATGAACCTCTCAAACATCTGGACAAATACAAGCACCCAAGAAGC
E L G K N Y E P L K H L D K Y K H P R S
GTGAAGACACTAGTCAAGTACTACAACCTCTGGACTGACGCTCCCCAGAGGTGCTATCTTC
V K T L V K Y Y N S G L T L P R G A I F
ACCCTGTTTAATGACAAGCATCGTGAGCAGATGATCACGGTCTTCGAGTCGCTGTTTTTC
T L F N D K H R E Q M I T V F E S L F F
GCACAAGACTGGGACACCTTCTGAGGACAGCCTGCTGGTGCCGTGACCGTGTCAACGAG
A Q D W D T F L R T A C W C R D R V N E
GGACAGTTCGTGTATGCTCTTAGCGTAGCTGTGTTGCACCGGGAAGATACCCGAGGAATT
G Q F V Y A L S V A V L H R E D T R G I
GTACTTCCACCACCCTACGAAATATACCCCCATCTTTTCGTCAACAGCGAAGTCATCCAT
V L P P P Y E I Y P H L F V N S E V I H
GCTGCATACAAGGCAAAGATGCGACAGGAGCCGGCAATTGTCAGAATGAACCTTCACAGGA
A A Y K A K M R Q E P A I V R M N F T G
ACGATCAGGAACCCCGAGCAGCGCGTGGCATACTGGGCGAGGATGTAGGCATGAATGCC
T I R N P E Q R V A Y L G E D V G M N A
CATCACAGCCACTGGCACATGGACTTCCCTTTCTGGTGGAAGCAGCAGGAATATGGCGTG
H S H W M D F P F W W K Q Q E Y G V
GAGAAGGATCGCAAAGGGAACTGTTTTACTGGATGCATCACCAACTTATTGCACGTTTC
E K D R K G E L F Y W M H Q L I A R F
GATTTGGAACGCTTTCCAACGACCTGCCCTTCGTGCAACCTCTGTCATGGGAAGATAAG
D L E R L S N D L P F V E P L S W E D K
ATCGTAGATGGATTCTACCCACAGACAACATACCGCGTAGGAGGCGAATTCCCCGCCAGG
I V D G F Y P Q T T Y R V G G E F P A R
CCCGACAACTTTGAGTTCCAGGATCTCGAGCACATCAAAGTAAAGGACATGGTGGACTAC
P D N F E F Q D L E H I K V K D M V D Y
ACCAGAAGGATTAGGGAAGCTATTTCTCAAATGGCCGTCGTAACCTAAGAGCGGAGAATAC
T R R I R E A I S Q M A V V T K S G E Y
TACTCTCTGAACAATACTAAAGGAATCAACACTTTGGGCGAAATCATGGAGCCTTCCTTT
Y S L N N T K G I N T L G E I M E P S F
GATTCCAAGCACCCAGAATACTACGGAGCTCTTCACAACCTACGGCCACATCATGTTGGGA
D S K H P E Y Y G A L N Y G I M L G
CAAATCGTCGACCCTAAGGGAAAATTCAATATGCCACCAGGTGTAATGGAGCACTTCGAG
Q I V D P K G K F N M P P G V M E H F E
ACGGCGACCCGTGATCCCGCTTTCTCCGATTGCACAAGTACATCGACAACCTGTTCAA
T A T R D P A F F R L K Y I D N L F K
ATACACAAGGACCTGCTGCCAGTCTACACACATGAGGAGCTTGACTTCCCTGGTCTGAAA
I H K D L L P V Y T H E E L D F P G L K
GTCGTAGATCTGGACGTGGATGTGCTGCAGACTTATTTGGAAGACTTCGACATTGACCTG
V V D L D V D V L Q T Y L E D F D I D L
CTGAACGCACTAGATGACACTCAGGAACTAGATGACGTGGAAATTAAGGCCCGAGTACGC
L N A L D D T Q E L D D V E I K A R V R
CGCCTCAACCATCAGCCATTACCTTCCGTGTTACAGTTTCAAGTGAAAAGAATGCAAT
R L N H Q P F T F R V T V S S E K N A N
GTCGCTGTCCGCGTGTTCCTTGACCCAAAGTACGACTGGTTTGGAGTACAGATCCCACTG
V A V R V F L G P K Y D W F G V Q I P L
AATGAGAAGCGACTCTACATGGTGGAAATCGACAAATTTTACACCACAGTGAACAAGGGC
N E K R L Y M V E I D K F Y T T V N K G
ACAACAGTGATAGAGAGGAAATCCAGTGAGTCTAGTGTGACAATCCCTGACCGAGAAACC
T T V I E R K S S E S S V T I P D R E T
ACAAAGGTGCTGCAAAAGAAAGTGGAAGATGCCATTCAAGGAAACGGAGCGTTTCTTGTC
T K V L Q K K V E D A I Q G N G A F L V
AATAAGGATGTGAGGCACTGTGGATATCCTGATAGACTGCTTCTGCCCAAAGGCAAGAAA
N K D V R H C G Y P D R L L P K G K K
GATGGAATGCCCTTCACTTTCTATGTGATTGTGACTGATTTTGATAAGGAGAAGGTCAAT
D G M P P F T F Y V I V T D F D K E K V N
GATGTTCTTTGGACTACAACCTATGGAGGATCCATCAGCTACTGTGGAACATTGGCAGGA

D V P L D Y N Y G G S I S Y C G T L A G
 CACAAGTATCCTGACAACAAGCCAATGGGATTCCCCTTCGACAGAAGAATCGATGAGGAC
 H K Y P D N K P M G F P F D R R I D E D
 CACTTCTTCACGCCCAACATGGGACAGAAAGAAGTAACCATCTCCTTCAAGGAAGTTCCA
 H F F T P N M G Q K E V T I S F K E V P
 TAAACGACCCTGAATTCAGCGATTGGATACAACACTACTCTAGAGTGATATCAAGGGAAACT
 * * *
 TCGTATTTAAACGGCATTATATACTCAATACATCCACATTTTGTATCCTCTAGGTTGCTCA
 ATAAAAATAGCATTTTAAATGCCTATTGTTATGTAACATTTATAGTTCTCGTTGATTTCAA
 TAAAGACAGTTCAGAAAGCCTAGTTCAGAGATATTATCTAATTAAAAATATTAGAGCAAA
 TCGTGAAACAAGTATACCTGTTTCGGCACAAAATGTTAAACAGATCCACGCTTAATAGAG
 TTAAGTGCAGTTTATTTTCTGAATTTTCTTCTTGTATTACATATTCAGTTTCACAGACT
 GCTAGAAACATTACAACCTCCTGAACAGCACACAGCTACATACAAAACCAAGGGTCTATG
 AATTCAGATCTTTGGAGATATTATTGCATAATTTTGTATTGTCTTTTAAGATGAATAAAC
 TGTATGTCACTAAAGAAAAAAAAAAAAAAAAAAAA [2592 bp]

Fig. S8. cDNA and deduced amino acid sequence of *H. membranacea* hemocyanin subunit 1 (HmeHc1, FM242642). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

GCAGTTGTAAAGTCGGACAGGAAGCA
ATTGCAAGAACTGACCGTGCCATCGGAGCTTTTTTTTTTAAAGAAAGGTGGTTCACGTTGT
ATGGCTGATGCAACATTCCCTGAAGAGACAACTCCATGTGTTGAAGCTCTTGCGTCACATT
M A D A T F L K R Q L H V L K L L R H I
TCCGAACCAATTACGGACCCTGAACTCCATGCCAAGAAGCAATCTGGGGATCCTGCAAAC
S E P I T D P E L H A K K Q S G D P A N
CACCTCAACAAATTTAATGACCCCAAAGCCGTTTCGACACTTTGTGAACCTCGTAAGAAAG
H L N K F N D P K A V R H F V N L V R K
GGAAGTCTGCTCAGCAGAGACACGGTATTAGTTTGTTCACCCAAAACACCGAGAACAA
G S L L S R D T V F S L F N P K H R E Q
ATGGTGGCCTTGTGTTGAAGTGCTCTACTCTGCAGTCGACTACGATACCTTCTTCAACACG
M V A L F E V L Y S A V D Y D T F F N T
GCAGTGTGGGCCAGAGACAGGATGAACCCCAGGCAGTTCCTCTACTCCTTCAGCGTTGCT
A V W A R D R M N P R Q F L Y S F S V A
CTTCTGCATCGTCCAGATTGCCGGGGCTTGACACTTCCACCTGCGTACGAGATAACGCC
L L H R P D C R G L T L P P A Y E I T P
CACATGTTCTCACCACGGATGTTGTGAGGGAGGCATACAAGGCAAAGATGACGAGGCAA
H M F L T T D V V R E A Y K A K M T R Q
CCCGTGTGATCCCAATGAAATTTACGGGATCAGTGAATAACCCAGAGCAGAGGGTGGCA
P V L I P M K F T G S V N N P E Q R V A
TATTTTGGAGAAGACATTGGCATGAACTCTCATCATGCACACTGGCATATGGACTTCCCT
Y F G E D I G M N S H A H W M D F P
TTCTGGTGGAGCCGATTATCAACACAACCATAGCAAAGATAGGAAGGGAGAAGTCTTC
F W W K P H Y Q H N H S K D R K G E L F
TTCTACATGCATCAGATGTTGCTCGATTGACGCCGAGAGACTTTCAAACAACCTT
F Y M H Q M V A R F D A E R L S N N L
CCTGTGGTGAACCTCTGGAATGGAACAGAAAGATAGTTGAAGGATTGCTCCAGGAGCC
P V V Q P L E W N R K I V E G F A P G A
ATGTACGATAACGGACAGGAGTTCCCATGCGCCCTGACAATATCGGTTTCCAGGATCTG
M Y D N G Q E F P M R P D N I G F Q D L
CCATGGCTCACTGTTGAGGATATGCTCATGTATGAAGGACGAATCCGGGATGCCATTACC
P W L T V Q D M L M Y E G R I R D A I T
GCAGGCTTTGTCAGGGCGTATGATGGAAGTACGTCTCGCTGAACAACACGGAGGGTATC
A G F V R A Y D G N Y V S L N N T E G I
AACGTGCTTGGAGAGATAGTCGAATCATCGGCCACAGCGTGAACCCAGCGTTCTACGGT
N V L G E I V E S S A H S V N P A F Y G
CAGCTCCACAACGACGGACACGTAATGTTGTCCAAAGTGACGGATCCCAAGCTTCGTTAT
Q L N D G V M L S K V T D P K L R Y
GGTCTCCCTCCTGGTGTGATGGAGCACTTCGAGACGGCCACACGCGACCCTGCCTTCTTC
G L P P G V M E H F E T A T R D P A F F
CGACTCCACAAGCACGTTGACAATCTCTTCAAAATGCACAAGGACTTTCTCCCACCCTAT
R L K H V D N L F K M H K D F L P P Y
CAAACCTCACGAGCTTCTTTTCCAGAAGTTACCGTGAAGCAGTGAAAGTGATAGGGGTA
Q T H E L L F P E V T V E A V K V I G V
TGCAAAGCCAGCATTCCCAATCATCTCATCACTTACTTTGACGAGTCTCACATTGATCTA
C K A S I P N H L I T Y F D E S H I D L
GGAAATGCTGTTTACACAAGTGGTCAGCGGGAACAAGTTGACATCAAGGCTGTTGTGAGT
G N A V Y T T G Q R E Q V D I K A V V S
CGACTCAACCACGAACCATTCAAGTATGTCATTACTGTGAAGTGAAGAAGGCAGTTCAT
R L N H E P F K Y V I T V N S K K A V H
GGAGTTGCTCGAATCTTTTATACCAAAGTACAAGTGGTTTGGTCAGAAGATTCTCTT
G V A R I F L S P K Y N W F G Q K I P L
CACGTCGCCCGTTGGGGAATGATCGAGCTGGACAGGTTCCCTGTCAAACCTGGAGGCTGGA
H V A R W G M I E L D R F P V K L E A G
GATAATGTCATAACACGAACTCAGAAGATTCCATTGTGACCATCCCAGAGCCACGAAGT
D N V I T R N S E D S I V T I P E P R S
TTCCCTCATTTGATCAACTATGTCGAGAGGGCCATTAATGGTGAAGAGGAATATGTTGTT
F P H L I N Y V E R A I N G E E E Y V V
GATAAATATCATCGCCACTGTGGATTCCCTCACCAGCTGCTCCTGCCTAAGGGAAAACCA
D K A Y H R H C G F P H R L L L P K G K P
GAAGGAATGACTTACAGAATTTACGTCATTATCACCGACTACAACAAGGATATTATTACA
E G M T Y R I Y V I I T D Y N K D I I T

TCTGACTGGAAACCAGAGGATATTGAGAACCTATCATCTCTCGGCTACTGCGGAGTATTG
 S D W K P E D I E **N L S** S L G Y C G V L
 GAGGGAAAGATTCCCTGATGCAAGGCCAATGGGATTCCCCTTCGACAGACGCATTACATCT
 E G K I P D A R P M G F P F D R R I T S
 GAAGAAGAGTTCTTCACGCAGAACATGAAGTTGTTGATGTTACCATCCCGAATGTGAAG
 E E E F F T Q N M K V V D V T I P N V K
 GAAAACCGGTAGAACAGCAATAAAGACTAATTTTGGACTACCTCTTTAATCTGTCACATT
 E N R ***
 ACGGTTCCCTTTGAGTCAAAAAGAGTTCTTGTTTTACATAAGCTCCAGGATCAGATATCATA
 TTGTGCCAAGGAGTTGTAAAAGAAACATAAATTTTATGAAAACAACTTTTCTGTAAACA
 CCTTTTCTATTTAGATTGTGATGTCATCCAGATGCAAAGATAGTCTAAGTAATCTTGTAT
 TGTGCATTTAAGATCTCCAAAAGATTGTACATGTATTTGTTTCGTTTCAGACATATGTCC
 TTTATGAAACATGTATGCTGCAGAATTGTAAGAAAACTTTTATTAATGAATAAAGAAAT
 GAGTGTGTAAAAAAAAAAAAAAAAAAAA [2456 bp]

Fig. S9. cDNA and deduced amino acid sequence of *H. membranacea* hemocyanin subunit 2 (HmeHc2, FM242643). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ACTCTGCTGTGGTGGGGCGGACATC
ATGAAGTGCATCCTGACTGTGGCGCTCCTAGGGGTGGCCGTGTGCCACCTGGCTCAGGCC
M K C I L T V A L L G V A V C H L A Q A
AAGGTCATGCGGGGCGAAGCACCAGCTGACAAGATCTTCCTGTCTGAAGCAGCGGGAGGTG
K V M R G E A P A D K I F L S K Q R E V
CTGAAGCTGCTCAATAAAGTTCACGAGCCCAACCGCTTCAAAGATCAAGCGGACCTGGGA
L K L L N K V H E P N R F K D Q A D L G
CGCAGCTACGATCCTACTCAGCACCTAAGCAAATATAAGAACGCGTTGCCAGTGAAGACC
R S Y D P T Q H L S K Y K N A L P V K T
CTCGTGAAGCAGTACACGTACAGGGCGTCTGCTGCCGCGAGGACGGATCTTCACTCTGTTT
L V K Q Y T S G R L L P R G R I F T L F
AACGACAGACAACGCGAGCAGATGATCACCATATTCGAGTCCTTGTTCTACGCGGAAGAC
N D R Q R E Q M I T I F E S L F Y A E D
TGGGACACCTTCTACCGCACGGCGTGCTATGTTTCGCGACAGGGTCAACGAGGGGCAGTTC
W D T F Y R T A C Y V R D R V N E G Q F
GTTTACGCCCTGAGTGTCTGCTGTACTTCACAGGGAGGACACGCGTGGAATGGTACTGCCG
V Y A L S V A V L H R E D T R G M V L P
CCGGCGTACGAGGTCTACCCACACCTCTTCGTCAACTCCGACGTTATCCACGCAGCTTAC
P A Y E V Y P H L F V N S D V I H A A Y
AGAGCCAAGATGAGGCAGGAGCCGGCCGTAGTGCGCATGAAGTTCACAGGAACGGTCAGG
R A K M R Q E P A V V R M **N F T** G T V R
AATCCTGAGCAGAGGGTGGCCTACCTGGGAGAGGACGTGGGCATAAACTCGCATCATAGT
N P E Q R V A Y L G E D V G I N S H S
CACTGGCACATGGACTTCCCCTTCTGGTGGAAGCAGGATGAGTATGGCGTGAAGAAGGAC
H W M D F P F W W K Q D E Y G V K K D
CGGCAGGGGGAGCTGTTCTACTACATGCATCACCAGCTGATCGCCAGACTCGACTTGGAG
R Q G E L F Y Y M H Q L I A R L D L E
CGCCTGTCCAACGATCTACCCTTCGTGAAGCCACTTTACTGGGAGGACAAGATCGAGGAT
R L S N D L P F V K P L Y W E D K I E D
GGCTTCTACCCACAGACAACGTACCGGGTGGGGGGAGAGTTCCCGGCCAGACCTGATAAC
G F Y P Q T T Y R V G G E F P A R P D N
TTCGAGTTTCAGGACCTGCAGGACATCAAAGTGAAGGACATGGTTGATTATACGCGCAGG
F E F Q D L Q D I K V K D M V D Y T R R
ATACGCGAGGCAATTTCTCAGCAAAGTGTCTAATCAGAGGTGGGGCCAGAATCTGCCCTG
I R E A I S Q Q S V I T R W G Q N L P L
AACGACACGATGGGCATCAACATACTCGGCAGCTTGGTAGAGCCGTCCCAAGAATCGCCG
N D T M G I N I L G S L V E P S Q E S P
AACCCGAGTACTACGGCGCGCTCCACAACCTACGGCCACATCCTGCTCGGACAGATCACC
N P Q Y Y G A L N Y G I L L G Q I T
GACCCCAAGGGAAAGTTCGACATGCCCCCTGGCGTCATGGAGCACTTCGAGACTGCGACG
D P K G K F D M P P G V M E H F E T A T
CGCGACCCTGCCTTCTCAGGCTGCACAAGTACATAGACAACCTGTTCAAGCTGCACAAG
R D P A F A L H K Y I D N L F K L K
GACATGCTGCCTCCTTACGCAAGAGCTCGAGTTCGCGGGCGTGGCGATTTCAGGAC
D M L P P Y A R E E L E F P G V A I Q D
GTCAGTGTGGACGACCTGGTCACGTACTTCGAGGACTTCGACATCGACTTGCTGAACGCC
V T V D D L V T Y F E D F D I D L L N A
CTGGATGACACCCTAGAGCTGAAGGACGTGGAGATAAAGGCCCGGGTGGTTCGCTCAAC
L D D T L E L K D V E I K A R V R R L N
CACCGCCCCCTTCACTTTACGATCACTGTGAACAGCAACCAGGAGCAAATGGCAGCGGTT
H R P F T F S I T V N S N Q E Q M A A V
CGGATCTTCTGGGCCCCAAGTTCGACTGGTTCGGGCAAGAGATCCCCATCAACGAGAAG
R I F L G P K F D W F G Q E I P I N E K
CGGCTCTACGTCATAGAGCTCGACAAGTTTGTGCTAAAGTGAACAAGGGGGCAACCGTT
R L Y V I E L D K F V A K V N K G A T V
ATACAGCGGAAGAGCAGCGAGTCCAGCGTCGCAATTCCGGACAGAGAGACGACCAAATA
I Q R K S S E S S V A I P D R E T T K I
CTCGTGGCGGGGTTGACGACGCGTGCAGGGGAAGGCGACGTACACAGTGAACAAGGAT
L V R R V D D A L Q G K A T Y T V N K D
GTGCGCCATTGCGGCTACCCAGCGCTGCTACTCCCCAAGGGCAAACGAGACGGCATG
V R H C G Y P E R L L L P K G K R D G M
CCCTTCAGCCTCTTCGTAATTCTGACTGACTTTGATAAGGAGAAGGTGAATGATTTACCC
P F S L F V I L T D F D K E K V N D L P

TGGGACTACGACTACGGGGGCTCCATTAGCTACTGTGGTACCGTTTCGGGACACAAGTAC
 W D Y D Y G G S I S Y C G T V S G H K Y
 CCGGACAGCAAGCCCATGGGATTCCCCTTCGATAGGCAAATCAACAGCGATAACTTCTTC
 P D S K P M G F P F D R Q I N S D N F F
 AGATCCAACATTTACCAGAAGGACGTTGTAATCACCTTCAAGGACAGCGAGTAGATCGGA
 R S N I Y Q K D V V I T F K D S E ***
 CAGCGCTTCCGGTCGATAGCTACACATATACTGAATTTTGCACAACTGTACATCAAATA
 ACATGTCCCATCATGGGTCTGTATATTATTACGTGCCTAATCAGTATGGCTTCATTTGTG
 TTTAGATTGAAAAGTTGTGTTTCAAATAATATTTGTCATACAATCATTGTCATAATATTT
 CTATCGGAATTAATTCAAATAAAACATAGTTTGTGTAAAAAAAAAAAAAAAAAAAAAAAAA [2305 bp]

Fig. S10. cDNA and deduced amino acid sequence of *C. secundus* hemocyanin subunit 1 (CseHc1, FM242644). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

GCTTGTTTGTCTCAGCCGCCGACGTCGGA
ATGAGGCCAGCTACTGTTCTGCTTATCGTCCTGGGCGCTTGCGCACTTGCGTCTGCCGAA
M R P A T V L L I V L G A C A L A S A E
CAAAAAGCGGACGCAATCTTTCTGAAGCGACAGCTGCAGGTACTGAACCTGCTACGTCAC
Q K A D A I F L K R Q L Q V L N L L R H
CTGAACCAGCCGCTGATCGATCCTGAACTCTTAGTCTTCAAGAACTCTTTCAGCCCACTG
L N Q P L I D P E L L V F K N S F S P L
AACAACTCTGAACAAGTTCAAGGACCCGACGGCAGTAAAGAACTTTCTGAACTGGTGACC
N N L N K F K D P T A V K N F L K L V T
GAGGGGAAGCTACTCGAACGTGATGCAATCTTCAGCTTGTTTCGACGCGCGCAACGCGCC
E G K L L E R D A I F S L F D A R Q R A
CAGATGGTGACGCTGTTTCGAGGTGCTGTATGGCGCCACTGATTACGATACCTTCCTCAAG
Q M V T L F E V L Y G A T D Y D T F L K
GCAGCCGCATGGGCGCGGACCGCACCAACCCTCGCCAGTTCCTGTACGCTTTCAGCGTG
A A A W A R D R T N P R Q F L Y A F S V
GCACTGTTGCACCGGGAAGACTGCAGGGGCGTGATACTGCCCCCGCGTACGAGATCACT
A L L H R E D C R G V I L P P A Y E I T
CCCCACATGTTCTTACCACCGATGTCGTCAGGAAGGCTTACCAGGCCAAAATGATGAGG
P H M F L T T D V V R K A Y Q A K M M R
ACGCCCACTGTGATTCCCATGAAATTCACGGGATCGGTAAACAACCCTGAGCAGCGGGTA
T P T V I P M K F T G S V N N P E Q R V
GCATACTTTGGAGAGGACATCGGCATGAACTCGCATCACAGCCATTGGCACATGGACTTC
A Y F G E D I G M N S H S H W M D F
CCCTTCTGGTGGAAGCAGGAATACACAGTGGACAAGGACCGGAAGGGAGAAGTGTCTTC
P F W W K Q E Y T V D K D R K G E L F F
TATATGCACCATCAGATGGTCGCCAGGTTTCGACGCCGAGCGTCTCTCAAACGACCTGCCA
Y M H Q M V A R F D A E R L S N D L P
GCGGTGGAGCCACTGGAGTGGGACCAGCAGATCACTGAGGGCTTTGCCCCGCGGGCGACC
A V E P L E W D Q Q I T E G F A P A A T
TACGAGAATGGGCAGGAGTTTCCCATGCGGCCAGATAACATGAAGTTCTGTGACTTGCCT
Y E N G Q E F P M R P D N M K F C D L P
TCCCTCTCAGTTCGAGATATGAAGATCGTCGAGGGTTCGCATCAGGGATGCTATCGCCTCT
S L S V R D M K I V E G R I R D A I A S
TCCTTCGTCAAGACGGCCGACTCGCGGCTCCTATCTATCAACAACACGGCCGGCGTCAAC
S F V K T A D S R L L S I **N N T** A G V N
CTGCTAGGGGAGATCATCGAGTCCTCAGCGAACAGTGTCAACCCTGTATACTACGGTCAG
L L G E I I E S S A N S V N P V Y Y G Q
CCGCACAACGACGGTCACGTGATGCTGTCCAAGGTGACGGATCCTCTCCAGCGCTACGGG
P N D G V M L S K V T D P L Q R Y G
GTCCCCCTGGCGTCATGGAGCACTTTGAGACTGCGACTCGCGATCCTGCCTTCTTCAGG
V P P G V M E H F E T A T R D P A F F R
TTGCACAAACATGTGGACAACCTCTTCAAGCTGCACAAAGACTTGCTCCTACCCTACGAC
L H K H V D N L F K L K D L L L P Y D
GTATCTGAGTTGGACTTTCTGGCGTGAAGATAGAAGCTGTAAAGGTGTTGGCAGCTGC
V S E L C D F A P G V K I E A V K V V G S C
AAGGCCAGCACACCAATCAGTGATCACTTACTTTGACGAGTCGCACATTGATCTGAAC
K A S T P N Q L I T Y F D E S H I D L N
AACATTGTGGAGAACGTAGGGCGGGAGCAACCCGTGGATATCAAGGCAGTGGTAAGCCGA
N I V E N V G R E Q P V D I K A V V S R
CTGAACCATGAACCCTTCAAGTACGTCATCACAGTCAACTCCAAGAAGCCTGTAAGGGGG
L N H E P F K Y V I T V N S K K P V R G
ATTGTCCGTATATTCTTGCACCTAAGTTGAACTGGTTCGGCCAGCGAGTGCCCCTGGAC
I V R I F L A P K L N W F G Q R V P L D
GTGGCCCGTTGGGGGTTTATTGAGTTGGACCGATTTCCAGTTAGACTGGCGAGCGGAGAC
V A R W G F I E L D R F P V R L A S G D
AATATAGTGTACGCAATTCTGTGGACTCCATCATCACCGTACCCGAGCCAGAAAGCTTC
N I V S R N S V D S I I T V P E P R S F
TCGAACCTTGTAAGCGATGTCGAGAAGGCAATACGCGGCGAACAGGTGTTTCAGCTACGAC
S N L V S D V E K A I R G E Q V F S Y D
AAGTATTATCGACACTGCGGTTTTCCCTCGCCGCTGTTGCTGCCCAAGGGAAAACCTGAA
K Y Y R H C G F P R R L L L P K G K P E

```

GGCATGGTTTATCAGCTGTACGTTGTTGTGACTGATTATGACAAAGACATTGTGGAACCA
G M V Y Q L Y V V V T D Y D K D I V E P
GGAATACAGCCGGAAGATGTTGAGAACCTGAGCTCCCTTGGCTACTGCGGTGTCATGGAT
G I Q P E D V E N L S S L G Y C G V M D
GGAAAGATACCTGATGGCAGGCCCATGGGATTCCCCTTCGATCGACGCATTGTGTCTCCC
G K I P D G R P M G F P F D R R I V S P
GAACAATTCCTTACTCCCAACATGGAGGTCGTGGATATCACTATCAAGAACGTGAACCAG
E Q F L T P N M E V V D I T I K N V N Q
AAGCCAGCAGAGTAGACAGCTCTTGAGATGGTGTCTGCCTTTAGGGAAGTGAAGTAAAT
K P A E ***
GGCATTCTACCAAGTTTCCTGTCTTTCCAGAATATGTGACTGGATAAAATTGCCACAAATG
ATTACATTTATAGCTCTGAAAGCTTTCTTTAGTCATTGAATATTTTTCTAACGTTATAAA
TAAAGTCTGGACTTTTCTGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA [2293 bp]

```

Fig. S11. cDNA and deduced amino acid sequence of *C. secundus* hemocyanin subunit 2 (CseHc2, FM242645). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATCCCAAC

ATGAAGACGCTGCTGTTCTACTACTAGGTGTGGCCTTGTACCATCAGGCCTGCGCAAGA
M K T L L F L L L G V A L Y H Q A C A R
CAGGCACAAATCAACAGGGGAGTTCCCGTTGACAGGGAATTCCTCACGCGACAGAATGAG
Q A Q I N R G V P V D R E F L T R Q N E
ATCTTGCATCTGCTCAACAAAGTCCACGAGCCCAACCGTTTCAAGGATCAATCGGAAGT
I L H L L N K V H E P N R F K D Q S E L
GGAAGAAGCTATGACCCCACTGCCCACTTCGACAAATACAAGAACGCAAGTCCTGTGAAG
G R S Y D P T A H F D K Y K **N A S** P V K
TCACTGGTGAAGCTGTACACATCTGGCCAGCTACTGCCTCGTGGTCACATTTTCACGCTG
S L V K L Y T S G Q L L P R G H I F T L
TTCAACGACAAACATAGAGAACAGATGATTACAGTTTTTCGAGTCATTATTTTACGCTCAA
F N D K H R E Q M I T V F E S L F Y A Q
GATTGGGACACATTCTATCGCACAGCCTGCTGGGCTCGAGATCGTGTCAATGAAGGACAA
D W D T F Y R T A C W A R D R V N E G Q
TTCGTCTATGCCCTAAGCGTGGCAGTGTGCACAGGGAAGACACGCGTGGTATCGTTCTT
F V Y A L S V A V L H R E D T R G I V L
CCCCCAGCATACGAAATTTATCCTCATCTCTTCGTGAAGTTCAGAAAGTTATCCATGCCGCC
P P A Y E I Y P H L F V N S E V I H A A
TACAGGGCCAAGATGCGTCAAGAACCTGCGGTCTGCGCATGAAGTTCCTGGAACCGTC
Y R A K M R Q E P A V V R M **N F T** G T V
AGGAACCCAGAGCAGAGAGTGGCTTATTTGGGAGAAGACGTGGGAATGAATTCACCAT
R N P E Q R V A Y L G E D V G M N S H
TCTCACTGGCAGATGGATTTCCCTTTCTGGTGGAAAGAACAAGAGTACGGAAGCCATAAG
S H W M D F P F W W K E Q E Y G S H K
GACCGCAAAGCGCAACTCTTCTATTGGATGCATCATCAACTTATCTCACGTTTCGATCTT
D R K G E L F Y W M H Q L I S R F D L
GAACGTCTTTCTAACGACCTTCCATTTGTGGAGCCTTTATACTGGGACGAGAAGATCGTC
E R L S N D L P F V E P L Y W D E K I V
GATGGATTTTATCCTCAAAGTACTTACAGAGTAGGTGGCGAATTCCTGCAAGGCCTGAC
D G F Y P Q T T Y R V G G E F P A R P D
AATTTTGAATTTCAAGACCTGAAGAACATCAAGGTGAAGGACATGGTGGACTACACAAGG
N F E F Q D L K N I K V K D M V D Y T R
AGAATCAGAGAAGCCATCTCACAGCAAGCCGTAATTACGAAAAACGGCGATTATTACTCG
R I R E A I S Q Q A V I T K N G D Y Y S
TTGAACGATACCACAGGAATTGACACTCTGGGCGCTCTGATAGAACCATCAGTGGACTCC
L **N D T** T G I D T L G A L I E P S V D S
AAACATCCGGAATACTATGGAGCTCTTCATAACTATGGACATATCATGCTGGGTCAGATA
K H P E Y Y G A L N Y G I M L G Q I
ACAGACCCCAAGGGGAAATTCAACATGCCTCCAGGAGTGATGGAACATTTTCGAAACAGCA
T D P K G K F N M P P G V M E H F E T A
ACTCGGATCCAGCTTTTTTTCAGACTCCACAAATACATCGACAATCTTTTCAAGCTGCAT
T R D P A F F R L K Y I D N L F K L H
AAGGATCTCCTGCCAAATATACTCAAGCAGAGCTCAGCTTTCTGGCGTAAAAGTGAGC
K D L L P K Y T Q A E L S F P G V K V S
GATGTAAGTGTGGACGAACTCGTCACCTACTTCGAGGATTTTCGATATCGACTTGTTGAAC
D V T V D E L V T Y F E D F D I D L L N
GCCCTGGATGATACTCAAGAATTAGATGACGTAGAAATCAAGGCACGCGTACGCAGGTTG
A L D D T Q E L D D V E I K A R V R R L
AACCACAAACCCTTCACTTTCCATATTACCATCAACAGTGAGCTAGACGCCACAGCAGCT
N H K P F T F H I T I N S E L D A T A A
GTGAGAGTTTTCTGGGTCCAAAATATGACTGGTTTCGGTCAAGAGATCCCCATCAATGAG
V R V F L G P K Y D W F G Q E I P I N E
AAGCGTCTCTACATGGTTCGAGATTGATAAGTTCGTTGCTAAAGTGAACAAAGGGACAACA
K R L Y M V E I D K F V A K V N K G T T
GTCGTGCAGCGTAAGAGTTCTGAGTCGAGTGTCAACATTCCGGACAGAGAACTACTAAG
V V Q R K S S E S S V T I P D R E T T K
GATCTCGTCAAGAAAAGTTCATGATGCTCTTGAAGGGAATCCAGCTTCTACGTGTGCAAG
D L V K K V H D A L E G K S S F Y V C K
GATGTGCGCCATTGTGGTTATCCTGACCGCTGCTTCTGCCCAAGGGCAAGAAGGAAGGA
D V R H C G Y P D R L L L P K G K K E G
ATGCCCTTCACTCTATATGTTATTGTACCGACTTCGAGAAAGAAAAGGTGAACGACTTA
M P F T L Y V I V T D F E K E K V N D L

CCTTGGGACTACGACTACGGTGGCTCAATATCATACTGCGGTGTGCTGTCCGGCCATAAG
 P W D Y D Y G G S I S Y C G V L S G H K
 TACCCAGACAGCAAGCCCATGGGCTTCCCATTCGACAGACATATTGATGGTGATCATTTC
 Y P D S K P M G F P F D R H I D G D H F
 TTCACTGAAAATATGTACCAGAAAGAAGTCAAAATTACTTTCCAGGACAATGAATGAATA
 F T E N M Y Q K E V K I T F Q D N E ***
 TTTTCATTACTTTATACCACTGATGCAATAAAATACATTTTACTTAAAAAAAAAAAAAA
 AAAAAAAAAA [2118 bp]

Fig. S12. cDNA and deduced amino acid sequence of *B. dubia* hemocyanin subunit 1 (BduHc1, FM242646). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATTGTTGGCCCTAGTC
ATGAAGTTGCTCGGGATTCTGGTGTCTGTCGCCTGCGTCGTCGCCGCCGGTTCGCCGTC
M K L L G I L V S C A C V V A A A V A V
CAACAAAAGGCGGACGGGGTTTTCTGAACCGGCAGATAAGCGTATTA AAACTACTGCAG
Q Q K A D G V F L N R Q I S V L K L L Q
CATCTGAACCAGCCACTGGTGGACTCTGATCTCACCGCGTATAAGGAATCATT CAGCCCG
H L N Q P L V D S D L T A Y K E S F S P
CTGAACAACCTGAACCGCTATCGAGATGCTGAAGTGGTGAAGCAATACGTGGGACTCGTC
L N N L N R Y R D A E V V K Q Y V G L V
GAGTCCGGCAAGCTCCTGCACCGCATGGAATCTTCGGCCTCTTCAACCCAGAGCATCGC
E S G K L L H R D G I F G L F N P E H R
CAACGATGATCGTGCTCTTCGAGACCTTGACGGAGCTTCAGACTGGGATACCTTCTTC
Q Q M I V L F E T L Y G A S D W D T F F
AAGACTGCAGTGTGGGCTCGAGATCGCACCAACCCTCGCCAGTTTCTCTATGCCTTTAGT
K T A V W A R D R T N P R Q F L Y A F S
GTTGCCCTTCTCCATAGGGAGGACTGCAGAGGAATCAAGCTACCGCCCCAGTACGAGATC
V A L L H R E D C R G I K L P P Q Y E I
ACCCCTCATGTTTTCTTACTACTGATGTCGTCAGGAAGGCTTACCAGGCCAAGATGACA
T P H V F L T T D V V R K A Y Q A K M T
AGGACTCCAACAGTAATTC CATGAAGTTCACTGGATCGGCGAACAACCCCGAACAAAGA
R T P T V I P M K F T G S A N N P E Q R
GTAGCGTACTTCGGAGAAGACATCGGGCTCAACTCTCATCACGCTCACTGGCACATGGAC
V A Y F G E D I G L N S H A H W M D
TTCCCGTTCTGGTGGAAACCTGAGTATGACGGAGACAAGGACCGCAAAGGAGAACTGTTT
F P F W W K P E Y D G D K D R K G E L F
TTCTACATGCATCACC AAGCAGTGGCAAGGTTTGATGCCGAGAGGCTTTCTAATGATCTG
F Y M H Q A V A R F D A E R L S N D L
CCCGTAGTGCACCCATTGGAATGGGACAAGAAGATCGTGGAAGGATTTGCACCTGGCGCA
P V V Q P L E W D K K I V E G F A P G A
TCCTATGAAAACGGACAAGAGTTCCCATGAGGCCTGACAACATGCAGTTCTCGGACTTG
S Y E N G Q E F P M R P D N M Q F S D L
CCCTGGCTTACTGTGCAGGATATGAGGTTGTACGAGGGACGCATTAGGGATGCTATTGCT
P W L T V Q D M R L Y E G R I R D A I A
TCTGGCTTCGTAAGGACGGCCGAGGGAAAATTGGTCAGTCTGAACAACACCGTTGGAATC
S G F V R T A E G K L V S L N N T V G I
AATTACCTGGGGGAAGTGGTGGAAATCGTCAGTTCACAGCGTGAATCCCGTCTTCTATGGC
N Y L G E V V E S S V H S V N P V F Y G
AAACTACACAATGATGCCACGTGTTGTTGTCCAAAGTGACTGATCCCAAGTTGCGGTAC
K L N D A V L L S K V T D P K L R Y
GGTGTACCACCAGGAGTAATGGAACATTTCGAAACTGCGACACGCGACCCAGCCTTCTTC
G V P P G V M E H F E T A T R D P A F F
AGATTGCACAAACACGTGGACAATATCTTCAAGTTGCACAAGGATCTGCTACCCCTTAT
R L K H V D N I F K L H K D L L P P Y
GATGTGTCCGAGTTGGAGTTTCCAGGAGTGAAGTAGAGGCGGTGAAAGTGATCGGCACT
D V A S E L E F P G A V K V I G T
TGCAAGGCTAGCACACCAATCAACTAATTACCTATTTTGACGAGTCTCACATCGACCTC
C K A S T P N Q L I T Y F D E S H I D L
GGAAATGCTGTTGATACAGCAACACACAAACAGGTTGTCGACATAAAGGCAGTTGTAAGC
G N A V D T A T H K Q V V D I K A V V S
CGTTTGAACCACGAACCTTCAAGTATGTGATCACCGTAAACTCGAAACGTCCAGTGAAT
R L N H E P F K Y V I T V N S K R P V N
GCCATCGTGCGCATTTTCTTGGCACCAAAGTATGACTGGTTCCGGACAGAAAGTGCCTCTC
A I V R I F L A P K Y D W F G Q K V P L
GACGCGCCAGATGGGGCGTCATTGAATTGGACCGCTTCCAGTCAAACCTTGTAACGGA
D A A R W G V I E L D R F P V K L V N G
GACACTGTAATCACACGTAGCTCCGACCAATCGACAGTGACTATTCCAGAGCCTAAGAGT
D T V I T R S S D Q S T V T I P E P K S
TATCCTGAATTGGTAAACGATGTGGAGCGTGCTGTTCTGTTGGAAGTGGACTTCGTGTAC
Y P E L V N D V E R A V R G E V D F V Y
GATAAGCACCATCGTCACTGCGGCCTCCCCACCGTCTCCTGCTTCCCAAGGGCAAGCCT
D K H H R H C G L P H R L L L P K G K P

GAGGGTATGGACTTCAAACCTCTATGTGGTCATCTCCGACTTCGACAAAGATGTTGTATCA
 E G M D F K L Y V V I S D F D K D V V S
 TCTGAAGTGAACGTTGACGACTTGCATCACATTTCTGTCTCTCGGCTACTGCGGAGTCCTC
 S E V N V D D L H H I S S L G Y C G V L
 GATGGCAAGATTCCTGATGGTAAACCAATGGGCTTCCCCTTCGATAGACGCATACCTTCA
 D G K I P D G K P M G F P F D R R I P S
 GAAGAACTCTTCCTCACTCCTAACATGAAGGTCGTGAACTTGGTGATCAAGAACGTCCGC
 E E L F L T P N M K V V N L V I K N V R
 CAGAAACCAAGTAGAATAAAGCATATTCCTGTATAGAGTGTTTCAATAAGATACCAATAAA
 Q K P V E ***
 ATGAACAACACATTGTTATTGAAATAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAA [2171 bp]

Fig. S13. cDNA and deduced amino acid sequence of *B. dubia* hemocyanin subunit 2 (BduHc2, FM242647). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

GAAAAATGTGTGGTGGGGCAACCAAC
ATGAAGTGGCAATTGGTGACCTTCTTGGGGGTGGCACTGTGCTACCTGGCTAGCGCCAAG
M K W Q L V T F L G V A L C Y L A S A K
CTCAGCGAAGTACCAGCCGACAAAGAGTTCTTGGTGAGGCAGAGGGAGGTGCTGAGACTG
L S E V P A D K E F L V R Q R E V L R L
CTCAACAAGGTTTCATGAACCCAACAGATTCAAAGAGCAAGCTGAGCTGGGACGCAGCTAT
L N K V H E P N R F K E Q A E L G R S Y
GATCCTACAGCACACTTCAGCAAGTATAAGAACGCGGTTCTGTGAAGAGGATGGTGAAA
D P T A H F S K Y K N A V P V K R M V K
GAGTATATGGCAGGTCATATGCTGCCTCGTGGCGCCATCTTCACCCTCTTCAACAATAAG
E Y M A G A G H M L P R G A I F T L F N N K
CACCGCGAGGAGATGATAACGGTGTTTCGAGTCCTTCTTCTTCGCGGAGGACTGGGACACC
H R E E M I T V F E S F F F A E D W D T
TTCTACCGCACGGCTTGCTGGGCCCCGCGATCGCGTCAATGAGGGGCAATTCATCTACGCT
F Y R T A C W A R D R V N E G Q F I Y A
CTGAGCCTCGCTGTGCTCCACAGGGAGGACACGCGCGGCATCGTGCTGCCACCTGCCTAC
L S L A V L H R E D T R G I V L P P A Y
GAGATCTACCCGCATCTCTTCGTCAACTCCGAGGTCATCCACGCTGCTTACAAGGCCAAG
E I Y P H L F V N S E V I H A A Y K A K
ATGAGGCAGGAGCCTGCCGTAGTCCGCATGAATTTTACAGGTACAATAAGGAATCCTGAA
M R Q E P A V V R M **N F T** G T I R N P E
CAGCGCGTCGCTTATTTGGGCGAAGATGTGGGAATGAACTCTCATCACAGCCACTGGCAC
Q R V A Y L G E D V G M N S H S H W
ATGGACTTCCCCTTCTGGTGGAACAGGAGGAGTACGGCGTGACAAGGACCGCAAAGGC
M D F P F W W K Q E E Y G V H K D R K G
GAGCTTTTCTACTACATGCACCACCAGCTGATCGCACGCTTCGACGCCGAGCGCCTGTCC
E L F Y Y M H Q L I A R F D A E R L S
AACGACCTTCCCATTGGAACCACTCTACTGGGACGAAAAGATCGTAGACGGCTTCTAC
N D L P I V E P L Y W D E K I V D G F Y
CCACAGACAACATACCGTGTGGGTGGGGAGTTCGCGCACGCCCGACAACCTTTGAATTC
P Q T T Y R V G G E F P A R P D N F E F
CATGACTTGGAACACATCAAAGTCAAGGATATGATGGATTACACGCGCAGGATCCGAGAG
H D L E H I K V K D M M D Y T R R I R E
GCCATCTCTCAACAGGCCGTGTCACGAGGACGGGTGAGATGGTGTCCCTGAACAACACG
A I S Q Q A V V T R T G E M V S L **N N T**
CAAGGCATCGACATTCTCGGCTGCATGGTGGAACCTTCACACGATTCCAAGCACCCCGAG
Q G I D I L G C M V E P S H D S K H P E
TACTATGGAGCTCTGCACAACCTACGGCCACATTATGCTGGGTGAGATTACCGATCCCAA
Y Y G A L N Y G I M L G Q I T D P K
GGAAAGTTCAACATGCCTCCTGGAGTAATGGAACATTTGAGACTGCAACCCGTGACCCC
G K F N M P P G V M E H F E T A T R D P
GCCTTTTTCCGCTGCACAAATATATCGACAATCTGTTCAAGCTTCATAAGGACTTGCTG
A F F R L K Y I D N L F K L H K D L L
CCCCATTACACTATGGAAGAGCTTGGATTAAACGGAGTGACCATCCAGGACTTGACCGTA
P H Y T T M E A T L G G F T N G V T I Q D L T V
GACGATCTGGTCACGTACTTCGAGGACTTCGACATCGATATGCTGAACGCACTGGATGAT
D D L V T Y F E D F D I D M L N A L D D
ACGGCTGAACTGCAGGACGTGGACATCAAGGCGCGAGTGCGCAGGCTGAACCACAAGCCC
T A E L Q D V D I K A R V R R L N H K P
TTCTCCGTGCAGGTGACCGTCAATAGCGAACGCGATGCCATGGCTGTCGTGCGAATTTTC
F S V Q V T V N S E R D A M A V V R I F
CTCGCGCCCAAGTACGATTGGTTTCGGCCAGGAGATTCCCATCAACGAGAAACGACTGTAC
L A P K Y D W F G Q E I P I N E K R L Y
ATGATCGAGATTGACAAGTTCGTACAAAAAGTAAACAAAGGAACAACAGTGGTACAGCGC
M I E I D K F V T K V N K G T T V V Q R
AAGAGCAGCGAGTCCAGTGTCACGATCCCTGACCGCGAGACCACGAAGGTCCTTACCAA
K S S E S S V T I P D R E T T K V L T K
CGAGTGGAAGAAGCACTGCAGGGCAAGACTACTTACATGGTGAACAAGGATGTGCGCCAC
R V E E A L Q G K T T Y M V N K D V R H
TGTGGCTACCTGATCGCCTGTTGCTGCCAAAGGGCAAGAAGGACGGCATGATCTTCACT
C G Y P D R L L L P K G K K D G M I F T

ATGTATGTCATCGTCACGGACTACGAGACGGAAAAGGTGAACGATCTACCCTATGATTAC
 M Y V I V T D Y E T E K V N D L P Y D Y
 GAGTATGGAGGTGCTATCAGCTACTGCGGCACTCTGGCGGGACACAAGTACCCCGTCAAT
 E Y G G A I S Y C G T L A G H K Y P V N
 AAGCCCATGGGCTTCCCTTCGACAGACAGATCGACGGAGATAATTTCTGCACACCCAAT
 K P M G F P F D R Q I D G D N F C T P N
 ATGTTCCAGAAGGATGTCATCATCACTTTTAAGGATACCATATAGAAGGGACCGGGAATT
 M F Q K D V I I T F K D T I ***
 TGAGTATTTTCTGGGAATACAATTCTGTTTAATAGCTCTATATTATTTTCTTTCTGTACA
 ATAAAAATAAAGACTAAAACGCTATTTGTTAATTCTGATTTCTTCAAGGCTAGTAAAAT
 TACTAATAAACATTTGAAATAAATTCATTTACAGCAAATATAGACTGAAACGGCTAGGAT
 ATAAATGTTTCATTGAAATAAATAAAGCAATTAATACGATATTAAGTAGTGTCAGAAAAT
 ACTGTTATTGCAGAGATTTTAGAAAATGAAAATTGGCTTCTAGATAATGGAAATGTTTTA
 TGATATATATGTACAATAATGTAAAATATATGAAACAGTGTAGCCTAATTCAGTTCTCAC
 CTAATCTCAGAAATGAAAAGTCTCAACAAACCTCATCTCCATTATCATTCAATTTCAA
 TGATATTGAGTAAACATTTACACGTTGTCTCGTAATGTAATCCTGTGTTTCAAACGCTG
 CTTTGTATAACCTGGAGAAATTGGAGTATTATTGTGAGAACATTCTATATATTATATGTG
 ACAACTTGTGGAGAACAGTACAGATCCATTTACGGAGACAGAATCCATAACCGGACCGAA
 CTGTCGTAAATCTCTTTAATGTGTGTAGTGTGGTTGCCTAATGAACTGGACTAAAGTCA
 TTGTAATTGTACATAACGTGTCTAAACCATGTATCATATTATGCATTCACTCTTTAACA
ATAAAGCAACATCATATCAAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA [2844 bp]

Fig. S14. cDNA and deduced amino acid sequence of *P. americana* hemocyanin subunit 1 (PamHc1, FM242648). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

GAAAAAGTTTCTTGTGAATTATCTTGTCAATCAGGACCGGCATCATCGAA
ATGAAGTTGGTCTGGGCGCTTGCTATAGCCCTTGGCGCCTGCGCCCTTGTCTGCTGCTGAC
M K L V W A L A I A L G A C A L V V A D
CAGAAGGCGGATATGACCTTCTGACGCGACAAGTGAAGTGCTGAAACTTCTGCGGCAC
Q K A D M T F L T R Q L Q V L K L L R H
CTGAACCAGCCTCTTAGCGATCCTGAACTTACGGCTATTAAGAATTCATTCAGCCCATTG
L N Q P L S D P E L T A I K N S F S P L
AACCACTTGAACAAATTCAAGGAACCCGAAACAGTGAACTCTTAGTAGGACTGATTACC
N H L N K F K E P E T V K L L V G L I T
AAGGGTAAGGTGCTGACGCGGGACACTGTCTTCAGCTTGTTTCACCCCCGGCACCGTTAT
K G K V L Q R D T V F S L F H P R H R Y
GAGATGGTGACGCTCTTCGAGACCCTGTACGGTGCGACGGACTATGACACGTTCTTCCGT
E M V T L F E T L Y G A T D Y D T F F R
ACTGCCGTGTACGCACGCGACCGCCTCAACCCGCGCCAATTCCTGTATGCGTTCAGCGTG
T A V Y A R D R L N P R Q F L Y A F S V
GCACTGCTGCATCGTCAGGACTGCAGGGGCATCTGCCTACCGCCGGCCTACGAGATCAGC
A L L H R Q D C R G I C L P P A Y E I T
CCCCACATGTTTCTGACCACCGATGTCGTCAGGAAGGCGTATAAGGCCAAGATGACAAGA
P H M F L T T D V V R K A Y K A K M T R
ACACCAACTGTCATCCCGATGAAGTTCACTGGCTCTGTAAACAATCCCGAACAGAGGGTA
T P T V I P M K F T G S V N N P E Q R V
GCCTACTTTGGGGAAGATATTGGCATGAACTCACACCACAGCCATTGGCACATGGACTTC
A Y F G E D I G M N S H S H W M D F
CCCTTCTGGTGGAAACAGGAGTATGCGGGCTACAAAGACCGCAAGGGGAGAACTGTTCTAC
P F W W K Q E Y A G Y K D R K G E L F Y
TACATGCATCATCAAATGGTTGCCAGGTTTCGATGCCGAGCGACTATCTAACGATTTGCCA
Y M H Q M V A R F D A E R L S N D L P
GTAGTAGAACCAATTGGAATGGGACCAAAAGATCGTGGAAGGTTTTGCACCCGGCGCCACC
V V E P L E W D Q K I V E G F A P G A T
TATGAAAACGCGCCAGGAGTTCCCCATGAGGCCGGATAATATGTGCTTCAGCGATCTTCCC
Y E N G Q E F P M R P D N M C F S D L P
TTCCTCACAGTTCAAGACATGAAAATATTGAGGGCCGCATCCGCGACGCCATCGCTTCT
F L T V Q D M K I F E G R I R D A I A S
GGATATGTGAAGACGGTTGACTCCCGATTGATGAGTCTGAATAACACGGCTGGCATAAAC
G Y V K T V D S R L M S L **N N T** A G I N
CTTCTGGGCGAGATTGTGGAGTCTTCGACACACAGTGTCAACCCCGTGTCTACGGCAAG
L L G E I V E S S T H S V N P V F Y G K
TTGCACAACGACGGTCACGTGATGTTGTCCAAAGTGACGGACCCACACCTTCGCTACGGT
L N D G V M L S K V T D P H L R Y G
GTTCTCCTGGCGTAATGGAACACTTCGAGACTGCAACCCGCGATCCCGCTTTCTTCCGC
V P P G V M E H F E T A T R D P A F F R
CTACACAAGCACGTGGACAACCTGTTCAAGTTTCACAAGGACTTGCTCGTTCCCTACAGC
L K H V D N L F K F H K D L L V P Y S
TTTCATGAGCTGGAGTTCCCGGGTGTAAGATCGAGGCTGTGAAGGTCATCGGCACTTGC
F H E L E F A P G V K I E A V K V I G T C
AAGGCCAGCAGCCAAACCATCTCATCTTACTTTCGACGAGTCACACATCGACCTCAAC
K A S T P N H L I T Y F D E S H I D L N
AATATAGTGGACACTTCCGGCCAGAAAGAAGTGGTGCACATCAAGGCACTGGTAGGCCGG
N I V D T S G Q K E V V D I K A L V G R
TTGAACCACGAGCCCTTCAAGTACGTGATCACGGTACACTCTGAGAAGCCGGTGCGCGCC
L N H E P F K Y V I T V H S E K P V R A
ATCGTGCGCTCTTCCCGCTACCCAAGTACAAGTGGTTCGGTCAGATGGTTCCACTGGAT
I V R V F P L P K Y N W F G Q M V P L D
GAGGCACGCTGGGGAGCCATCGAGCTGGACCGCTTCCCTGTCTCTTGGTGCACGGTGAC
E A R W G A I E L D R F P V S L V H G D
AACATCGTTACACGTAAGTCTGAAGATTCCACAGTGACCATCCCAGAGCCACGCAGCTTC
N I V T R N S E D S T V T I P E P R S F
CCAGATCTCATGAGCGACGTGCAAAAAGCCATCCGAGGAGAGCAGGATTTCAAATACGAC
P D L M S D V Q K A I R G E Q D F K Y D
ATTCATCATCGCCACTGTGGCTTCCCACACCGCCTATTGCTGCCTAAGGGGAAACCGGAA
I H H R H C G F P H R L L L P K G K P E

```

GGCATGGTCTACAGGCTGTTCGTCGTGGTCACCGACTATGACAAGGACGTCGTGGCGCCA
G M V Y R L F V V V T D Y D K D V V A P
GGAATGTCTCCAGATGACGTAAATACACTGTCATCTCTGGGTTACTGTGGTGTGATGGAC
G M S P D D V N T L S S L G Y C G V M D
GGAAAGATGCCGGACGGTAAACCGATGGGATTCCCCTTCGACCGACGTGTGCCCTCGGAA
G K M P D G K P M G F P F D R R V P S E
GAGACCTTTTACACACCCAACATCAAGACCATTGACCTCACCATCAAGAACGTCAAGCAG
E T F Y T P N I K T I D L T I K N V K Q
AAACCTACCGAGTAGATTAGTAAGCGCCTGCCTTGTGCAGTTATGATGTAATACGTACAG
K P T E ***
TAGGGAAACTTCATCTGTTTCCTTTGTCTGTATAGTCCTGACAGTGTTAGCATTTTATACA
ACGTTTGGATCAATAATGAAGGATATTCTTTCTTTCTATCTTTATAACAGATAATGGAAC
TAATTTTAATGTGGAAGATGAACTGCATTAAATATATAACTTAAAAAAAAAAAAAAAAAAAA
AAAAA [2334 bp]

```

Fig. S15. cDNA and deduced amino acid sequence of *P. americana* hemocyanin subunit 2 (PamHc2, FM242649). Copper-binding histidines are shaded in black, the putative signal peptide is in boldface, and the potential N-glycosylation sites (NXS/T) are boldface and underlined; in the cDNA, the putative polyadenylation signal is underlined.

ATGGACTTCCCGTTTTGGTGGAACAGGAGGAGTACGGCGTGCACAAGGACCGCAAAGGC
 M D F P F W W K Q E E Y G V H K D R K G
 GAGCTCTTCTACTACATGCACCACCAGCTGATCGCACGCTTCGACGCCGAGCGCCTGTCC
 E L F Y Y M H Q L I A R F D A E R L S
 AACGACCTGCCCATCGTGGAACCGCTCTACTGGGACGAGAAGATCGTAGACGGCTTCTAC
 N D L P I V E P L Y W D E K I V D G F Y
 CCTCAGACAACATACCGTGTGGGAGGGGAGTTCCCGGCACGTCCCGACAACCTTTGAATTC
 P Q T T Y R V G G E F P A R P D N F E F
 CATGACTTGGAACACATCAAAGTCAAGGATATGATGGATTACACGCGCAGGATCCGAGAG
 H D L E H I K V K D M M D Y T R R I R E
 GCCATCTCTAACAGGCCGTCGTCACGAGGACGGGCGAGATGGTGTCCCTAAACAACACG
 A I S Q Q A V V T R T G E M V S L **N N T**
 CAGGGCATCGACATTCTCGGCTGCATGGTGGAACCTTCACACGACTCCAAGCACCCCGAG
 Q G I D I L G C M V E P S H D S K H P E
 TACTATGGAGCTTTGCACAACTACGGCCACATCATGCTGGGCCAGATTACCGATCCCAAG
 Y Y G A L N Y G I M L G Q I T D P K
 GGGAAAGTTCAACATGCCTCCTGGAGTAATGGAACACTTCGAAACCGCCAC **[530 bp]**
 G K F N M P P G V M E H F E T A

Fig. S16. Partial cDNA and deduced amino acid sequence of *S. lateral* hemocyanin subunit 1 (SlaHc1, FM242652). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined.

ATGGACTTCCCGTTTTGGTGGAACAGGAGTATGCGGGCTACAAAGATCGCAAGGGAGAA
 M D F P F W W K Q E Y A G Y K D R K G E
 CTGTTCTTCTACATGCATCATCAAATGGTTGCCAGGTTTCGATGCCGAGCGACTATCTACC
 L F F Y M H Q M V A R F D A E R L S T
 GATTTGCCAGTAGTAGAACATTAGAATGGGACCAAAAGATCGTGGAAGGTTTTGCACCC
 D L P V V E P L E W D Q K I V E G F A P
 GGCGCCACCTATGAAAACGGCCAGGAGTTCCCATGAGGCCGGATAACATGCGCTTCAGC
 G A T Y E N G Q E F P M R P D N M R F S
 GATCTTCCCTTCCTCACAGTTCAAGACATGAAAATATTCGAGGGTCGCATCCGCGACGCC
 D L P F L T V Q D M K I F E G R I R D A
 ATCGCTTCTGGATATGTGAAGACGGCTGACTCGCGACTGATGAGTCTGAATAACACGGAT
 I A S G Y V K T A D S R L M S L **N N T** D
 GGCATAAACCTTCTGGGCGAGATTGTGGAGTCCTCGGCACACAGTGTCAACCCTGTGTTT
 G I N L L G E I V E S S A H S V N P V F
 TACGGCAAGTTGCACAACGACGGTCACGTGATGCTGTCCAAGGTGACGGACCCACACCTT
 Y G K L N D G V M L S K V T D P H L
 CGCTACGGCGTTCCTCCTGGCGTAATGGAACATTTTGAGACAGCGAC **[527 bp]**
 R Y G V P P G V M E H F E T A

Fig. S17. Partial cDNA and deduced amino acid sequence of *S. lateral* hemocyanin subunit 2 (SlaHc2, FM242653). Copper-binding histidines are shaded in black and the potential N-glycosylation sites (NXS/T) are boldface and underlined.

EpiHc : M-----*-----20-----*-----40-----*-----60-----*-----80-----*-----100-----*-----
LpoHcI : M-----TVKEKQSRLLPLFKHLT-----SLTRD-QLPVG-LR-----DORLKDVG-----HLPRGTLFSCF : 49
TtrHcA : M-----TIKEKQASLTALFEHL-----SVPKQ-HIPEK-ER-----DNRLHDVG-----HLSRGKLFSCF : 48
EcaHcE : M-----PDKQKQLRVLSLFEHMT-----SIN-T-PLPRD-QI-----DARLHLHG-----RLPQGEFSCF : 48
NinHcE : M-----SVKEKQQRVCSLFTHLT-----SISR-APLA-DR-----DPRLHGIG-----LQAQGEFSCF : 49
EcaHcG : M-----ASIPEKQALILPLFEKLT-----SLTKE-TPPRA-QW-----DPRLAGVG-----VLPGRGTLFSCF : 50
NinHcG : M-----SIKEKQDQILPLFEKLT-----TLTRQ-QIPPE-QR-----DPRLIGVG-----VLPGRGTLFSCF : 49
CsaHc1 : M-----ATKDKVEKLPLFEKLT-----SLTKQ-QLAPD-QR-----DPRLASVG-----ILPRGTLFSCF : 49
AauHc6 : M-----TVADKQARLMPLFKHLT-----ALTRE-KLPLD-QR-----DERLKGVG-----ILPRGTLFSCF : 48
LpoHcIV : M-----TLKEKQDRILVLFLEHT-----SLTKH-QLPED-QR-----DPRLHDLG-----HLHRGELFSCF : 49
EcaHcF : M-----TVQDKQRQILPLFEHLT-----SLTRG-GLDRT-ES-----DVRLRRVG-----RLPRGTLFSCF : 49
NinHcF : M-----PAQDKQRIRLPLFEHLA-----SLTRA-VLPPE-ER-----DPRLKRLG-----RLPRGTLFSCF : 49
EcaHcD : M-----TIADHQARILPLFKKLT-----SLSPD-PLPEA-ER-----DPRLKGVG-----FLPRGTLFSCF : 49
NinHcD : M-----SVKDKQAQILPLFKNLT-----ALAPE-TLPEA-ER-----DPRLKGVG-----FLPRGRFLFSCF : 49
LpoHcII : M-----TLHDKQIRVCHLFEQLS-----SATVI-GNGDKHKK-----SDRLKNVG-----KLQPGAIFSCF : 50
EcaHcA : M-----TILHDKQVQALKFEKLS-----VAATGE-PVPAD-QI-----DERLKNIT-----TLGPNEFFSCF : 51
NinHcA : M-----SLLHEKQVRVLLKLFERLS-----VAASGE-KIPAD-QI-----DARLSTVG-----ELPNSAFFSCF : 51
LpoHcVI : M-----VLGALEKQLRVLPFEYAS-----IPTKE-KFALKAHR-----DPKLQGLG-----VLGRGQLFSCF : 52
EcaHcB : M-----PSTAEEKQRRLPFQFQTS-----LSTKD-KFGLVQR-----DPRLAGLG-----VLGRGVLFSCF : 51
NinHcB : M-----SSASEKQVRVLPFQYVS-----FSTKD-KFGRMVRQ-----DPRLSGLG-----VLGRGVLFSCF : 51
EcaHcC : M-----PSDANEMQARLLQFHEST-----LSTKA-KFGLRVIR-----DPKLAGIG-----TLGRGKIFSCF : 52
ScoHcA : M-----WS-LALATLFLVGTGIRADQC-----PPVPADTKDKLEKLTLEIGHVNRPLTPETE-----PAG-YD-----ETLKGKLG-----TLPGHEIFSCF : 76
ScoHcB : M-----VA-KWCVLAMCLLVAVG-----ADKC-----PRA-TDQEAQKRMLEVLQHVNPYVAETKDPKIPAG-----S-----EDVFKHLG-----LLEKHEVFSF : 77
ScoHcC : M-----GAMKVMVFFAIALVAVKAYDEEAKC-----MSHEDSQVKERQILHIVDSINKPSPDFRA-----PRGVID-----EHLRGLG-----TLKKREIFSCF : 83
ScoHcD : M-----DT-RVLRLTLALVALSGVLADESCE-----SAIVNDYKLGKQEIQLHVDITINKPYPDFKD-----TRGIIID-----ESKLKGLG-----TLPRREVFSF : 82
SpiHc1 : M-----KLFLAVLLATTAT-----AEKCP-----KP-VNVEAKQKMVDITLSYINKPANVPV-----PKPID-----PKKLGIG-----TLPKDQVFSF : 73
GpuHc1 : M-----KLLLLCL-AATSAFAWPSV-----NDAAPSLAVRQAVVNHLLRYRTEPL-----RFDDLKAAAASFN-----PVADKSIYTDGGVA-----AQHLVDEIADHRVLEKHHWFSF : 95
GroHc1 : M-----KLLLLCL-AATSRFAWPSV-----NDAAPSLAVRQAVVNHLLRYRTEPL-----RFDDLKAAAASFN-----PVADKSIYTDGGVA-----AQHLVDELADHRVLEKHHWFSF : 95
CscHc : M-----RTLAFCL-LVASAAWPHFQFN-----DHPTDSLTRQHQDINQILTRITEDL-----TGDAKALAESFD-----PIADTSIYKDDGKA-----AKKLVGEIKDERVLERHHWFSF : 96
EpuHc1 : I-----VLLLVSTAAAWPNWA-----NDAPVSSAKKQDILRLLYKVTHEHTRERFTDLKDAKSNFD-----PLADLTHYTDGGAS-----VKTLMQELNDHRVLEQHHWFSF : 93
PleHc2 : M-----MVLVKLLVCLCS-LVAAAAWPSF-----GGDGFQSDAP-----DATLAKRQVQVNHLLDRIYDHL-----HYTDLKGIAGTFS-----PEADTSIYTDGGAA-----AHVLMELRDGRVLEQHHWFSF : 104
HamHcA : M-----LVLVLF-LVAAAAWPSFYQN-----DASGDSTAKRQDQVNHLLDRIYDHL-----HYSDLKQISETF-----PEADTSMYTDGGTA-----AHVLMELNDHRVLEQHHWFSF : 91
PinHcB : M-----LVLVLF-LVAAAAWPSFYQN-----DALGTGNANKQDQVNHLLDRIYDHL-----KYPDLKQIAENFN-----PLGDTSIYNDHGA-----AEALMKELNDHRVLEQHHWFSF : 75
PinHcB : M-----LVLVLF-LVAAAAWPSFYQN-----DALGTGNANKQDQVNHLLDRIYDHL-----KYPDLKQIAENFN-----PLGDTSIYNDHGA-----VETLMKELNDHRVLEQHHWFSF : 75
PeHc1 : M-----KVLVLC-LVATAAAMPFSGELTNGFQS-----DVHSSDNAHKQDQVNHLLDRIYDHL-----KDEELHNIQTFD-----PVADTSIYDVGAA-----AKTLMQELNDHRVLEQHHWFSF : 103
PeHc4 : M-----KVLVLC-LVATAAAMPFSGELTNGFQS-----DVHSSDNAHKQDQVNHLLDRIYDHL-----KDEELHNIQTFD-----PVADTSIYDVGAA-----AKTLMQELNDHRVLEQHHWFSF : 104
PeHc3 : M-----KVLVLC-LVATAAAMPFSGELTNGFQS-----DVHSSDNAHKQDQVNHLLDRIYDHL-----KDEELHNIQTFD-----PVADTSIYDVGAA-----AKTLMQELNDHRVLEQHHWFSF : 103
PeHc2 : M-----KVLVLC-LVATAAAMPFSGELTNGFQS-----DVHSSDNAHKQDQVNHLLDRIYDHL-----KDEELHNIQTFD-----PVADTSIYDVGAA-----AKTLMQELNDHRVLEQHHWFSF : 103
CmaHc6 : M-----KLLVLF-LVAAAAWPSFGIMAD-----SAGGAFDAQQHDVNSALWKVYEDT-----QDPHLIQLSQFSD-----PLS-----GHYDDGVA-----AKRLMKELNDHRVLEQHHWFSF : 97
CmaHc4 : M-----KLLVLF-LVAAAAWPSFGIMAD-----SAGGAFDAQQHDVNSALWKVYEDT-----QDPHLIQLSQFSD-----PLS-----GHYDDGVA-----AKRLMKELNDHRVLEQHHWFSF : 96
CmaHc5 : M-----RLVLVLC-LVAAAAWPSFGIMAD-----SPGASDVQKQHDVNSALWKVYEDT-----RDSNLIQLSQFSD-----PLS-----GHYDDGVA-----AKRLMKELNDHRVLEQHHWFSF : 97
CsaHc : M-----KLLVLC-LVAGASAWPNFGMMAD-----SPGASNAQKQHDVNSALWKVYEDT-----RDPHLKELSQFSD-----PLS-----GHYDDGVA-----AKRLMKELNDHRVLEQHHWFSF : 97
CmaHc3 : M-----RLVLVLC-LVAAAAWPSFGIMAD-----CDHHSVLAHKQHDVNSALWKVYEDT-----RDPHLKELSQFSD-----PLS-----GHYDDGVA-----AKRLMKELNDHRVLEQHHWFSF : 92
PvaHc1 : L-----LALVLA-LVAAAAWPNFGFQS-----DAGESDAQQHDVNSALWKVYEDT-----RYSDLKAKASFD-----PAGRFGSYDGGGA-----VQKLVREVKGKGLVLEQHHWFSF : 93
PvaHc : M-----RVLVLC-LVAAAA-----FQVASADVQKQHDVNSALWKVYEDT-----QDGLLATANSFD-----PVGNLSYDGGGA-----VQKLVQDLNKGKGLVLEQHHWFSF : 89
PinHcC : A-----DCQAGDSADKLQKQHDVNSALWKVYEDT-----RDDLKELGETFN-----PQCDLLLYHDNGAS-----VNTLMADFQDGRVLEQHHWFSF : 81
PleHc : M-----QWTVLVAALLVAT-----VSADTVAHQQVIAKINRLYKSHIKSFTDLKAAETWN-----PRDHTKSDSGGA-----IKHLMDELADHRVLEQHHWFSF : 91
CmaHc1 : M-----KWAVVAALLVLA-----VAEGADLAHRQAVNRLLYKTIYSPISAFALKKEHATTFN-----PRDHISHCNDGNS-----VNHLMDEIAHRVLEQHHWFSF : 91
CmaHc2 : M-----KWAVVAALLVLA-----VAEGADLAHRQAVNRLLYKTIYSPISAFALKKEHATTFN-----PRDHISHCNDGNS-----VNHLMDEIAHRVLEQHHWFSF : 91
HamHcP1 : M-----SLVAAAAASPYSGSHDFSQFQDEPDGVPYTAQKQHDINFLHLKLEPL-----HEANLKALEDSD-----PLHTANMPPDAGVA-----VNMKMEQVKTQHLERHHWFSF : 99
HamHc2 : V-----LLCSLVAATAAMPYFG-----GFORDEPDGVPYTAQKQHDINFLHLKLEPL-----HEANLKALEDSD-----PLAHTANMPPDAGVA-----VNMKMEQVKTQHLERHHWFSF : 97
CmaC1 : D-----EPDGVSQYKQHDVNSALWKVYEDT-----RESNLVEKAATFN-----PVGDVSMYKDDGGA-----VQNLMYELTHQRLLEKHHAVAT : 76
CmaC2 : M-----KVLVLF-LVAAAAWPSFGIMAD-----EPDGVSQYKQHDVNSALWKVYEDT-----RESNLVEKAATFN-----PVGDVSMYKDDGGA-----VQNLMYELTHQRLLEKHHAVAT : 95
PmaHc2 : M-----KFALVLIGLCAALAA-----DTVDDTFYKHQTQVHLVHEIYEPISSHYEDLTIYKTYD-----PLQHLNDFKD-----GDH-----VKHFVHDVTGKGLVLAHEIFNIF : 92
PgrHc2 : M-----KLALVLIGLCAALAA-----DAVDDTFYKHQTQVHLVHEIYEPISSHYEDLTIYKTYD-----PLQHLNDFKD-----GEQ-----VKHFVHDVTGKGLVLAHEIFNIF : 92
CseHc2 : M-----RPATVLLTVLGACALASAE-----QKADAIKRLQQLVNLRLHNLQPLD-----PELLVFNKSF-----PLNMLNKFKD-----PTA-----VKNFLKLVTEGKLLERDAIFSL : 94
PamHc2 : M-----KLWMLATLALGACALVAD-----QKADMTFLTRQLQVNLRLHNLQPLD-----PELLVFNKSF-----PLNMLNKFKE-----PET-----VKLLVGLTIGKGLVLRDQTVFSF : 94
BduHc2 : M-----KLLGILVSCACVAAAVAVQ-----QKADGVFLNRQISVLKLLQHLNQLPVD-----SDLTAYKESF-----PLNMLNRYRD-----AEV-----VKQVGLVSGKGLVLRDQTVFSF : 95
HmeHc2 : M-----MADATFLKRLQVNLRLHNLQPLD-----PELHAKKQSGD-----PANHLNKFND-----PKA-----VRHFVNLVRKGLVLRDQTVFSF : 73
TdoHc2 : M-----MADETFLKRLQVNLRLHNLQPLD-----PELVQIKDSYS-----PLNLSVYKQ-----PQL-----VQKVVRLKLEQMLNRLNIFVLF : 73
SamHc1 : M-----RLMVLAVLVAVALGGAQAGID-----VPADKPFMLKQREVLRLLNKVHEPNRY-----KDQAEVLNAYE-----PLNLSLPRYR-----PAP-----VKQLVRLYRAGSLVLRGAIFTFL : 98
BduHc1 : M-----KTL-LFLLGVALYHQAACARAQINRG-----VPVDRFELTRQNEILHLNLKVKHEPNRF-----KDQSELGRSYD-----PTAHFQKYKN-----ASP-----VKSLLVLYTSQGLVLRGAIFTFL : 101
PamHc1 : M-----KWQL-VTFGLVALYHQAACARAQINRG-----VPADKEFLVRQREVLRLLNLKVKHEPNRF-----KDQSELGRSYD-----PTAHFQKYKN-----AVP-----VKRMVLYTSQGLVLRGAIFTFL : 97
CseHc1 : M-----KCLITVALLGVAVCHLAQAKVPR-----GE-----APADKIFLKSQREVLRLLNLKVKHEPNRF-----KDQADLGRSYD-----PTQHLKSKYKN-----ALP-----VKTLLVQYTSQGLVLRGAIFTFL : 100
HmeHc1 : M-----GLVKYTG-----ASADKFLVQRQVNLRLHNLQPLD-----PELHAKKQSGD-----PANHLNKFND-----PKA-----VRHFVNLVRKGLVLRDQTVFSF : 83
CmoHc1 : M-----QWVI-AAVLLAGLQGLVAAKYTR-----GE-----VPADKDFLLKQREVLRLLNLKVKHEPNRY-----KDQADSGSKYE-----PSNNLNRYKN-----PVP-----VKTLLVRYNSHLLVLRGAIFTFL : 99
CacHc1 : M-----KLLPACLVAVCIAYCGVS-----VPADSGLVVRQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 95
TdoHc1 : M-----RAVL-LLLLGVGTAHLVVRGVR-----GA-----VPADKEFLKQREVLRLLNLKVKHEPNRF-----KDQSELGRSYD-----PLNMLNRYRD-----AEV-----VKQVGLVSGKGLVLRDQTVFSF : 99
PmaHc1 : M-----KWL-LTGLVNLVLAQAKILTR-----GS-----VPADQDFLTRQNEILHLNLKVKHEPNRY-----QEQDVLVKDYD-----PSVA-----GKFKD-----VTP-----IKRLMKYNNKGLVLRGAIFTFL : 98
PgrHc1 : M-----KWL-LTGLVNLVLAQAKILTR-----GS-----VPADHEFLVRQREVLRLLNLKVKHEPNRY-----QEQDVLVKDYD-----PSVA-----GKFKD-----VTP-----IKRLMKYNNKGLVLRGAIFTFL : 98
ScuHc1 : M-----ARFALTIACTLVAVAA-----TPADQEFLLTKQREVLRLLNLKVKHEPNRY-----QEQDVLVKDYD-----PSVA-----GKFKD-----VTP-----IKRLMKYNNKGLVLRGAIFTFL : 93
MgeHc1 : M-----KTVFCLLLALAASLAPVTKKD-----VPADHEFLVRQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 96
TdoHex1 : M-----AQTLVLITLTVAVFAATNIHYNRRYDDRVADNHLNKLQVITLKLHRRVHNPQY-----RDLADIANNYN-----IEQNVDKYQN-----EAA-----VRKFVYQKGLVLRGAIFTFL : 104
PmaHex1 : M-----KSAFVLLALAFAAASV-----PGDTHDKRVADKEFLVKQKNNLQFVRPHQYNLY-----KEQAEIGKYS-----LESHLADYTN-----AKA-----VKQFLVHYKQGLVLRGAIFTFL : 100
BdHex : M-----KTALVLI-LATLAVAPSSQDQY-----VADKQFELTRQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 99
RfHex2 : M-----RTAVLLVLAALAAANSPSQESR-----IADKQFLQRQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 100
PamHex12 : M-----KTALFAAVVA-FVAARFPDHQYK-----QLADKQFLAQRQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 98
RfHex1 : M-----NTALLFATVAVVLVCGAFSDHHVKG-----KVADKPFMLKQREVLRLLNLKVKHEPNRY-----KDQAEIGNSFYD-----LSTALTGFKN-----PNA-----AKALLKGYKAGSLVLRGAIFTFL : 99
LmiHex1 : M-----RTATVVLVLLAALAAATAVPP-----SEADKELLEKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIENIDNYQ-----KDK-----VEVFWKYTYEYGVFPRGEVFSF : 97
LmiHex2 : M-----RTATVVLVLLAALAAATAVPP-----SEADKELLEKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KDK-----VELFWKYTYEYGVFPRGEVFSF : 97
RmiHx1 : M-----RTATVVLVLLAALAAATAVPP-----SEADKELLEKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KEK-----VELFWKYTYEYGVFPRGEVFSF : 97
TmoHex2 : M-----RLTLVALAGLCAVSAVS-----HSTNYAQKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KEK-----VELFWKYTYEYGVFPRGEVFSF : 92
AgeHex : M-----KFSLLVGLCALASAVLHGHK-----VKVQDQEVYKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KEK-----VELFWKYTYEYGVFPRGEVFSF : 97
LdeDp1 : M-----KISLVFLVGFCITAVLSNPV-----ADTNYLKREQQILKLHYHNPQSTY-----PEHVEIGKSCMFCCKGSLYDKYK-----REV-----VDNVWFQWRFG-----LLPRGEVFSF : 94
AgeHex2 : M-----KFAEKDWLHKQEDLLLFRHLYEKOWN-----PQLVAYAKNFH-----FKESAEFLQHEPTVQWAKKLWSYEHG-----LLPKEGEVFSF : 76
MseAryla : M-----KTVVILAGLVALALSSAVPPPKYQHHTYK-----SPVDAIFVEKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KEK-----VELFWKYTYEYGVFPRGEVFSF : 103
MseArylb : M-----KTVVILAGLVALALGSEVPVKHSFKV-----KDVDAAFVERQKQKILRLFYHQVQPTIT-----PEEQEIAKSYK-----PIEHVDNYQ-----KEK-----VELFWKYTYEYGVFPRGEVFSF : 100
CfeHex2 : M-----LKELLALLVALCTTGLHITG-----KTADLFLHKKQILYELLFVKQNT-----LTMDFHKGIDRYN-----IESNIDSYND-----KLI-----VQDFLHLHKKG-----MLSRNAIYSF : 97
BheHex : M-----LKKVFLASLAIIVIKADDFEYTD-----VIADQFLKQKQVFLQYLVHSQPD-----ISNPELFQEGLAYD-----IGANIDSYND-----KDA-----VNHFEELYKFG-----FLPRGAIYSF : 100

EpiHc : IKRHREESTKVFEFLDSDKYKGFLENAASLRPVVNEILFYVCSVALYQSPHTTLLNLAAPPIMQTPYDFF-DQRVENQIN-GIIES---TK-----GAPDKQIQVILEPTGN : 158
LpoHcI : HAKHLEEAHTYELLYGAKDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
TtrHcA : HREHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
EcaHcE : HREHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
NinHcE : HEKGLAEATKLYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
EcaHcG : HEKHLLEATKLYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
NinHcG : HEKHLLEATKLYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
CsaHc1 : HEKHLLEATKLYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
AauHc6 : HARHLEAEATLYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
LpoHcIV : HKEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 150
EcaHcF : HSEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 151
NinHcF : HTEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 150
EcaHcD : HSEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 150
NinHcD : HEDHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 149
LpoHcII : HPDHEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 153
EcaHcA : YPDHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 154
NinHcA : LPAHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 154
LpoHcVI : HAEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 154
EcaHcB : HEHLEEAHTYELLYHAAKNFDFDOLMLLEQARTVNEGFMFYAAVAVLHRDDCGVTPVPIIEQVPPDRFIPVETIRHADKEA-AAH-P-DK-----DIVFEETGN : 152

NinHcB : HEDHLKEATQLYEVLTITAPTFEEFDLCHQCQRDYVNEGLFAYAVSAILHRKDCRGNVLPPVQEIFPDKFVPVETLYEAYKET - KLHKEDE ----- DVIINIQTGKN : 152
EcaHcC : HEDHLEASRLAEVLGAETDFEEDLCHQCQRDVNEALFVYSLVAILHRPDCGHSLPPIQEIFPDKFVPVETIYKAFKEA - TRHADKTD ----- DIIVDMEATGT : 154
ScoHcA : DERTWPEATKAAEFLEATDFEHFTRQADVLRRHINEDMFMYALNVAVLHRKDRTRGVQVPRHIIYDPDKFLQDILVEVREKV - NHGECK ----- PLVDATLEHQN : 176
ScoHcB : DERQWDEATTAAYVMTAPSFDEFIDRAEIVRHRINEDMFYAFSAIAHRRDDTRGILNPRHIEIYDPDKFLKHVIVEVKNSI - NSGQED ----- PLIDATHEFTD : 177
ScoHcC : DERNWDEASKVRLLLDPAKDFDDFIDAEVIRLVRNEELFLYAFSAVAMHRGDTQGLQVPRHIDIFPDKFLKEDVIHRLLELS - NRGEHYDRI ----- PIDATQTSNH : 186
ScoHcD : DERNWAEAAKVVELLLEPKTFREFTHLADIHHRVNDLFLYALSVATAHRRPDCQGVQVPRVLDIYDPDKFLRKEVHIKIKEVS - NEGAYLDKV ----- PVDATEVSDN : 185
SpiHc1 : DERTWDEVSEVLKRFMAAPTFEEDFKVAETLYPHVNEDLFLGLSVAIVHRPDANGVHVPRVHEVDPKFLTNDVLINIRKQI - IKGHKN ----- PVVKDTIYYHN : 173
GpuHc1 : NTRQREALLFLFDVLTHSKTWTALNNAAYFREKMNNEGFEVYALYAAVHISKLGAIVLPPLEYVTPHFLTNSVEIQKAYTAQ - M ----- TQTPTGTFKQDTFS : 193
GroHc1 : NTRQREALLMFDVFMHCKTWEAALNNAAYFREQMNNEGFEVYALYAAVHISKLGAIVLPPLEYVTPHFLTNSVEIQKAYTAQ - M ----- TQTPTGNFKMDFTS : 193
CcsHc : NTRQREALLMFDVFMCKTWEAAMNNAAYFRHSMNEGFEVYALYAAVHISLHGAGIVLPPLEYVTPHMTNSEVISKAYTAQ - M ----- TQTPTGKFSMEFTGS : 194
EpuHc1 : NNRHREALLMFDALMCHNDWHTAVSAAAYFREKLNNEGFEVYALYAVIHADLGGGLVLPPLYEIVTPHFLTNSVEINKAYSAK - M ----- TQTPTGKFRMSFTGT : 191
PleHc2 : NTRQREAIMLFEVLTHCKSWCVFKNAAAYFREHMNEGFEVYALYVAVIHTDLGHGIVLPPLEYVTPHFLTNSVEINKAYSAK - M ----- TQTPTGKFNMDFGT : 202
HamHcA : NTRQREALLMFDVLHCKSWCECFVNDAAAFRRERMNEGFEVYALYAVIHSELGGGIVLPPLEYIPTHMFTNSEIHKAYTAK - M ----- TQTPTGRFEMKFTGT : 189
PinHcB : NTRQREALLMFAVLNCKQKEWYCFRNSAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYIPTHMFTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 173
PinHcA : NTRQREALLMFAVLNCKQKEWYCFRNSAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYIPTHMFTNSEVIDKAYSAK - M ----- TQKQGTFNVSFTGT : 173
PleHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- THKEGTFNMSTGT : 201
PleHc4 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- THKEGTFNMSTGT : 202
PleHc3 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- THKEGTFNMSTGT : 201
PleHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- THKEGTFNMSTGT : 201
CmaHc6 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQTAAKIESFTGS : 195
CmaHc4 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQTAAKIESFTGS : 194
CmaHc5 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQTAAKIESFTGS : 195
CsaHc : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQTAAKIESFTGS : 195
CmaHc3 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQTRAKIESFTGS : 190
PvaHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 191
PvaHc : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 187
PinHcC : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 179
PleHc : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 189
CmaHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 189
CmaHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 189
HamHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 197
HamHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 195
CmaC1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 175
CmaC2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 194
CmaC3 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 190
PgrHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 190
CseHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 192
PamHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 192
BduHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 193
HmeHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 171
TdoHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 171
SamHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 196
BduHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 199
PamHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 195
CseHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 198
HmeHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 181
CmoHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 197
CacHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 193
TdoHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 197
PmaHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 196
PgrHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 196
ScuHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 191
MgeHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 194
TdoHex1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 202
PmaHex1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 211
BdiHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 201
RfLHex2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 202
PamHex12 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 200
RfLHex1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 201
LmiHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 201
LmiHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 201
RmiHx1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 201
TmoHex2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 198
AgeHc1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 206
LdeDp1 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 204
AgaHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 184
MseAry1a : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 218
MseAry1b : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 215
CfeHc2 : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 205
BheHex : NTRQREALLMFTVLNCKQKWNCFNLNNAAYFRERMNEGFEVYALYVSVIHSKLDGGIVLPPLEYVTPHMTNSEVIDKAYSAK - M ----- TQKPGTFNVVSFTGT : 203

240 * 260 * 280 * 300 * 320 * 340
EpiHc : --NR--DPEKLSYFREDIAGNAHHMHVVPATW--GSPNDKK--INRKGELFYMHQMCARYDCERLNSGMRRVKSLSHDYDE-IIPEGYYPRLSKISITYESAQRANT-- : 261
LpoHcI : --TL--DPEYRLAYFREDIAGNAHHMHVVPATW--RPEFMHKE--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNDFHEEL-EG-YSPLHTSLISGLNYASRPDGL-- : 253
TtrHcA : --TL--DPEYRLCYFREDIAGNAHHMHVVPATW--RSEVMGKT--KNRKGELFFYMHQMCARYDCERLNSGMHRMSPFNSHEEL-EG-YAPHLTSLISGLHYGRSPDGL-- : 251
EcaHcE : --TL--DEEYKLYFREDIAGNAHHMHVVPATW--DPAFMGRM--KDRKGELFYMHQMCARYDCERLNSGMHRMIPFNSFDEKL-EG-YSAHLTSLISGLHYASRPDGL-- : 252
NinHcE : --TL--DPEYKMSYFREDIAGNAHHMHVVPATW--RPEVMCKL--KDRKGELFYMHQMCARYDCERLNSGMHRMIPFNFEEEL-EG-YSAHLTSLISGLHYASRPDGL-- : 253
EcaHcC : --TL--DPEYKLYFREDIAGNAHHMHVVPANW--DAVFTGKT--KDRKGELFYMHQMCARYDCERLNSGLTRMIPFNFKEKL-EG-YAPHLTSLISGLHYASRPDGL-- : 254
NinHcG : --TL--EPEYKLYFREDIAGNAHHMHVVPANW--SVELTGKV--KDRKGELFYMHQMCARYDCERLNSGLNRMVAFHNFEEKL-EG-YAPHLTSLISGLHYASRPDGL-- : 253
CsaHc1 : --TR--DPEYKLYFREDIAGNAHHMHVVPANW--DESLTGKV--KDRKGELFYMHQMCARYDCERLNSGLNRMVAFHNFEEKL-EG-YAPHLTSLISGLHYASRPDGL-- : 253
AauHc6 : --TL--DPEYKLYFREDIAGNAHHMHVVPATW--NPTVMGKE--KDRKGELFFYMHQMCARYDCERLNSGLQRMIPFQNFDEPL-EG-YAPHLTSLISGLHYASRPDGL-- : 253
LpoHcIV : --TL--NPEQRKLYFREDIAGNAHHMHVVPATW--RPEVMGKV--KDRKGELFYMHQMCARYDCERLNSGLQRMIPFQNFDEPL-EG-YSAHLTSLISGLHYASRPDGL-- : 254
LpoHcV : --TL--DPEHKLAYFREDIAGNAHHMHVVPATW--DASVMGKV--KDRKGELFYMHQMCARYDCERLNSGLQRMIPFNFDEKL-EG-YSPLHTSLISGLHYASRPDGL-- : 255
NinHcF : --TL--DPEYKLYFREDIAGNAHHMHVVPATW--RPEIMGKV--KDRKGELFYMHQMCARYDCERLNSGLQRMIPFNFNEKL-EG-YSPLHTSLISGLHYASRPDGL-- : 254
EcaHcD : --TR--DPEYKLYFREDIAGNAHHMHVVPATW--DAIFGKV--KDRKGELFYMHQMCARYDCERLNSGLQRMIPFQNFDEKL-EG-YSPLHTSLISGLHYASRPDGL-- : 254
NinHcD : --II--DPEYKLYFREDIAGNAHHMHVVPATW--RPDFFGKV--KDRKGELFYMHQMCARYDCERLNSGLQRMIPFQNFDDKL-EG-YSAHLTSLISGLHYASRPDGL-- : 253
LpoHcII : --TL--DPEYKLYFREDIAGNAHHMHVVPSTW--NPKYFGKK--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNDFDEPL-AG-YAPHLTSLISGLHYASRPDGL-- : 257
EcaHcA : --TL--DPEYKLYFREDIAGNAHHMHVVPSTY--DPAFFGKV--KDRKGELFYMHQMCARYDCERLNSGLNRMIPFNFNEPL-EG-YAHLTHVAGSRHYAQRDGL-- : 258
NinHcA : --TL--DPEYKLYFREDIAGNAHHMHVVPSTY--DSKFFGKV--KDRKGELFYMHQMCARYDCERLNSGLTRMVPFNFEEPL-EG-YAHLTHVAGSRHYAQRDGL-- : 258
LpoHcVI : --TL--DPEYKLYFREDIAGNAHHMHVVPATW--RPEVIGKV--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YSSHLSSSINGLPYSPRDPGL-- : 258
EcaHcB : --VK--DPEYKLYFREDIAGNAHHMHVVPATW--RPEVHRI--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YSSHLSSSINGLPYSPRDPGL-- : 256
NinHcB : --TM--DPEYKLYFREDIAGNAHHMHVVPATW--RPEVIGRI--KDRKGELFFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAHLSSSINGLPYSPRDPGL-- : 256
CmaHcC : --TM--DPEYKLYFREDIAGNAHHMHVVPATW--DSVKMHMR--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YNPHLSQNGLPYAFRDPGL-- : 258
ScoHcA : --QL--DPNYSYLFLEDIAGNAHHMHVVPATW--LPKHGPRK--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAPHLTHVAGSRHYAQRDGL-- : 280
ScoHcB : --LR--DPNYSYLFLEDIAGNAHHMHVVPATW--LPKHGPRK--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAPHLTHVAGSRHYAQRDGL-- : 285
ScoHcC : --YL--DPNYSYLFLEDIAGNAHHMHVVPATW--VSELGNEK--KDRKGELFFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAPHLTHVAGSRHYAQRDGL-- : 290
ScoHcD : --HL--DPNYSYLFLEDIAGNAHHMHVVPATW--LPKHGPRK--KDRKGELFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAPHLTHVAGSRHYAQRDGL-- : 289
SpiHc1 : --DF--DPFRRYDYFLEDIAGNAHHMHVVPATW--TQDIEGK--SKLGEFYMHQMCARYDCERLNSGMHRMFPNFHEPL-EG-YAPHLTHVAGSRHYAQRDGL-- : 276
GpuHc1 : --KK--NPEQRKLYFREDIAGNAHHMHVVPATW--EDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 295
GroHc1 : --KK--NPEQRKLYFREDIAGNAHHMHVVPATW--EDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 295
CcsHc : --QK--NPEQRKLYFREDIAGNAHHMHVVPATW--DDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 296
EpuHc1 : --KK--NPEQRKLYFREDIAGNAHHMHVVPATW--KDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 293
PleHc2 : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--KDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 304
HamHcA : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--KDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 291
PinHcB : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--KDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 275
PinHcA : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--KDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 275
PleHc1 : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--DDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 303
PleHc4 : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--DDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 304
PleHc3 : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--DDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 303
PleHc2 : --KK--NKEQRKLYFREDIAGNAHHMHVVPATW--QDSY-GYH--KDRKGELFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 303
CmaHc6 : --KS--NPEQRKLYFREDIAGNAHHMHVVPATW--DDAHENHH--IERKGESFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 298
CmaHc4 : --KS--NPEQRKLYFREDIAGNAHHMHVVPATW--DDAHENHH--IERKGESFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 297
CmaHc5 : --KS--NPEQRKLYFREDIAGNAHHMHVVPATW--DDAHENHH--IERKGESFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 298
CsaHc : --KS--NPEQRKLYFREDIAGNAHHMHVVPATW--DDSHENHH--IERKGESFFWAHQLTVRFDAERLNSGLDVLDEL-YWDR-PIEGFAPHTTY-RYGGEFFPRPDV-- : 298

CmaHc3 : -KS-NPEQRVAYFGEDIGMNTHHVTHLEFPFW--DDAHENHH--IERKGENFFWWHQLTVRFDAERLSNHLDPVDEL-HWDD-VIHEGFDPQAMY-KYGGFYPSRPDNI-- : 293
PvaHc1 : -KK-NPEQRVAYFGEDIGMNTHHVTHMEFPFW--DDKY-SHH--LDRKGNFFWWHQLTVRFDAERLSNLYDPVEEL-SWDK-PIVQGFAPHTTY-KYGGFYPSRPDNI-- : 293
PvaHc : -KK-NPEQRVAYFGEDIGLNTHHVTHWMEFPFW--NDAY-GHH--LDRKGENFFWWHQLTVRFDAERLSNLYDPVDEL-QWNK-PIVQGFAPHTTY-KYGGFYPSRPDNI-- : 289
PinHcC : -KK-NPEQHVAIFGEDVGMNTHHVTHWMEFPFW--EDSS-GRH--LDRKGSFFWWHQLTVRFDAERLSNHLDPVEEL-SWNK-AIDEGFAPHTAY-KYGGFYPSRPDNI-- : 281
PleHc : -KK-NREQRGAYFGEDVGLNSHHVHWHMDFPFW--N---GAK--IDRKGLFFWAHHQLTARYDAERLSNFLPAVDEL-YWDR-PIKDGFAPIHTY-KYGGFFPARPDNK-- : 288
CmaHc1 : -KK-NPEQRVAYFGEDIGMNSHHVHWHMDYFPFW--H---GQE--IDRKGLFFWAHHQLTARFDAERLSNHLPVDEL-YWDR-PIKDGFAPIHTY-KYGGFFPARPDNK-- : 288
CmaHc2 : -KK-NPEQRVAYFGEDIGMNSHHVHWHMDYFPFW--H---GQE--IDRKGLFFWAHHQLTARFDAERLSNHLSPVDEL-YWDR-PIKDGFAPIHTY-KYGGFFPARPDNK-- : 288
HamPhc1 : -QN-NPEHKIAYFGEDIGLSTHYINWHIEYFPFW--NETF-GYQ--IERRGENFYFWHQLVNRFEAERISNHLQKIEKL-HWER-NLHEGFDPHITS-KNGNPFPRFHDDI-- : 299
HamPhc2 : -QN-NPEHKIAYFGEDIGLSTHYINWHIEYFPFW--NETF-GYQ--IERRGENFYFWHQLVNRFEAERISNHLQKIEKL-HWER-NLHEGFDPHITS-KNGNPFPRFHDDI-- : 297
CmaCC1 : -PN-DIEHRVAYFREDIGVGTTHLMIHLENPCW--KDTY-GYH--IDRKGENFFYAHQLLNRYEAERISNHLPLEEL-KLDK-PLKQGFAPQTTY-NFGGFPATRNDNV-- : 277
CmaCC2 : -PN-DIEHRVAYFREDIGVGTTHLMIHLENPCW--KDTY-GYH--IDRKGENFFYAHQLLNRYEAERISNHLPLEEL-KLDE-PLKQGFAPQTTY-KFGAPFIRNDDI-- : 296
PmaHc2 : -IK-NPEQRVAYFGEDIGVNSHHSHWMDFPFW--KR-SYDVT--KDRRGELFFYMHQMVARFDAERLSNMLPQVEPL-NWHH-EIEEGFAPAAMY-FNGQEFPMRPDGI-- : 292
PgrHc2 : -IK-NPEQRVAYFGEDIGVNSHHSHWMDFPFW--KR-SYDIT--KDRRGELFFYMHQMVARFDAERLSNFLPQVEPL-DWHH-EIEEGFAPAAMY-FNGQEFPMRPDGM-- : 292
CseHc2 : -VN-NPEQRVAYFGEDIGMNSHHSHWMDFPFW--KQ-EYTV--KDRKGELFFYMHQMVARFDAERLSNDLPVEPL-EDWQ-QITEGFAPAATY-ENGQEFPMRPDNI-- : 294
PamHc2 : -VN-NPEQRVAYFGEDIGMNSHHSHWMDFPFW--KQ-EYAGY--KDRKGELFFYMHQMVARFDAERLSNDLPVEPL-EDWQ-KIVEGFAPGATY-ENGQEFPMRPDNI-- : 294
BduHc2 : -AN-NPEQRVAYFGEDIGLNSHHAHWHMDFPFW--KP-EYDGD--KDRKGELFFYMHQMVARFDAERLSNDLPVQPL-EDWK-KIVEGFAPGASY-ENGQEFPMRPDNI-- : 295
HmeHc2 : -VN-NPEQRVAYFGEDIGMNSHHAHWHMDFPFW--KP-HYQNHSGKDRKGELFFYMHQMVARFDAERLSNMLPVQPL-EWNR-KIVEGFAPGAMY-DNGQEFPMRPDNI-- : 275
TdoHc2 : -VR-NPEQRVAYFGEDIGLNSHHAHWHMDFPFW--KP-EYDGE--KDRKGELFFYMHQMVARFDAERLSNMLPVQPL-DWEK-PIVQGFAPGAMY-ENGQEFPMRPDNI-- : 273
SamHc1 : -TR-NPEQRVAYLGEDLGMNSHHSHWMDFPFW--KQEDYGR--KERKGELFFYMHQLIARFDERLSNDLPVEPL-YWTE-RIKDGFPQTTY-RVGEFFPARPDNI-- : 299
BduHc1 : -VR-NPEQRVAYLGEDVGMNSHHSHWMDFPFW--KQEQYGS--KDRKGELFFYMHQLIARFDERLSNDLPVEPL-YWDE-KIVDGFYPQTTY-RVGEFFPARPDNI-- : 302
PamHc1 : -TR-NPEQRVAYLGEDVGMNSHHSHWMDFPFW--KQEEYGVH--KDRKGELFFYMHQLIARFDERLSNDLPVEPL-YWDE-KIVDGFYPQTTY-RVGEFFPARPDNI-- : 298
CseHc1 : -VR-NPEQRVAYLGEDVGMNSHHSHWMDFPFW--KQDEYGVK--KDRKGELFFYMHQLIARFDERLSNDLPVEPL-YWED-KIEDGFYPQTTY-RVGEFFPARPDNI-- : 301
HmeHc1 : -TR-NPEQRVAYLGEDVGMNAHSHWMDFPFW--KQEQYGV--KDRKGELFFYMHQLIARFDERLSNDLPVEPL-SWED-KIVDGFYPQTTY-RVGEFFPARPDNI-- : 284
CmoHc1 : -TR-NPEQRVAYLGEDLGMNSHHAHWHMDFPFW--KPEEYGD--KDRKGELFFYMHQLIARFDERLSNDLPVEPL-GWRE-RIVDGFYPQTTY-RVGEFFPARPDNI-- : 309
CacHc1 : -IK-NPEQRVAYLGEDIGMNAHSHWMDFPFW--KEHEYGTH--KDRKGELFFYMHQLIARFDERLSNMLPVEPL-SWDQ-KIANGFYPTTY-RVGEFFPARPDNI-- : 296
TdoHc1 : -TR-NPEQRVAYFGEDIGMNSHHSHWMDFPFW--KP-EYDGE--KDRKGELFFYMHQLIARFDERLSNMLPVVRL-DWEK-PIVNGFYPTTY-RVGEFFPARPDNI-- : 299
PmaHc1 : -TR-NPEQRVAYFGEDVGLNSHHAHWHMDFPFW--KAAEYGE--KDRKGELFFYMHQMVARFDERLSNMLHVEPI-SFED-EIEHGFYPQTTY-RVGEFFPARPDNI-- : 299
PgrHc1 : -VR-NPEQRVAYLGEDVGLNSHHAHWHMDFPFW--KAAEYGE--KDRKGELFFYMHQMVARFDERLSNMLHVEPI-SFEE-KIEHGFYPQTTY-RVGEFFPARPDNI-- : 299
ScuHc1 : -VR-NPEQRVAYFGEDVGLNSHHAHWHMDFPFW--NEEKYGLT--KDRKGELFFYMHQMVARFDERLSNDLNEVAL-HWDK-PIVNGFYPTTY-RVGEFFPARPDNI-- : 294
MgeHc1 : -TR-NPEQRVAYFGEDIGLNTHHHYWMDFPFW--SP-HYDTK--FDRAGEMFYMHQLIARFDERLSNMLPVEPL-EWFK-PIKSGFAPLTY-RVGEFFPARPDNI-- : 296
TdoHex1 : RQLK--NPEQVAYFGEDVGLNSHHAHWHMDFPFW--KK-DYGT--KDRKGELFFYMHQMVARFDERLSNDLNEVAL-HWDK-PIVNGFYPTTY-RVGEFFPARPDNI-- : 306
PmaHex1 : YTTQ--YPEQFISYFTEDVGNFAHAYHMDYFPFW--NSKYVNVK--FDRRGELFFYMHQLIARFDERLSNMLPVEPL-EWFK-PIKSGFAPLTY-RVGEFFPARPDNI-- : 318
BdiHc1 : -VA-NPEQLYVYFEDVGLNSHHAHWHMDFPFW--NETEYGVH--FDRRGELFFYMHQMVARFDERLSNDLPVEPL-HYDR-PFQTEYYPRLY-SNGQEVPPRPPEYSR-- : 306
RfLhex1 : -VH--TPEQVLSYFTEDVGLSTYYLYYYYNYPTFF--NSTEYGVH--FDRRGELFFYMHQMVARFDERLSNDLPVEPL-HYDK-PFQTAAYPKLY-ANGQEVPPRPPEYSR-- : 307
PamHex12 : -LR--TQDHLAYFEDVGLNFAHAYHMDYFPFW--NTTLYGHN--IDRRGEFYFYYHQLIARFDERLSNDLPVEPL-YYSK-PVKSAYNPLRY-NNGEFPVRPDKFA-- : 303
RfLhex1 : -MK--SKENLLYFTEDVGLNFAHAYHMDYFPFW--NVTEYGVH--FDRRGELFFYMHQMVARFDERLSNDLPVEPL-VYNK-PLKTAYNPLRY-HTGQEMPVRPDKFA-- : 304
LmiHc1 : PVAN--NPEQVLSYFTEDVGLNSFFAYLHYKSPFW--NPANYSLP--ASKRRGDSFFFLQQLLARYLERLSNMLPDKV--DYAN-PVLGYYPPLR--QNGEAPARPEGI-- : 307
LmiHc2 : PVAS--NPEELVSYFTEDIGLNSFFAYLHYKSPFW--NPKNYSLP--EYQGRGDSFFFLQQLLARYLERLSNMLPDKV--DYAN-PVLGYYPPLR--QNGEAPARPEGI-- : 307
RmiHc1 : PVAN--NPEELVSYFTEDVGLNSFFAYLHYKSPFW--NPKNYSLP--EYQGRGDSFFFLQQLLARYLERLSNMLPDKV--DYAN-PVLGYYPPLR--QNGEAPARPEGI-- : 307
TmoHex2 : YLNL--HPEQVLSYFTEDVGLNSFFAYLHYKSPFW--SGEENFLK--YDNRGEAFYMHQMVARFDERLSNDLPVEPL-DWEV-PFETGYPPLRY-PNGLHFTPRPNYA-- : 303
AgeHc2 : YLNL--HPEQVLSYFTEDVGLNSFFAYLHYKSPFW--SGEENFLK--YDNRGEAFYMHQMVARFDERLSNDLPVEPL-DWEV-PFETGYPPLRY-PNGLHFTPRPNYA-- : 311
LdeDp1 : HMNL--GHEQVLSYFTEDVGLNSFFAYLHYKSPFW--SSEEFGLK--DSNRGLYFHAHQQLIARFDERLSNDLPVEPL-DVDS-PIQTYPNPMY--PCGAEFPPRPDKFA-- : 309
AgaHc2 : YMHT--NAEQKISYFTEDIGLNSFFAYLHYKSPFW--GHEGLFGL--KDRRGELFFYMHQMVARFDERLSNDLPVEPL-SWWT-PIATGYYPPLR--QNGEAPARPEGI-- : 289
MseAry1a : LSYP--NEEERIASYFTEDIGLNSFFAYLHYKSPFW--NSEKYGVF--KERRGEIYFYYHQLIARFDERLSNDLPVEPL-SWYS-PVKTGYPPLRY--GSYFPAQRPNY-- : 322
MseAry1b : LVYN--NEEQRLSYFTEDIGLNSFFAYLHYKSPFW--NSERYGAL--KSRRGEIYFYYHQLIARFDERLSNDLPVEPL-SWYS-PVKSGYYPPLR--SYYPFPAQRPNY-- : 319
CfeHc2 : LVPR--YLDYKLYFEDVGLNFAHAYHMDYFPFW--SMKNFDIP--AQFRGHFYFHHQLLNRYLERLSNDLPVEPL-DWNR-PFYGPHSTLY--NMGVWPQRPSY-- : 311
BheHex : YLNRAYDEMRNVYFEDIGLNTFFYFMRDNPFWL--SSEEFGLQ--KNLGRGEFLYHKLTLNRYLERLSNDLPVEPL-LWEG-EFYGYPPTMY--GNGLAYQRPGRMS-- : 310

EpiHc : * 360 * 380 * 400 * 420 * 440 * 460
LpoHcI : -RMDLTD--NTVIELERFRERIEHAISVGYIEL-PD-NT-KTPLNNE--QGIDILADLIEASN : 317
TtrHcA : -KLHDLH--VTIQDMEWRERIHDAIDLKMVDH-HH-GK-EVAVDDE--HGIDILGALVEASH : 309
EcaHcE : -KLHDIH--VTIQDMEWRERIHDAIDLKMVDH-HH-GK-EVAVDDE--HGIDILGALVEASH : 305
NinHcE : -CLHDLK--IDLKEMFRWRERILDAIDSGYIID-NE-GH-QVKLDIV--DGINVLGALIESSF : 308
NinHcC : -RLHDLK--VDVQDMIRWRERILEAIDLGVED-ES-HQ-QIKLTD--NGIDILGALIESSV : 309
EcaHcG : -CLDLSE--LEVQDVERWRERILDAIHLNVHVD-RE-NN-DVLDDE--HGADILGALIESS : 310
NinHcG : -SLRDLT--VDVQDMIRWRERILEAIDLKHVD-SK-GN-EVPLDE--NGANILGTIEASS : 309
CsaHc1 : -SLRDLV--VDVQDMIRWRERILKAIDLHFVHG-AN-GQ-TLRDDE--HGADILGALIESS : 309
AauHc6 : -SIHDLSD--VDVQDMIRWRERILDAINMHYIVD-KD-NN-KIPLDE--HGADILGALIESS : 309
LpoHcIV : -HIRDLND--VSLQEMERWRERILEGINLGHVHD-DH-GK-EVELDE--HGIDILGALIESS : 310
EcaHcF : -HLRDLVDF--VDVQDMARWRERILLSIDIGHVID-HE-GQ-EIPLDAE--NGINILGALIESS : 312
NinHcF : -SLHDLNL--VDVQDMARWRERILLEAHLGVVID-DH-GN-EIPLTE--HGIDILGALIESS : 311
EcaHcD : -HLRDLND--CSVQDMERWRERILDAIHTGLVTD-SH-GK-EIKLTE--NGINILGALIESS : 310
NinHcD : -TLRDRE--VDVQDMERWRERILSAIHTGQVID-SN-GQ-EVPLDE--RGLDILGALIESS : 309
LpoHcII : -KRLDLG--EISEVMRWRERILDSIHLGYVIS-ED-GS-HKTLDE--HGIDILGALIESS : 313
EcaHcA : -AMHDVR--VDVQDMERWRERIMEAIDLRVIS-PT-GE-YIPLDE--HGADILGALIESS : 314
NinHcA : -AMTDLRL--VDVQDMQRWRERILEAHLGKVID-SE-GN-DIILDE--HGADILGALIESS : 314
LpoHcVI : -TLQDLK--VSVQDLERWRERILDAIHTGHVID-EN-GS-HVPLDDT--RGLDILGALIESS : 314
EcaHcB : -ILKDMRE--VSVQDLERSRERILLDAINLGYVVD-PN-GL-ETPLDE--HGADILGALIESS : 312
NinHcB : -KLADLEE--CAVQDMERWRERILQINLGVVVD-PN-GH-ETVLEI--HGADILGALIESS : 312
EcaHcC : -VLCMDH--VSVQDLERWRERILDAINLGEVTD-PN-GD-EYALDE--FGIDVLGALIESS : 314
ScoHcA : -IVRDLPE--LNKMKRQWKSRIHLGIHLNVLYA-EN-GT-KISLDE--HGIDILGALIESS : 336
ScoHcB : -VVRDLPE--LTKNMRQWRERILYAVHRGEALA-AN-GS-SVSLDE--RGLDILGALIESS : 341
ScoHcC : -TLRDLPN--LPKMKVWAKRVLYSIHSGIFHF-SN-GT-DAHLDE--HGIDILGALIESS : 346
ScoHcD : -IVHDLPN--LPKTKFEWKNRMVGIKGLSLIS-AN-KT-QVPLND--HGIDILGALIESS : 345
SpiHc1 : -EIEDLPE--LTKGELMQWKNRIMEAIAHNTIL-KGGPGK-KLRTTPE--NGIDILGALIESS : 334
GpuHc1 : -HFEDVGL--VRVDMIIHETRIRDAIAHGYITA-AD-GS-HININDE--HGIDILGALIESS : 352
GroHc1 : -NFEDVGL--VRVDMIIHETRIRDAIAHGYITA-AD-GS-HININDE--HGIDILGALIESS : 352
CcsHc : -NFEDVGA--VRVDMIIHETRIRDAIAHGYITS-ED-GS-RVDIRE--NGINILGALIESS : 353
EpuHc1 : -RFEDVDGV--ARVDMIIHETRIRDAIAHGYITA-KD-GS-HIDIRNE--HGIDILGALIESS : 350
PleHc2 : -HFEDVDGV--ARVDMIIHETRIRDAIAHGYITA-HS-GN-KIDIKNE--HGIDILGALIESS : 361
HamHcA : -HFEDVDGV--AHVRDMIIHETRIRDAIAHGYITA-NH-GD-NINIRND--HGIDILGALIESS : 348
PinHcB : -HFEDVDGV--AHVHLEITESRIHDAIDHGYITA-SD-GH-TIDIRQP--KGIELLGDIESS : 332
PinHcA : -HFEDVDGV--AHVHLEITESRIHDAIDHGYITA-SD-GH-TIDIRQP--KGIELLGDIESS : 332
PeLHc1 : -RFEDVDGV--AHVHMEITENRIRDAIAHGYITA-TD-GH-TIDIRQP--NGIELLGDIESS : 360
PeLHc4 : -RFENVDGV--AHVHMEITENRIRDAIAHGYITA-TD-GH-TIDIRQP--NGIELLGDIESS : 361
PeLHc3 : -RFKNVDGV--ARVHMEITESRIHDAIAHGYITA-TD-GH-TIDIRQP--NGIELLGDIESS : 360
PeLHc2 : -QFKNVDGV--ARVDMIIHETRIRDAIAHGYITA-TD-GH-TIDIRQP--NGIELLGDIESS : 360
CmaHc6 : -HFEDVDGV--ARVDMIIHETRIRDAIAHGYVTD-GD-GS-IISISDS--HGIDILGALIESS : 355
CmaHc4 : -HFEDVDGV--ARVDMIIHETRIRDAIAHGYVTD-GD-GS-IISISDS--HGIDILGALIESS : 354
CmaHc5 : -HFEDVDGV--ARVDMIIHETRIRDAIAHGYVTD-GD-GS-IISISDS--HGIDILGALIESS : 355
CsaHc : -NFEDVDGV--ARVDMIIHETRIRDAIAHGYVTD-GD-GS-VISIRDA--HGIDILGALIESS : 355
CmaHc3 : -HFEDVDGV--ADVDMIIHETRIRDAIAHGYVTD-RN-GK-IVDLRNN--DGIDILGALIESS : 350
PvaHc1 : -DFDMDGV--ARIDRLIIHETRIRDAIAHGYIVD-KV-GN-HIDIMNE--HGIDILGALIESS : 350
PvaHc : -KFEDVDGV--ARIDRMVIESRIRDAIAHGYIVD-SE-GK-HIDISNE--KGIDILGALIESS : 346
PinHcC : -HFSDVDGV--ARVDMMSMTEDRIRDAIAHGYIDA-LD-GH-SIDIMMS--HGIEFLGDIESS : 338
PleHc : -EFEDVDGV--ARIDRLKELESIRIRDAIAHGYIIN-AD-GT-KTDINNE--HGIDILGALIESS : 345
CmaHc1 : -NFEDVDGV--ARIDRLKEMENRIRDAIAHGYADN-TD-GT-HLDINND--EGINVLGALIESS : 345
CmaHc2 : -NFEDVNGV--ARIDRMKIMESRIRDAIAHGYADN-TD-GT-HLDINND--EGINVLGALIESS : 345
HamPhc1 : -HIEDVDKV--AEVRDMIVMENRIRDAIAHGYVID-KE-GN-KVDINNE--HGIDILGALIESS : 356
HamPhc2 : -HIEDVDKV--AEVRDMIVMENRIRDAIAHGYVID-KE-GN-KVDINNE--HGIDILGALIESS : 354
CmaCC1 : -HLQDQNVN--GRIEHIVHMDRIRDAIAHGYVID-EQ-GN-KINIEEN--QGIDILGALIESS : 334
CmaCC2 : -QLHDERV--GRITHEIVHLEDRIRDAIAHGYVID-EQ-GN-KINIEEN--QGIDILGALIESS : 353
PmaHc2 : -HFHDLPW--FTVKDTEEDYEDRIRDAIAHGYVKA-SD-GH-TIFLNGT--EGINVLGALIESS : 348
PgrHc2 : -HFHDLPW--FTVKDTEEDYEDRIRDAIAHGYVKA-KD-GH-VFLNGT--EGINVLGALIESS : 348
CseHc2 : -KFCDLPS--LSVRDMKIEGRIRDAIAHGYVKT-AD-SR-LLSINNT--AGINVLGALIESS : 350
PamHc2 : -CFSDLPF--LTVQDMKIEGRIRDAIAHGYVKT-VD-SR-LLSINNT--AGINVLGALIESS : 350
BduHc2 : -QFSDLPW--LTVQDMKIEGRIRDAIAHGYVKT-AE-GK-LVSLNNT--VGINVLGALIESS : 351
HmeHc2 : -GFQDLPW--LTVQDMKIEGRIRDAIAHGYVKA-YD-GN-VYSLNNT--EGINVLGALIESS : 331
TdoHc2 : -HFHDLPW--LTIEDMKGYEDRIRDAIAHGYVID-DS-LVYLNNT--RGLDILGALIESS : 329
SamHc1 : -AFHDLQN--IKVQDMIDYTRIRRECISSQAVIT-ET-GD-LYNISPP--EGINVLGALIESS : 355
BduHc1 : -EFQDLKN--IKVQDMIDYTRIRRECISSQAVIT-KN-GD-VYSLNNT--TGIDILGALIESS : 358
PamHc1 : -EFHDLHE--IKVQDMIDYTRIRRECISSQAVIT-RT-GE-MVSLNNT--TGIDILGALIESS : 354
CseHc1 : -EFQDLQD--IKVQDMIDYTRIRRECISSQAVIT-RG-GQ-NLPLNNT--MGINILGALIESS : 357
HmeHc1 : -EFQDLHE--IKVQDMIDYTRIRRECISSQAVIT-KS-GE-VYSLNNT--KGINTLGEIPESS : 340

CmoHc1 : : : : : EFQDLSS : : : : : IKVQDMVDYERRVREAIISQAVIT -MS -GD -YFSLNDT -TGINTLGEMMEPST : 356
CacHc1 : : : : : AFHDLEN : : : : : IKIKMDIDYTRRIRERIAIQAQAVIT -KS -GE -YFYLND -KGIEVLGDLMEPSF : 352
TdoHc1 : : : : : YFQDLLETPY : : : : : IKVKHMIAYSQIRERIAIARGVLT -KS -GE -LYSINDT -TGINTLGELIEPSV : 357
PmaHc1 : : : : : HFHLEH : : : : : IKIKMDLDYTRRIKEAISQKQVRS -KN -GE -KIPLDVA -HGIDILGDLMEPSV : 355
PgrHc1 : : : : : HFHLED : : : : : IKIKMDLDYTRRIKEAIVKHTVIN -KN -GE -HIPLDVA -HGIDILGDLMEPSV : 355
ScuHc1 : : : : : MFHDLK : : : : : HRVADLEAYERIRERIASAGVIFDAHD -PNTFISLNT -EGIDKLGATIEASS : 352
MgeHc1 : : : : : LLHDHLD : : : : : LKVHDVVFEERIKEAIDRGFVYN -KW -IGE -KIRLNNS -IEGIEITLGRMIEASR : 354
TdoHex1 : : : : : QKLDIDYKEEPEK : : : : : LKVRDNDYERRRDAIDRGYIFT -DE -NN -QVPLNND -KGIEVLGKIVGAT - : 367
PmaHex1 : : : : : SILNN : : : : : YHVFEALERRIHDAIDLGFVFG -KD -GQ -KISLKEK -EGISILGDMIEGT : 372
BdiHex : RSLYYSNNGSYYYGNYYGGNNYYTGNYYTGNYPHSYYYGYSTEHDY : : : : : YYPEDLIQYDRRLVDSIDYGVVFS -YP -DH -KYPLEDYETKGDYLDGVIEGNG : 404
RfLhex2 : RNLNYYNGQGYYYGNYYGGNNYYTGNYYTGNYPHSYYYGYSTEHDY : : : : : YYPEDIKSYEGIRIDGIDFYVFS -EG -GQPKYPLYDEYSKGINYLGDIEGNG : 406
PamHex12 : : : : : YVTNFDL : : : : : YYIADIKNYEKRVDAIDFGYAFD -EH -MK -PHSLYHD -VHGMEYLADEMEGNG : 360
RfLhex1 : : : : : LVTNFDT : : : : : YTIEDIKNYERRVADIDFGYFKD -EH -LK -VHSMYED -NGNINYLQGMIEGYS : 361
LmiHex1 : : : : : YPSNFDL : : : : : LFVERIQNYERRIRDAVDFGLY -YD -FK -TFNLNEK -DLTDIIGNVIEGNA : 363
LmiHex2 : : : : : FARNVDI : : : : : LYVEEIRNYERRIRDGIDYGLAG -YN -YE -KYNVREK -DYTNVLGNLIEGNE : 363
RmiHx1 : : : : : FPRNTDI : : : : : SYVEEIKDYRRIRDSIDYGVFAG -YN -YE -KYNVREK -DHTNILGNLIEGNA : 363
TmoHex2 : : : : : HLYEYFNYGQKYGF -NKYA -NSYSYVKDFERRIHVIDSGYAHT -ES -GQ -TVELFSH -HGMNVLGNLIEGNE : 371
AgeHx : : : : : KLYEYFNYGQSWTIRNEYG -YSPTFKDAERRINDAIDSGVAF -HE -GK -MIDLYS -HGMNVLGNLIEGNA : 380
LdeDP1 : : : : : KLYEYFNYGQKWSW -SRFGTNSFDMADYDRRMDAIDQGFVFD -TK -GN -KIELDTP -EGFNVLGNLIEGNS : 378
AgaHex2 : : : : : HVEFQSNY : : : : : RETRQVDYVKRLRDVIDQGFVFL -PD -GK -KVNLSPT -EISDVLGNLIEGNA : 346
MseAry1a : : : : : DIHNDKNY : : : : : EQIRFLDMFEMTLQYLQKHGFKAFDK - - - -EINFHDV -KAVNFVGNWQANA : 378
MseAry1b : : : : : NVHSEENY : : : : : EKVRFLDTYEMSFLQFLQKHGFKAFDQ - - - -KIDHFD -KAINFVGNWQANA : 375
CfeHex2 : : : : : NVPFYKY : : : : : KYLKEVEVLEKRIMDAIDLGVYHD -KD -GK -QVNIYTS -EGLSILANLIEGNE : 367
BheHex : : : : : RIPPYKY : : : : : HYLRYIHDIEDRISTADLGYIID -SD -GS -HNNISSP -EGLNVLGNLIEGNE : 366

EpiHc : Y - - - - SLNGAYY - - - - GD LHNTGHVMLATAHDPKRFNSSDGMVGFQVTLARLDP - FYRWKHKIDNLLQNHKR - TLQPYTDKELI - ADKIEIKASITSSQTKDPKNKLYTFFD : 420
LpoHc1 : E - - - - SVNGQFY - - - - GSHLHWGHVLTARAHDPGKGFHENPGVMSDSTSLRDP - FYRWHRITLNLFLQYKE - SLSPYTKHEELS - FPGVEVLSATV - KAKT - - - -DNVITSM : 408
TtrHcA : D - - - - SINPEYY - - - - GSHLHWGHVLTARAHDPGKGFHENPGVMSDSTSLRDP - FYRWHRIDNVFDQYKE - SLHPYTKHEELS - FTGKVVNVSV - NAKVITARNVITTFK : 407
EcaHcE : E - - - - TKNKLYY - - - - GSHLHWGHVMMARLQDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKIA - TLPHYTPEDLT - CPGVHVVNVTV - NAKV - - - -PNVVTTFMK : 407
NinHcE : E - - - - SKNKLFF - - - - GSHLHWGHVMMANITQDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYQY - TLKPYDKAQLD - FPGVQVNVTV - NAKV - - - -PNLVITTFMK : 408
EcaHcG : D - - - - SVNRFFY - - - - GSHLHWGHVMMARMTDPRSFEEPNVMSDSTSLRDP - FYRWHRFVNDIFQYKA - TLPSTADDLN - FPGLRIVSVQV - NAKS - - - -QNRVRTFLK : 409
NinHcG : D - - - - SPNKGFF - - - - GSHLHWGHVMMARMHDPGRFQENPGVMSDSTSLRDP - FYRWHRFVNDIFQYKYS - TLQPYTWEQLS - FPGVKAVCEV - RAKV - - - -NNVITTFMK : 408
CsaHc1 : S - - - - SPNKGFF - - - - GSHLHWGHVMMARMHDPGRFQENPGVMSDSTSLRDP - FYRWHRFVNDIFQYKYS - TLHPYTKHEELS - FAGVKIVNAYV - KAKK - - - -TNVVTTFMK : 408
AauHc6 : E - - - - SKNVEYY - - - - GSHLHWGHVMMANITQDPGRFQENPGVMSDSTSLRDP - FYRWHRFIDNIFQYKYS - SFHPYTKHEELS - FPGVEVGVSTV - NSKT - - - -ANVVTTFMK : 408
LpoHcIV : E - - - - SKNEEYY - - - - GSHLHWGHVLMARVHDHGRFHENPGVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLHMYTKTELT - FPEISISHVT - KAKV - - - -PNIHTTFMK : 409
EcaHcF : D - - - - SLNDYDY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - TLPCYKQDLE - FSGVEIVNCTV - NAKA - - - -PNLVITTFMK : 411
NinHcF : E - - - - SKNREYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - TLPCYKQDLE - FSGVEIVNCTV - NAKA - - - -PNLVITTFMK : 410
EcaHcD : D - - - - SVNKPFF - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - RLPSYTHQDL - FPGVRISRVTV - RSKV - - - -NNVITTFMK : 409
NinHcD : E - - - - SLNKGYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - RLPSYTHQDL - FPGVRISRVTV - RSKV - - - -PNLVITTFMK : 408
LpoHcII : E - - - - SVNHEYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - RLPSYTHQDL - FPGVRISRVTV - RSKV - - - -PNLVITTFMK : 408
EcaHcA : E - - - - SKNRQYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - TLPSYTHQDL - FPGVRISRVTV - RSKV - - - -PNLVITTFMK : 413
NinHcA : E - - - - SKNRQYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - TLPSYTHQDL - FPGVRISRVTV - RSKV - - - -PNLVITTFMK : 413
LpoHcVI : E - - - - SINQAFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - TLPHYNHSDLE - FPGIEITKVL - NAEH - - - -TNNTITTFK : 413
EcaHcB : D - - - - SVNKSYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 410
NinHcB : D - - - - SVNPFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
EcaHcC : D - - - - SKNRESY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
ScoHcA : L - - - - SVNRFFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
ScoHcB : Q - - - - SINRPFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
ScoHcC : T - - - - SLNRDYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
SpiHc1 : N - - - - SVNRFFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 411
GpuHc1 : Y - - - - SPNAQYY - - - - GALHNDHVLTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTAEEIG - FPGVHVTGVS - - - -EGELETFE : 447
GroHc1 : Y - - - - SPNAQYY - - - - GALHNDHVLTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTAEEIG - FPGVHVTGVS - - - -EGELETFE : 447
CscHc : Y - - - - SLNAGYY - - - - GALHNDHVLTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTAEEIG - FPGVHVTGVS - - - -EGELETFE : 447
EpuHc1 : Y - - - - SPNAGYY - - - - GALHNDHVLTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTAEEIG - FPGVHVTGVS - - - -EGELETFE : 447
PleHc2 : Y - - - - SPNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 456
HamHcA : Y - - - - SPNAQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 443
PinHcB : Y - - - - SPNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 427
PinHcA : Y - - - - SSNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 427
PeHc1 : Y - - - - SSNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 455
PeHc4 : Y - - - - SSNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 456
PeHc3 : Y - - - - SSNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 455
PeHc2 : Y - - - - SSNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 455
CmaHc6 : Y - - - - SPNPEYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 450
CmaHc4 : Y - - - - SPNPEYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 449
CmaHc5 : Y - - - - SPNPEYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 450
CsaHc : Y - - - - SPNPEYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 450
CmaHc3 : Y - - - - SPNPEYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 445
PvaHc1 : Y - - - - SPNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 445
PvaHc : Y - - - - SPNVQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 441
PinHcC : Y - - - - SANPQYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 433
PeHc : Y - - - - STNAAPPY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 439
CmaHc1 : S - - - - STNPSYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 439
CmaHc2 : S - - - - STNPSYY - - - - GALHNTAHVTLGRQADPHGKFLNPPSGMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 439
HamPhc1 : Y - - - - NPNYEEY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 451
HamPhc2 : Y - - - - NPNYEEY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 449
CmaCC1 : Y - - - - SPNRKYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 429
CmaCC2 : Y - - - - SPNRKYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 448
PmaHc2 : H - - - - DYNRHFF - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 451
PgrHc2 : H - - - - DYNRHFF - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 451
CseHc2 : N - - - - SPNVVYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 453
PamHc2 : H - - - - SVNPVYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 453
BduHc2 : H - - - - SVNPVYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 454
HmeHc2 : H - - - - SVNPVYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 434
TdoHc2 : H - - - - SQNKELY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 432
SamHc1 : G - - - - SKHPEYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 449
BduHc1 : D - - - - SKHPEYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 452
PamHc1 : D - - - - SKHPEYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 448
CseHc1 : E - - - - SPNPQYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 451
HmeHc1 : D - - - - SKHPEYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 434
CmoHc1 : T - - - - SKHRDYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 450
CacHc1 : D - - - - SKHPEYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 446
TdoHc1 : F - - - - SRHREYY - - - - GALHNYGHIMLGQITDPKGFNMPPGVMEHFETATRDP - FFRLLHKYMDNIFKEHKO - SLPPYTKEDIA - FPGVVLDSVT - - - -DGLKTTFE : 451
PmaHc1 : E - - - - SPHEYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 449
PgrHc1 : E - - - - SPHEYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 449
ScuHc1 : C - - - - STNPSYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 446
MgeHc1 : L - - - - SPIDIDY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 448
TdoHex1 : H - - - - DENDKYGNKHYH - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 466
PmaHex1 : D - - - - STNKGYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 466
BdiHex : D - - - - TVNRYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 498
RfLhex2 : D - - - - TVNRYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 500
PamHex12 : D - - - - SPNYFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 453
RfLhex1 : N - - - - SPNYFY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 455
LmiHex1 : E - - - - SVNPEYY - - - - GSHLHWGHVMMARHDPGRFNEPNVMSDSTSLRDP - FYRYHRFIDNIFQYKYS - SLTPYS - SGLQ - FKGVIKSVCV - RAKT - - - -ADVVETFA : 457
LmiHex2 : D - - - - SINKEYY - - - - GAFYNNLISLFGHIVDPVHRGVPASVLEQPETQLRDL - FYRIGKRVLSIFHYHKN - LLPRYTHELD - LPGAIVQDVT - - - -DGLKTTFE : 457
RmiHx1 : E - - - - SINKEYY - - - - GAFYNNLISLFGHIVDPVHRGVPASVLEQPETQLRDL - FYRIGKRVLSIFHYHKN - LLPRYTHELD - LPGAIVQDVT - - - -DGLKTTFE : 457
TmoHex2 : D - - - - SPYHY - - - - GAFYNNLISLFGHIVDPVHRGVPASVLEQPETQLRDL - FYRIGKRVLSIFHYHKN - LLPRYTHELD - LPGAIVQDVT - - - -DGLKTTFE : 466
AgeHx : D - - - - SPYQYY - - - - KALWYARHLLGYSQPLDKYKVASALEHFTSMRDPVFVFLQKKIVLQFQRYKA - RLHPYTHELD - FPGVKVLEW - - - -DGLKTTFE : 475
LdeDP1 : D - - - - SPNRKYY - - - - KALWYARHLLGYSQPLDKYKVASALEHFTSMRDPVFVFLQKKIVLQFQRYKA - RLHPYTHELD - FPGVKVLEW - - - -DGLKTTFE : 472
AgaHex2 : D - - - - SYHRFFY - - - - KYVAMYARHLLGYSQPLDKYKVASALEHFTSMRDPVFVFLQKKIVLQFQRYKA - RLHPYTHELD - FPGVKVLEW - - - -DGLKTTFE : 440
MseAry1a : DLYNEEVTKLQY - - - - RSYEYIARQVLGAAPKPFDKYTFMPSALDFYQTLRDL - FYQLNKRILKIVYEQYK - YLQPYSSSEKAL - FKGVKVLEW - - - -DGLKTTFE : 476
MseAry1b : DLYGEEVTDQY - - - - RSYEYIARQVLGAAPKPFDKYTFMPSALDFYQTLRDL - FYQLNKRILKIVYEQYK - YLQPYSSSEKAL - FKGVKVLEW - - - -DGLKTTFE : 473
CfeHex2 : D - - - - SCNRFFY - - - - GMYDALARDLIFGNDYKNNKVPASLQCYVTSRDP - FYRLTKRIMVFFRYK - NMPQYQDEL - - - -DGLKTTFE : 461

BheHex : D---SCNKNFY-----HSLDWYGRKVLGFNLEPKTPYQVIPSALSFSTCMRDP A-FYRLYNRYLSYWYRFKE-TLKPYSKNEIV--FSDLKFESIAV-----DKLITYFD : 460

	700	720	740	760	780	800		
EpilHc	: DKFQ138LSPGNNITRKSEDSNVTLHKEMTWEMENQPD	-----	TSGSPSOHYCACGPHHLLTPKGSAAQTAFAHCIMITD	TDWKDEVKSSKPP	-----	TCQDAVSYCGII	-NDLYP : 602	
LpHcIV	: DKFKHKLKAGKNIETNRNSDVSVTSLTLRDELQESSV	-----	GVNENADEYCSGCGPWKMLVPRGNKGMFTFLVMLTDW	EDNVNKGDDGCK	-----	NHMDCKDAVSYCGAK	-DSKYP : 590	
TrfHCa	: DKFKHKLISQNIITHNAASSVSTLSTLRTFEDLKAGK	-----	GVSEDATEYCSGCGPWMLIPRGKNGMVFELVMLTDW	EDNVNKGDDG	-----	FCSDAVSYCGAK	-DHKYP : 585	
EcaHcE	: DKFKATLDPGKNVTRDRHRNSTVVEQSVPPVKLLREE	-----	GV-----	AGEYCSGCGPEHMLTPRGKNGRMDGDFL	FVIVTDYQDAVANGHEN	-----	FCSDAVSYCGAK	-DQKYP : 585
NiHcE	: DKFQVELAAGKNATIRDRHSSVTVSETHTFKQLLAGE	-----	GVSENTTEFCSCGPEHMLTPRGSHKSGEFLD	FVILTDYEQDAVGEASG	-----	ECVDAISYCGAK	-DQKYP : 588	
EcaHcG	: DKFKHKLGPGLNITNRNSDVSVTAHTVTFDELREGK	-----	GVSEDATEYCNCGPWPHMLTPKGTIRHGMFQFVML	TDYTVDNPCGGAG	-I	VCDAVSYCGAK	-DQKYP : 590	
NiHcG	: DKFKHKLAPGKNVITNRNAISDNVLSHTVTFDELRAQ	-----	SASADASEYCSGCGPEHMLVPRGTTHKGMDQFVML	TDYSDENPEAGANV	-I	CINDAVSYCGAK	-DQKYP : 589	
CsahC1	: DKFKHKLPGKNGVTHRAAFNSVTLSTHTVTFDELRSGL	-----	SASQDATEYCSGCGPEHMLTPRGTKTIGEFELVMA	TDYTHDNPEADVK	-I	ICGDAVSYCGAK	-DQKYP : 589	
AauHcE	: DKFFYVTLPGKNTIVRNRHSDSVTSIKSVRTFDQLAGE	-----	GVSEDSYEYCSGCGPEHMLTPRGSHKGMFELFVML	TDHEDDVAUGLEN	-----	VCSDAVSYCGAR	-DRDYP : 589	
LpHcIV	: DKFFHHTIGKHNITNRNSDVSVTKITHNFLNKAQV	-----	GVDEKNTIEYCSGCGPWMLTPRGKNGMDQFVLMV	TDYEDHVDNGNEPSL	-----	ICDAVSYCGAK	-DHKYP : 587	
EcaHcF	: DKFKHAELAAGKNITRKSESSVTVSHPTFEELQREG	-----	GVIDENTTEFCSCGPEHMLVPRGTYKGMDFQFVML	TDYEDHVDGSHNGQ	-I	LCDAVSYCGAK	-DSKYP : 590	
NiHcF	: DKFKHAELAPGKNVTRKSESSVTPVKPKFSQLQSGE	-----	GVSENTEYFCSCGPEHMLVPRGTTHKGMDYFVLM	TDYEDHVDNGLNEK	-S	ICDAVSYCGAK	-DQKYP : 589	
EcaHcD	: DKFKHRRLEPGKNVLVRASGDSVTLSKVPTFEELSEG	-----	NANVNPNNEYCSGCGPEHMLVPRGKERGMDFVLM	TDYEDDQYGAQEG	-I	ICSDAVSYCGAK	-DQKYP : 587	
NiHcD	: DKFTTTLTLPGRNITNRNSDVSVTLSHVPTFEELKGE	-----	NIPNTRTEYFCSCGPEHMLVPRGTGRGMVFHFLVML	TDYEDDQYLDASAT	-I	VCSDAVSYCGAK	-DQKYP : 587	
LpHcII	: DKFKTDLHPGKNITVRRSLDSVSLTHQPTFDELLHGV	-----	GLNEHKSEYCSGCGPSSLVPGKNIKGMEYHFLVML	TDWDDKKYDVSQSG	-----	ACVDAVSYCGAK	-DHKYP : 592	
EcaHcA	: DKFYHTLTPGKNTIVRNSDTSVTLSSVHTFKELLREG	-----	DLVEGQTEYFCSCGQPQLHLPVKGNEKGMDQFVLM	TDASVDRYQSGDGT	-P	LCADALSYCGVL	-DQKYP : 594	
NiHcA	: DKFTDTLRPGKNIIVRSSTDSVTSSITSKFLLHAGE	-----	DLEEDRSFEYCSGCGPQLHLPVKGNEKGMDVFLVIM	TDADAEKKQVTSGKK	-----	LCNDALSYCGVM	-DEKYP : 593	
LpHcVI	: DKFDHAEHPGNTIVRHSKDSVTSISARSFGQLAHGE	-----	GINEANEYCSGCGPDHLLVPRGDEKGMDFHFLVML	TDWDDQVGDHGT	-G	MCTDAVSYCGAK	-DQLYP : 594	
EcaHcB	: DKFTTTLVLPGENVIRLSDSDSVTLREQYTVLRQDQGR	-----	-SNREQTEYCSGCGPWDLHLPVKGNEKGKFRFLVIM	TDADVQGVGDHGT	-G	LCTDAVSYCGAK	-DQLYP : 590	
NiHcB	: DNHFHVLPGKNTIVRRSLDSVSLTKQPTLLKQGLQK	-----	GTDRRDGEYCSGCGPDHLLTPKGDSGGMKHLFAL	FTNYFEDVNDHGR	-I	ECVDAVSYCGAK	-DQLYP : 592	

EcaHcC : DRFRVELKPGVNNHHRSDSSVTISKQPKFDELLKGGK-----GTNN-ANEYCSCGWPDHLLVPKGNDRGMPFHLVCMITDLYLDLVGDYGYD-T-PCVDVAVSYCGAK-DSLYP : 593
ScaHcA : DKFVATLEPGHNVIERKSKSSVTMSKDHFVGEIRNLA-----DDHQSCSGWPDYLLLPKGGYEGMTYQLFVAVTDYEDHVEDAGE--ECQCRDSMSYCGSV-EHKL : 614
ScaHcB : DKFVATVPAGSSTVEQHSSKSSVTQSDNDFYGSATRS-----SENHCSCGWPDYLLLPKGNHQQGVQFNVYVIATSYDEDHVES-DE--SCHCGDLSLYCGAL-YDKYP : 617
ScaHcC : DKFIEKLYPGKNTIERHSEDSTVTPMSSASIFSDISSEQ-----SEDHCSGCGWPDYLLVPKGNFEGFPMFVFIVTDYEDKVEGPDE--GCACHDALTYCGGI-DYHFP : 624
ScaHcD : DKFTVNLKPGKNSVRQPCYNAVATKPYDVFYGDVESQK-----PQEGCNCGWPDYMLLPKGGYEGLRFRVFAIVTNHDEKVKSD-QE--TCLCGDAVAYCGAH-NQKYQ : 622
SpiHc1 : DKFKATLPGHNVIRCSNDSVTMPDHYIGDITRTV-----DEDHCHCGWPEYLLVPQGDYDGMHYELFVMATNYDEKQAQDITDL--SCKCH-AHSYCGNI-LGEYL : 611
GpuHc1 : DKFWTKLAAGDNH1KRKSSSESVTPDVPFSFASLIHDADAASVAGSELHLEDFDRSCGLPQRMLLPKGTEEGDLFLLVAVSDGTTDAQHDALE--AVDAHGHQAQGVH-GEKYP : 636
GroHc1 : DKFWTKLAAGDNH1KRKSSSESVTPDVPFSFASLIHDADAASVAGSELHLEDFDRAGCLPQRMLLPKGTEEGLEFLLVAVSDGTTDAQHDALE--AVDAHGHQAQGVH-GEKYP : 636
CscHc : DKFWTKLAPGVNSITRKSSESVTPDVPFSFKRLIADTKAALAEKVEKFEEYHRSCTGPNRLLLPKGSSEGMKFLVAVATEASSDAHEPIL--IQGAHSHAQCGVQ-GEKYP : 638
EpuHc1 : DKFWRTLHSGDNH1ERKSSDASTVTPDIPSYESLMKMAAASVAGEELHIEQFTRSCGIPNRLLPKGTPEGMEFALFVAVTDGSKDAIEGLE--KDEHGGSQAQCGIH-GEVYP : 636
PleHc2 : DKFYKSLAPGINH1VRKSTDSSTVTPDRPSFNILKSKTDRAVSSSHKLDLHKFERSCTGPNRLLLPKKGARGMDFALFVGVTNGDEAKTVDDPE--ELGATHQAQCGIH-GEKYP : 645
HamHcA : DKFYKQLKTGSKNIDRKSTESSVTPDPSLSLAELHEKTDQAVASGSKLDLHEYERSCTGPNRLLLPKGTVNGMDFLFAVATNGDEDKVDTHPE--TDEHGGSQAQCGAH-GEKYP : 634
PinHcB : DKFFQKVPSPGETIERSSKSSSVTPDMPFSQSLKEQADNAVNGHDLDSLAYSERSCTGPNRMLLPKSKPEGMEFNLYAVATDGDKDTTEGHNG--GHYDGGTHQAQGVH-GEAFP : 617
PinHcA : DKFFQKVPKGPETIERSSKSSSVTPDMPFSQSLKEQADNAVNGHDLDSLAYSERSCTGPNRMLLPKSKPEGMEFNLYAVATDGDKDTTEGHNG--GHYDGGTHQAQGVH-GEAYP : 617
PeHc1 : DKFWKRLANGKNLIERSSKSSSVTPDMPFSFESLKHQADEAVKNGEDLDSLAYSERSCTGPNRMLLPKSKPQGMFENLYVAVTDGDKDTGSDGD--HDHGGTHQAQCGIH-GEQYP : 644
PleHc4 : DKFWKLLANGKNLIERSSKSSSVTPDMPFSFESLKHQADEAVKNGEDLDSLAYSERSCTGPNRMLLPKSKPQGMFENLYVAVTDGDKDTGSDGD--HDHGGTHQAQCGIH-GEKYP : 645
PleHc3 : DKFWKLLGNGLNMIERSSKSSSVTPDMPFSFESLKHQADEAVKNGEDLDSLAYSERSCTGPNRMLLPKSKPQGMFENLYVAVTDGDKDTGSDGD--HDHGGTHQAQCGIH-GEQYP : 645
PleHc2 : DKFWTKLAGEKNTIERSSKSSSVTPDMPFSFESLKHQADEAVKNGEDLDSLAYSERSCTGPNRMLLPKSKPQGMFENLYVAVTDGDKDTGSDGD--HDHGGTHQAQCGIH-GEQYP : 644
CmaHc6 : DKFWKLASAGTNTITRKSSESVTPDVPFSQDLIEKTEAVSSGSELHLEEVYSLGLPNRFLPKGTGTQGMKFLVVFVSDGAKDAADIGLH--ENTSFNHY-GAH-NGKYP : 638
CmaHc4 : DKWVTKLTPGNTITRKSSESVTPDVPFSFLITLIEKTEAVSSGSELHLEEVYSLGLPNRFLPKGTGTQGMKFLVVFVSDGAKDAADIGLH--ENTSFNHY-GAH-SGKYP : 637
CmaHc5 : DKFWTKLAPGNTITRKSSESVTPDVPFSFLITLIEKTEAVSSGSELHLEEVYSLGLPNRFLPKGTGTQGMKFLVVFVSDGAKDAADIGLH--ENTSFNHY-GAH-NGKYP : 638
CsaHc : AKFWTKLNGSNIERSSKDFGVTPDVPFSQTLMDMTDEALGSSGSLHEDYVSGGLPDRFLLPKGTTEGMDFHVVVFVSDGAKDAADIGLH--ESTYFNHY-GCH-DGTP : 638
CmaHc3 : DKFWVMLHPGHNHIERSSDSSATVPDIPSFQIKORTNEAIAQNKLHIEEFESGLGLPNRFLIPKGNVKGDLMDVMVAITSGEADAIVEGLH--ENTSFNHY-GCA-DGTP : 633
PvaHc1 : DKFWKLP6GTHHIERKSESAVTPDVPFSFATLFEKTEAL--GAADSLKDFESATGIPNRFLLPKGNEEGLFEDLVAVTDGAADAADVGLH--ENTEFNHY-GSH-GVYP : 639
PvaHc : DKFWVSLKGGNTIERKSSDSTVTPDVPFSQDLIEKTEAVSSGSELHLEEVYSLGLPNRFLPKGTTEGMDFHVVVFVSDGAKDAADIGLH--ENTYFNHY-GSH-GVYH : 622
PinHcC : DRFWVVKHGGHQTIRQSTESSVTPDVPFSQTLIDRADAASGSCALHLEDYSEALGLPNRFLLPKGAQGMFENLYVAVTDGRTDAALDDLH--ENTKFTHY-G-Y-DQYP : 629
PleHc : DTFTVEVHAGHNVIERKSSSESVTPDRVSFQLIHQADEAVANGAELP-HKESRSCTGPNRLLLPKSGKEQGMDFWLDIIITSGD-DAVQDILT-VNAGHGHQAQGVH-GEKYP : 626
CmaHc1 : DTFTVEVAGHGHNNHRSFGESNAIIPDRISFEELIKETDEAVNNLELSLNSG-RSCGHPQRLLLPKGNEEGLQFWLVVYVTSGE-DAVHDDLH--TDHGH-NHGYCIGI-QAKYP : 625
CmaHc2 : DTFWVDVAGHGHNVIERKSSSESVTPDRISFEELIKETDEAVNNLELSLNSG-RSCGHPQRLLLPKGNEEGLQFWLVVYVTSGE-DAVHDDLH--TDHGH-NHGYCIGI-QAKYP : 626
HamHc1 : DKFWKYLAPGENEVTIRKSDSSVTPDVPFSLSLHEQAEAAIAGSELNLEEFVSATGLPNRLLLPKGAAGVEFCLVAVTDGEADSVNDEIN--LTKFHHY-GHH-G-VYL : 641
HamHc2 : DKFQKELIPGNTITRKSSESVTPDVPFSLSLHEQAEAAIAGSELNLEEFVSATGLPNRLLLPKGAAGVEFCLVAVTDGEADSVNDEIN--LTKFHHY-GHH-G-VYL : 639
CmaC1 : DLFVTKLTSGNIEDRKSSSESVTPDVPVTKYKLVKTE--ENRENLEMYESSGTIPNRLLPKGCCEGTFRPLVAVTDAENDVNDESII--TMQKYHHY-GVR-G-VQP : 611
CmaC2 : DLFVTKLTSGNIEIRKSSSESVTPDVPVTKYKLVKTE--EDHQSLERYESATGIPNRLLPKKGNEKGETFRVLVAATNAEEDINDESII--TMNKHYYH-GVR-G-VQP : 630
PmaHc2 : DRFPVKLTAGENVITRSGKSVVTDDEPMSFAET---HKAVADKDATFHFKFRHCGFPHRLLLPVKGPRPEGMHYKLMVVIITDYHKDVVVPMDVHEHMDKLSQVSGYGVH-EGKTP : 636
PgrHc2 : DRFPVKLTAGENVITRSGKSVVTDDEPMSFAET---HKAVADKDATFHFKFRHCGFPHRLLLPVKGPRPEGMHYKLMVVIITDYHKDVVVPMDVHEHMDKLSQVSGYGVH-EGKTP : 636
CseHc2 : DRFPVRLASAGENVITRNSVDSIITVPEPRSFNLVSDVEKAIERGEQVFSYDKYRHCGFPHRLLLPKKGPEGMVYQLVYVVTDYDKDIVEPQIPEDEVNLSLGLCYGVH-DGKTP : 644
PamHc2 : DRFPVSLVHGDINVTIRNSDSTVTPDPSFPLDMSDQVKAIRGEQFVYDHYRHCGFPHRLLLPKKGPEGMVYQLVYVVTDYDKDVAAPGMSPDVNTLSLGLCYGVH-DGKTP : 644
BduHc2 : DRFPVKLVNGDVTIRSSDQSVTTPPEKSPYELVNDVERARGEVDFYDKHRRHCGLPNRLLPKKGPEGMDFCLVYVVSDFDKDVSSEVNVDDLHHTISLGLCYGVH-DGKTP : 645
HmeHc2 : DRFPVKLEAGDNVITRNSDSTVTPDPSFPLINVERAINGEEYVYDKYRHCGFPHRLLLPKKGPEGMVYQLVYVVTDYDKDVAAPGMSPDVNTLSLGLCYGVH-EGKTP : 625
TdoHc2 : DRFPVHLRQGDNVIRSSQDSTVTPPEPRSPYELIKEIEDALNGQDEYAIKYRHCGFPHRLLLPKGTSGGMVFNLYVITITNYEQDKVHPGVTVDEVDNLTSGLCYGVH-DGVT : 623
SamHc1 : DKFVAKVLAGENVITRKSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 638
BduHc1 : DKFVAKVNGKTTVQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 638
PamHc1 : DKFVTKVNGKTTVQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 638
CseHc1 : DKFVAKVNGKTTVQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 638
HmeHc1 : DKFYTTVNGKTTVQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 624
CmoHc1 : DKFVTKVNSGKTTQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 639
CacHc1 : DKFVTKLAGENVITRKSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 635
TdoHc1 : DKFVAKVAGSQTPQVRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 640
PmaHc1 : DKFVTKVAGKSVIRKSSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 639
PgrHc1 : DKFVAKVLAGENVITRKSSESVTPDRETTKVLQVRVEDAIGAKVLTYNKDVHRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 635
ScaHc1 : DKFVAKLTAGENVITRKSSESVTPDVPVSTHDLRLKLDVDAIAGTAEVLDKVRHCGVPDRLLLPKGGRRDGMPTFLFVLTDEYKDKVNDLPF--DYDGGSVSYCGTL-GHYP : 635
MgeHc1 : DKFLTHIPAGSVIERLSTESTVTPDYESFRSLVNRVDEALKGNKEFIDEERFHCGLPDRLLLPKGNNEEGYVKFVFIVTDWEEDKVNQVE--NHRYGNVSYCGTYGDRPL : 638
TdoHex1 : DRFPVTLTSGKMNQRESKDFNSIVPDHPIKTNLKHMKVSAVEGEQTFYVQKGRQGLPEHLLLPKGGKAGLPFTLFMTENENNEQD----NVNYQOPFLICGVK-EGKYP : 669
PmaHex1 : DRFTFKMTSGKMNQRESKDFNSIVPDHPIKTNLKHMKVSAVEGEQTFYVQKGRQGLPEHLLLPKGGKAGLPFTLFMTENENNEQD----NVNYQOPFLICGVK-EGKYP : 669
BdiHc : DRFPYKQAGKTTITRNSRESSVSHDPSFRTLKRVFDAYEGKEQFYDYKSERYCGYPERLLLPKGGTGGQTYTFVYVMTVPYVKQDHDHF--EPYNYKFSYCYGVANHKFP : 688
RfLhex2 : DRFPYKQAGKTTITRNSRESSVSHDPSFRTLKRVFDAYEGKEQFYDYKSERYCGYPERLLLPKGGTGGQTYTFVYVMTVPYVKQDHDHF--EPYNYKFSYCYGVANHKFP : 688
PamHc12 : DRFPYHVGAGKTVIERNSHDSNIAPERDSYRTFYKQVQAEYEGKSQYVYDKGNHYCGYPERLLLPKGGKGGQTYTFVYVMTVPYVKQDHDHF--EPYNYKFSYCYGVANHKFP : 642
RfLhex1 : DRFRHHYVAGKSVIERNSHDSNIAPERDSYRTFYKQVQAEYEGKSQYVYDKGNHYCGYPERLLLPKGGKGGQTYTFVYVMTVPYVKQDHDHF--EPYNYKFSYCYGVANHKFP : 645
LmiHc1 : DIFNQKLNAGENEIRKSSKEFTLFAKEAPSYDYLYQTFRALKGEDKFLDKFRSHFGPQRLALPRGTRSGLLPLSVFAITPAVQSGSEHPV--PYH : 631
LmiHc2 : DIFNQKLNAGENEIRKSSKEFTLFAKEAPSYDYLYQTFRALKGEDKFLDKFRSHFGPQRLALPRGTRSGLLPLSVFAITPAVQSGSEHPV--PYH : 631
RmiHc1 : DIFNQKLNAGENEIRKSSKEFTLFAKEAPSYDYLYQTFRALKGEDKFLDKFRSHFGPQRLALPRGTRSGLLPLSVFAITPAVQSGSEHPV--PYH : 631
TmoHex2 : DHFVYKLKAGDNISIRSDSHETYNMHORTSYEYLQKAFGVHDQSHSFQFHHDQFWGYGQRYMLPMGSRGGMTYQFYVMYSGPKYKTH-----TTPVPMIGS-GMQVY : 648
AgeHc : DHFVYDLKTGENVITRNSDNYFFTPORTSYEYMYQVQLGASNGQ-DYHVDGRQNYFYFPQRYMLPKGTEDGMTQFYFYVIVSPYKAFHGHQHH--DQDHYFNYPKVT-AGQVY : 666
LdeDp1 : EHFVYDLKTGENVITRNSDNYFFTPORTSYEYMYQVQLGASNGQ-DYHVDGRQNYFYFPQRYMLPKGTEDGMTQFYFYVIVSPYKAFHGHQHH--DQDHYFNYPKVT-AGQVY : 666
AgeHc2 : DIFVQELTAGKNVLTIRSSVFNFGYKORTSFYELYRSVMAGYGAEKFLDMSSEAHCGFPHRLLLPKGGKGGMPQLYVIVSPYKAPQVS--QYSGFDPVLSGCVGSGARYM : 642
MseAry1a : DIFVYKVMGPNQNHVIRKSSDFFLFKEDSLPMSEIYKLDGKIPS-----DMSNSDTPQLRMLPRGTDGQYFPQLFVFVYPPQAVPKMEFP--KSVYP : 649
MseAry1b : DWFYTKFVAGDNKIVRNSNDFTLFKDDSDVPMTELYKLEQNKVPH-----DMSDYGVLPRMLPRGTEGGFPFQFVYVYFVFNADSKDLAPF-----EAFQI : 647
CfeHc2 : DHEFTVLKQGTNTIERHSSDSTVTPDPLVSDSTFYKLERAMSGSAPFTY--VEKFTLPLERLVLPMKGPMEGRMKMFYSLTDSGSKVRNVLP-IF-----GKTLT : 641
BheHex : DYWITDYNAGLNKLEIRTSDFIFLMADRDPSIILYKRVKALDGSSEKFMY--KNKLYGPERLLLPKGRAGSIFQLFAYVSPVTPYVYKSRVF-----GSYQY : 636

820 * 840 * 860 * 880 * 900 * 920
EpiHc : DKKPMGFPDRILIPSDLATFVRSNMKIFDISIIHREKK----- : 641
LpoHcI : DKKPMGFPDRIRIDAHDIIEFLTPNMALTDVKIKFEG----- : 627
TtrHcA : DKKPMGFPDRKIENEHLEDFLTNMGVSDIKIQFMG----- : 622
EcaHcE : DKKPMGFPDRVIEGLTLEEFLTSPMSCTDVIKITYDIK----- : 624
NinHcE : DKKPMGFPDRVITADTVAEFLTPNMSCTDVIKIHVI----- : 625
EcaHcG : DKKPMGFPDRPTKIHTAEELTPNMSLTDVVIQYVGHE----- : 629
NinHcG : DKKPMGFPDRPLLSSVAALPTENTCTDVIKIFLG----- : 626
CsaHc1 : DKKPMGFPDRPTTARSAGQELLTDNMSLTTVKIQYLGKRYACVGN----- : 634
AauHc6 : DKKAMGFPDRKIEARTAAEFLTPNMGLTDIKIKFHG----- : 626
LpoHcIV : DKKPMGFPDRPIDARTPSQFATPNMSFTDIRIQFSH----- : 624
EcaHcF : DKRAMGFPDRVIRKARTVADFRITNMSFTDVKIQKFDQV----- : 629
NinHcF : DKKSMGFPDRVIRKARTLAEFSSANMKFQEVKIQFKE----- : 626
EcaHcD : DKKAMGFPDRPIQVTRPSQFKTPNMAFQEIITIQYEGHKH----- : 627
NinHcD : DRRAMGFPDRPIKARTPSQFKTNMSFTVIRIQYGGYKA----- : 627
LpoHcII : DKKPMGFPDRPIHTEHISDFLTNNMFIKDIKIKFHE----- : 629
EcaHcA : DKRAMGFPDRKIDATDHEEFLTNMNIISHTVRFQD----- : 631
NinHcA : DKRAMGFPDRITITAQSHIEFTPNMKISNITVRFQD----- : 630
LpoHcVI : DRRPMGFPDRHIELEHLKDWLPNMKDTLITVTHHEEVHNNH----- : 638
EcaHcB : DRYPMGFPDRDIDKADSIPEWLPNMFSEVITITHQ----- : 627
NinHcB : DKRAMGFPDRDIEVANDFNWRQPNMISIPDIVHS----- : 628
EcaHcC : DRRAMGFPDRPIPEGHASNLHQPNSVFSQIKIQH----- : 629
ScaHcA : DNKPLGYFPDRIDGTGFEFEKTNQMYGDDVVIQFTGETVTH----- : 656
ScaHcB : DRRPMGFPDRHADAQTFDEFTKNMNSVTVTIKHTGEVKDA----- : 659
ScaHcC : DKRAMGFPDRPIKQNFNAFKTKNMKGKTVDVKFTGETIAPE--DFHNQH----- : 673
ScaHcD : DKKPMGFPDRDIDERTFEHFTPNMIATDVIKFTGEFLPPKGD----- : 668
SpiHc1 : DKRPLGYFPDRKIKATGWAEFKTNMCDADVIKFSGKNK-----DH----- : 653
GpuHc1 : DHQPMGFLDRIPDDRIFL-SADNIEKLIVTVHKKH----- : 672
GroHc1 : DHQPMGFLDRIPDDRIFL-SADNIEKLIVTVHKKH----- : 672
CscHc : DSKPMGFLDRIDIEDERILL-SSSNLYTVVVKITHK----- : 674
EpuHc1 : DKRPLGFLDRQIPDERILE-HFPNLKHSIVVHHEEHKDH----- : 677
PleHc2 : DKKPMGYPDVRSDPNRVL-ESPNIKRVYVVFHDEHDEHQ----- : 687
HamHcA : DKKPMGYPLERSIPDERVH-DVANINVTWKVFFKEEH----- : 672
PinHcB : DNRLGYPLERRIPDERVID-GVSNIKHVVVIVHLEHHD----- : 657
PinHcA : DNRLGYPLERRIPDERVID-GVSNIKHVVVIVHLEHHD----- : 657
PeHc1 : DNRLGYPLERRIPDERVID-GVSNIKHALVKIVHDPDLRA----- : 684
PeHc4 : DNRLGYPLERRIPDERVID-GVSNIKHALVKIVHDPDLRA----- : 685
PleHc3 : DNRLGYPLERRIPDERVID-GVSNIKHALVKIVHDPDLRA----- : 685
PleHc2 : DNRLGYPLERRIPDERVID-GVSNIKHALVKIVHDPDLRA----- : 684
CmaHc6 : DKRPHGYPLDRVRDDKRIFS-GVTNICKMNVKVVHYVETO----- : 676
CmaHc4 : DKQPHGYPLDRVRDDKRIFT-GVTNFKGMDVKVYHVEEQ----- : 675
CmaHc5 : DKQPHGYPLDRVRDDERIT-GVTNFKGMDVKVYHVEEQ----- : 676
CsaHc : DNQPHGYPLDRVRDDERIT-GVSNFKAVDVKVVYHVEDH----- : 676
CmaHc3 : DKRPHGYPLDRVDDERIT-DLHNFHKKIVQVVFHHA----- : 669

```

PvaHc1 : DKRPHGYPLDRKVPDERVFE-DLPNFGHIHLKVFNHGEHIHH----- : 671
PvaHc : DKRPHGYPLDRKVPDERVFE-DLPNFKHQVKVFNHGEHIH----- : 662
PinHcC : DKRPHGYPLDRRVDDERIFE-ALPNFKQRTVKLYSHGVDG-G----- : 661
PleHc : DKRPMGFPPDRPIDLRVFK--VQNQGQVVKIFHH----- : 660
CmaHc1 : DKRPMGFPPERRVPDIRVK-NLPNFFGKIVDVYHKDH----- : 662
CmaHc2 : DKRPMGFPPYERRIPNKFLVL-GLPNHHSQEVNVYHKEE----- : 663
HamPhc1 : DKKSHGYPLDRRVDERLHF-EIPNFGETIVKVFNHNEHVHRHE----- : 684
HamPhc2 : DKKPHGYPLDRRVDERLHF-EIPNFGETIVKVFNRDEHVYHH----- : 681
CmaCC1 : DNRPFGYPLDRRVDEHIVD-EVSNIKATMKVYNHNVFIPLS----- : 653
CmaCC2 : DKRPFGYPLDRRVDEHIVN-EVSSIKETMKVYNHNDVLIRLPHY----- : 674
PmaHc2 : DGKPMGYPPDRRISCEESFI--TKNMKFVDITVKTRV----- : 671
PgrHc2 : DGKPMGFPPDRVPCEETFI--TKNMKFVDITVKTRV----- : 671
CseHc2 : DGRPMGFPPDRRIVSPEQFL--TPNMEVVDITIKNVNQKPAE----- : 684
PamHc2 : DGKPMGFPPDRRVSEETFY--TPNIKTIDLTIKNVKQKPT----- : 684
BduHc2 : DGKPMGFPPDRRIPSEELFL--TPNMKVNLVIKNVRQKPE----- : 685
HmeHc2 : DARPMGFPPDRRITSEEEFF--TQNMKVVDVITPNVKENR----- : 663
TdoHc2 : DGRPMGYPPDRIPDKEVFN--IPNSKIVIEVIKNAK----- : 658
SamHc1 : DAKPMGFPPDRRIDQD-SFF--TKNIYQRDVITIFKGSV----- : 674
BduHc1 : DSKPMGFPPDRRIDGD-HFF--TENMYQKEVKITFQDNE----- : 678
PamHc1 : VNKPMGFPPDRRIDGD-NFC--TPNMFQKDVITIFKDTI----- : 674
CseHc1 : DSKPMGFPPDRQINSD-NFF--RSNIYQKDVITIFKGDSE----- : 677
HmeHc1 : DNKPMGFPPDRRIDED-HFF--TPNMGQKEVITISFKEVP----- : 660
CmoHc1 : DSRPMGFPPDRRLSTE-EFS--MPNFCSDKDVITIFKGDSS----- : 676
CacHc1 : DSKPMGYPPDRPI-GR-EFY--YPMNFEKDVVITHKES----- : 669
TdoHc1 : DTQPMGYPPDRRIENVEEFL--TPNMYVKDVVITHN-DAE----- : 677
PmaHc1 : DTKAMGFPPDRRISREDF--TDNMYTKDVITIFKFNHHH----- : 678
PgrHc1 : DTKAMGFPPDRRISREDF--TDNMYTKDVITIFNENHHK----- : 678
ScuHc1 : DKKPMGWPLDRKICKSFHD--VTNMYFRDVEIKFEHTH----- : 672
MgeHc1 : DKKAFLYPDRVICKDVNVFK--TPNMFQKDVSIYHKDISELNRVVPVVEE----- : 686
TdoHex1 : DTKPMGYPPDRKFDERDIKS-ATNVYVKDVLIYHKRDVSESSDKYSQYDDL----- : 711
PmaHex1 : DSYSMGYPPDRITIEHY-DFFYNAPMNHFDVVIIFHKTAELINIKVHH----- : 702
BdiHex : DDKPFGYPPDRVLYSQ-EFV--TPNMYFKDVVIYHKKYEEINAATVQ----- : 733
Rf lHex2 : DYKPLGYPPDRPIYGS-DFY--TTNSYFKDVVIYHKKKEEVNTAITQ----- : 734
PamHex12 : DNKPLGYPPDRKIYSN-DFY--TPNMYFKDVIIHKKYDEVGVQGH----- : 685
Rf lHex1 : DMPLGLFPDRQIHSK-DFY--THQHVLRGCTIFPQETRRSQYSHPLGQQCEISTSHLLSGNGFSISKPKATDILQSDVLTLVNIIILRFSGMCCRILW----- : 740
LmiHex1 : DNQAAGFPDRRVVEF-EFD--VPNVYFGETYVVRHRRVEDINTTA----- : 673
LmiHex2 : DLHAAGFPDRRVVEF-EFD--VPNAHFDETFVVRHRRVEDLNATA----- : 673
RmiHx1 : DNHAAGFPDRRIVEL-EFD--VPNAHFDETFVVRHRRLEDINSTA----- : 673
TmoHex2 : DSYPMGYPPDRPVYEQVFY-DIPNSYFHEVKIYHEHHDHGIEQHLEATHYMV----- : 702
AgeHex : DGYALGHPDRHIKFEKMMH-NVPNSYFYETKIYHKHSYDDHIHHHLHHEHYTEL----- : 720
LdeDP1 : DNRTAGFPDRVIRFEKMWY-PLANGQFQEAQVYFKDIYDINAPHH----- : 702
AgaHex2 : DSYAFGYPPDRPIDEKLFYA--VPNAFFQDVSIYHKSE----- : 678
MseAryla : DSKPFGYPPDRPVHPEYFKQ---PNMHFEDVHVYHGEQFPYKFNVVPFVYPQKVEV----- : 702
MseArylb : DNKPLGYPPDRPVVDAYFKQ---HNMFFKDVHVYHGEQFPYKFNVPSHVMHSNVVPKH----- : 703
CfeHex2 : DEKPLDFPLDKPMHW-KFF--TPNMLKDVYIYHIPGNTINENRMTDDTMIENWMNENWLNEDRVNEDRLNEDTMNKSWLKEHRMNKDTINESWLKENRMKDTMDENTMNL----- : 750
BheHex : YMKPGGFPLDRPIYP-HFQG--PNMFFKDTIYHKTDVDPNATT----- : 678
*****

```

Fig. S18. Multiple alignment of the putative hemocyanin subunits and the previously identified insect hemocyanin subunit types 1 and 2, selected insect hexamerins, crustacean hemocyanins, crustacean pseudohemocyanins, chelicerate hemocyanins, myriapod hemocyanins, and one onychophoran hemocyanin used for the phylogenetic analyses. The N- and C-terminal sequences were deleted from the alignment for phylogenetic analyses (denoted by asterisks)