J. Biochem. **138**, 303–312 (2005)
DOI: 10.1093/jb/mvi130

Structure of Hemocyanin Subunit CaeSS2 of the Crustacean Mediterranean Crab *Carcinus aestuarii*

Pavlina Dolashka-Angelova[1,*], Alexandar Dolashki[2], Savvas N. Savvides[3], Rumyana Hristova[1], Jozef Van Beeumen[3], Wolfgang Voelter[2], Bart Devreese[3], Ulrich Weser[4], Paolo Di Muro[5], Benedetto Salvato[5] and Stefan Stevanovic[6]

[1]Institute of Organic Chemistry, Bulgarian Academy of Sciences, G. Bonchev 9, Sofia 1113, Bulgaria; [2]Abteilung für Physikalische Biochemie des Physiologisch-chemischen Instituts der Universität Tübingen, Hoppe-Seyler-Straße 4, 72076 Tübingen, Germany; [3]Laboratory of Protein Biochemistry and Protein Engineering, Ghent University, K.L. Ledeganckstraat 35, 9000 Ghent, Belgium; [4]Anorganische Biochemie, Physiologisch-chemisches Institut der Universität Tübingen, Hoppe-Seyler-Straße 4, 72076 Tübingen, Germany; [5]Department of Biology and CNR Institute for Biomedical Technologies, Section of Padova, University of Padova, Via Ugo Bassi 58/B, I-35131 Padova, Italy; and [6]Department of Immunology, Institute for Cell Biology, University of Tübingen, Auf der Morgenstelle 15, D-72076 Tübingen, Germany

Received November 5, 2004; accepted June 13, 2005

Arthropodan hemocyanins are giant respiratory proteins responsible for oxygen transport. They exhibit unusual assemblies of up to 48 structural subunits. Hemocyanin from *Carcinus aestuarii* contains three major and two minor structural subunits. Here, we reveal the primary structure of the γ-type 75 kDa subunit of *Carcinus aestuarii* hemocyanin, CaeSS2, and combine structure-based sequence alignments, tryptophan fluorescence, and glycosylation analyses to provide insights into the structural and functional organisation of CaeSS2. We identify three functional domains and three conserved histidine residues that most likely participate in the formation of the copper active site in domain 2. Oxygen-binding ability of *Carcinus aestuarii* Hc and its structural subunit 2 was studied using CD and fluorescence spectroscopy. Removing the copper dioxygen system from the active site led to a decrease of the melting temperature, which can be explained by a stabilizing effect of the binding metal ion. To study the quenching effect of the active site copper ions in hemocyanins, the copper complex $Cu^{II}(PuPhPy)^{2+}$ was used, which appears as a very strong quencher of the tryptophan emission. Furthermore, the structural localization was clarified and found to explain the observed fluorescence behavior of the protein. Sugar analysis reveals that CaeSS2 is glycosylated, and oligosaccharide chains connected to three *O*-glycosylated and one *N*-glycosylated sites were found.

Key words: *Carcinus aestuarii*, Crustacea, glycosylated sites, hemocyanins, primary structure, tryptophan distribution.

Abbreviations: *Asn-Gly-Ser*, Asparagine-Glycin-Serin; CaeSS2, *Carcinus aestuarii* structural subunit 2; CFA, Complete Freund's adjuvant; Hc, hemocyanin; IFA, Incomplete Freund's aduvant; LPS, *Escherichia coli* lipopolysaccharide; Man, D-mannose; GlcNAc, *N*-acetyl-D-glucosamine; PNGase-F, peptide *N*4-(*N*-acetyl-b-glucosaminyl) asparagine amidase F; MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry; SI, stimulation indexes; Trp, tryptophan; Tyr, tyrosine.

Hemocyanins (Hcs) are large copper-containing respiratory proteins that act as the predominant oxygen carriers found in several species of molluscs and arthropods (*1*, *2*). They have been reported to contribute to 90% of the total hemolymph protein content. The properties of these large extracellular oligomers, found in three classes of Arthropoda (Crustacea, Chelicerata, and Myriapoda), have been extensively studied for many years, as was done for the structurally distinct, but functionally similar molluscan hemocyanins (*3–5*). They exist as 16S multimeric species that can be organized as higher aggregation forms in arthropodan Hcs: 24S dimers (or 2 × 6-mers), 37S tetramers (or 4 × 6-mers), and 62S octamers (or 8 × 6-mers) which are stabilized by Ca^{2+}. A specific aggregation form is typical for each species of the phylum. The higher oligomers can be reversibly dissociated under mild conditions into the 16S species (M_r < 450,000, 6 × 5S subunits with M_r < 75,000) by removing Ca^{2+} at neutral pH. The 5S subunit consists of a single polypeptide chain of about 660 residues folded into three structural domains (*6*, *7*). The subunits contain one oxygen-binding site each and are arranged as a trigonal antiprism (*5*, *8*, *9*). Different types of 5S subunits exist in the various Hc species that differ from each other in their association–dissociation behavior. The dissociation, however, is not always reversible.

The hexamer is the predominant form in most primitive crustacean Decapoda, such as *Penaeus setiferus* (*10*) and *Penalus monodon* (*11*), and consists of three subunits (a, b and c subunits) (*5*, *8*). Thus far, the primary structures of several hemocyanin subunits from chelicerata and crustacea have been determined (*12–23*). In addition, detailed three-dimensional structures of hemo-

*To whom correspondence should be addressed. Tel. +359-2-9606163, Fax: 00359-2-8700225, E-mail: pda54@yahoo.com

cyanin subunits from the crustacea *Panulirus interruptus* (*24*, *25*) and the chelicerate *Limulus polyphemus* (*26*) have been established based on X-ray diffraction studies.

In the present communication we focus on the structural subunit CaeSS2 of *Carcinus aestuarii* hemocyanin (previously referred to as *Carcinus maenuas*), for which no primary structure and other structural/biochemical properties are known. Hemocyanin (Hc) of *C. aestuarii* contains three major and two minor electrophoretically separable polypeptide chains with different N-terminal amino acid sequences (*27*). Usually, the reconstitution of the native aggregates requires the full complement of subunits. However, in some instances (*i.e.*, *C. aestuarii* Hc) the higher aggregation forms (24S dodecamer or above) are scarcely populated in reassociation experiments (*28*). Studies of the stability of the native *C. aestuarii* Hc dodecamer towards various denaturants (temperature and guanidinium hydrochloride), using different techniques, indicate that the quaternary structure is stabilized by oligomerization between structural subunits, and that the sugar moieties may have a structural role (*29–31*). Under appropriate experimental conditions CaeSS1 and CaeSS3 are able to reassociate to the hexameric form (*28*), while CaeSS2, which has been found to be glycosylated, is unable to reassociate after its separation from the subunit pool. It retains its monomeric state also at neutral pH and in the presence of Ca^{2+}.

Here we report the complete primary structure of this structural subunit of crustacean *C. aestuarii* hemocyanin as deduced from amino acid sequencing and matrix-assisted laser desorption/ionization mass spectrometry. We also report structure-based sequence alignments with homologous proteins, tryptophan fluorescence experiments, and analyses of the glycosylation content of CaeSS2, which were performed in an attempt to gain insights into the structure–function relationships of hemocyanins.

MATERIALS AND METHODS

Isolation of C. aestuarii Hc and Structural Subunit CaeSS2—Native Hc from the crab *C. aestuarii* was prepared from the hemolymph obtained from the dorsal lacuna of living animals collected in the lagoon of Venice. The protein was stored at –20°C in the presence of 18% (w/v) sucrose. Before use, Hc was exhaustively dialysed against 20 mM phosphate buffer containing 10 mM EDTA and 5 mM hydroxylamine hydrochloride at pH 7.5. Protein concentration was determined using the absorption coefficient at 278 nm of E_{278} 0.1% = 1.24 mg^{-1} ml cm^{-1} at pH 7.5 and 20°C. A molecular mass of 75 kDa was assumed for the structural subunit containing one active site. The degree of oxygenation was determined using the absorbance ratio of A_{337}: A_{280} = 0.21 for a preparation containing 100% oxy-Hc. The native protein was dissociated into its subunits by dialysis for 24 h against 100 mM sodium bicarbonate buffer, pH 9.5, containing 20 mM EDTA and 1 M urea. The subunits were isolated by FPLC ion exchange chromatography using a Resource column eluted with 50 mM Tris/HCl buffer, 1 M urea and a non-linear gradient of 0.0 –0.5 M NaCl.

Modification of CaeSS2 by Reduction and S-Pyridylethylation—Before each enzymatic cleavage, portions of 5 mg of CaeSS2 were dissolved in 3.0 ml of 0.25 M Tris-HCl buffer, pH 8.5, 6 M guanidine-HCl, 1 mM EDTA. An ethanolic solution of 2-mercaptoethanol (10% v/v in water, at 100-fold molar excess over the cysteine residues) was added. The mixture was incubated under nitrogen for 2 h at room temperature in the dark. Neat 4-vinylpyridine (100-fold molar excess over the expected cysteine residues) was added, and the mixture incubated under nitrogen for 2 h at room temperature in the dark. The reaction was stopped and the pyridylethylated protein desalted by reverse-phase HPLC on an Aquapore RP-300 column (2.1 × 30 mm; Applied Biosystems, Weiterstadt, Germany). The elution was performed using the following conditions: eluent A, 0.1% TFA in water; eluent B, 0.085% TFA, 80% acetonitrile and 20% water; gradient program: 0% B for 5 min, then 0–100% B in 10 min. A flow rate of 0.5 ml/min was used, and the absorbance of the eluate recorded at 214 nm. The fraction containing the modified protein was collected and recovered by lyophilization.

Enzymatic Hydrolysis of CaeSS2 Structural Subunit—Chymotryptic cleavage of the *S*-pyridylethylated protein was performed at 37°C for 50 min by the addition of chymotrypsin to a final concentration ratio of 1:40 (w/w) in 100 mM ammonium bicarbonate, pH 7.8. The chymotryptic digest was fractionated *via* HPLC on a LiChrospher 60 RP column (250 × 4 mm; Merck, Darmstadt, Germany) by elution with a mixture of water and acetonitrile (eluent A, 0.1% trifluoroacetic acid in water; eluent B, 80% acetonitrile in 0.1% trifluoroacetic acid/water), using a linear concentration gradient from 5 to 100% B in 100 min at a flow rate of 1.0 ml/min. The UV absorbance of the eluate was monitored at 214 nm.

One milligram of modified CaeSS2 was dissolved in 1 ml of 5 mM ammonium bicarbonate buffer, pH 8.2, and incubated with 20 μl of trypsin solution (1 mg ml^{-1}) at room temperature for 15 h, followed by a further addition of 20 μl of trypsin solution. Then the reaction mixture (enzyme:protein ratio of 1:30, w/w) was incubated overnight at 37°C. The generated peptides were separated by reverse phase HPLC on a Nucleosil 100 RP-18 column (250 × 10 mm; Macherey-Nagel, Germany) by applying the following gradient: from 5% B to 90% buffer B within 110 min and a flow rate of 1.0 ml min^{-1} (eluent A, 0.1% trifluoroacetic acid in water; eluent B, 80% acetonitrile in 0.1% trifluoroacetic acid/water).

Mass Spectrometric Analysis—Mass spectrometric analysis of the HPLC fractions containing the peptide fragments resulting from the cleavage was done by MALDI-MS (Voyager, PerSeptive Biosystems, Wiesbaden, Germany). Peptides (10–50 pmol) were dissolved in 0.1% (v/v) TFA and applied to the target. Analysis was carried out using α-cyano-4-hydroxycinnamic acid or 2,5-dihydroxybenzoic acid as a matrix. Solutions of human substance P (1,347.7 Da) and bovine insulin (5,733.6 Da) were used to calibrate the mass scale. The mass values assigned to the amino acid residues are the average masses.

Amino Acid Sequence Determination—Peak fractions were dried, dissolved in 40% methanol/1% formic acid, and subjected to automated Edman N-terminal sequencing

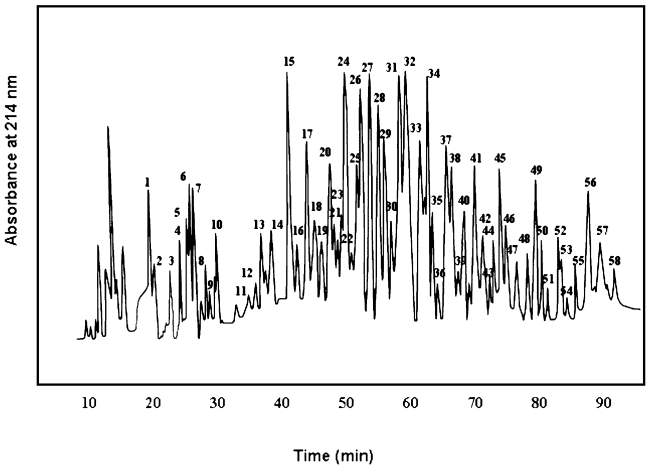


Fig. 1. **HPLC separation of tryptic fragments of CaeSS2.** Column: Nucleosil 100 RP-18 (250 × 3 10 mm; Macherey-Nagel, Germany); solvents: A, 0.1% trifluoroacetic acid in water, and B, 80% acetonitrile in 0.1% trifluoroacetic acid/water; gradient: from 5% B to 90% buffer B in 110 min; flow rate 1.0 ml min^{-1}; detection: UV, λ = 206 nm.

(Procise 494A Pulsed Liquid Protein Sequencer, Applied Biosystems GmbH, Weiterstadt, Germany).

Spectroscopic Measurements—CD measurements were performed with a Jasco J-720 dichrograph and the far UV CD spectra were recorded between 200 and 250 nm at apo- and oxy-forms of protein solutions of 0.25 mg ml^{-1} in 50 mM Tris/HCl buffer, pH 8.9 and cuvettes of 0.1 cm.

Fluorescence Spectroscopy—The apo-form (copper-deprived) of native protein and dissociation products was obtained by overnight dialysis against 25 mM KCN in 100 mM Tris/HCl, pH 8.0, at 4°C. The proteins were then dialyzed against the same buffer without KCN containing 10 mM EDTA, and finally against 100 mM Tris/HCl, pH 7.0.

Fluorescence spectra were recorded with a Perkin Elmer LS 5 spectrofluorimeter. Protein solutions had an absorbance at excitation wavelengths lower than 0.05 to minimize the inner filter or self-absorption effects. The relative quantum yields (Φ) were measured by comparing the integrated corrected fluorescence emission spectra of Hcs with those of *N*-acetyltryptophanamide, normalized to the same absorbance at the excitation wavelength (295 nm). The quantum yield of the standard was 0.13 at 21°C (*32*).

Fluorescence quenching experiments were performed with a copper complex [Cu^{II}(PuPhPy)^2] (*33*) as an external quencher in the concentration range from 1×10^{-7} M to 20×10^{-7} M. Emission spectra were recorded at a speed of 60 nm min^{-1}, and emission slit widths were adjusted to 5.0 nm band pass.

The results of the quenching reactions between the excited tryptophyl side chains and acrylamide were analyzed according to the Stern-Volmer equation (*32*):

$$F_0/F = 1 + K_{sv}[X]$$

where F_0 and F are the fluorescence intensities at an appropriate emission wavelength in the absence and presence of quencher, K_{sv} is the dynamic quenching constant and [X] the quencher concentration. The inner filter effect due to acrylamide was corrected by the factor:

$$Y = \text{antilog}\ (Abs_{exc} + Abs_{emiss})/2$$

where Abs_{exc} and Abs_{emiss} are the absorbance at the excitation and emission wavelength respectively.

$$K_{sv} = K_q\tau$$

where K_q is the quenching rate constant and τ is the fluorescence lifetime. The quenching rate constant K_q, which is the rate constant for diffusional collision of the quencher with the tryptophyl side chains, was obtained by dividing the apparent K_{sv} values by the longer lifetime.

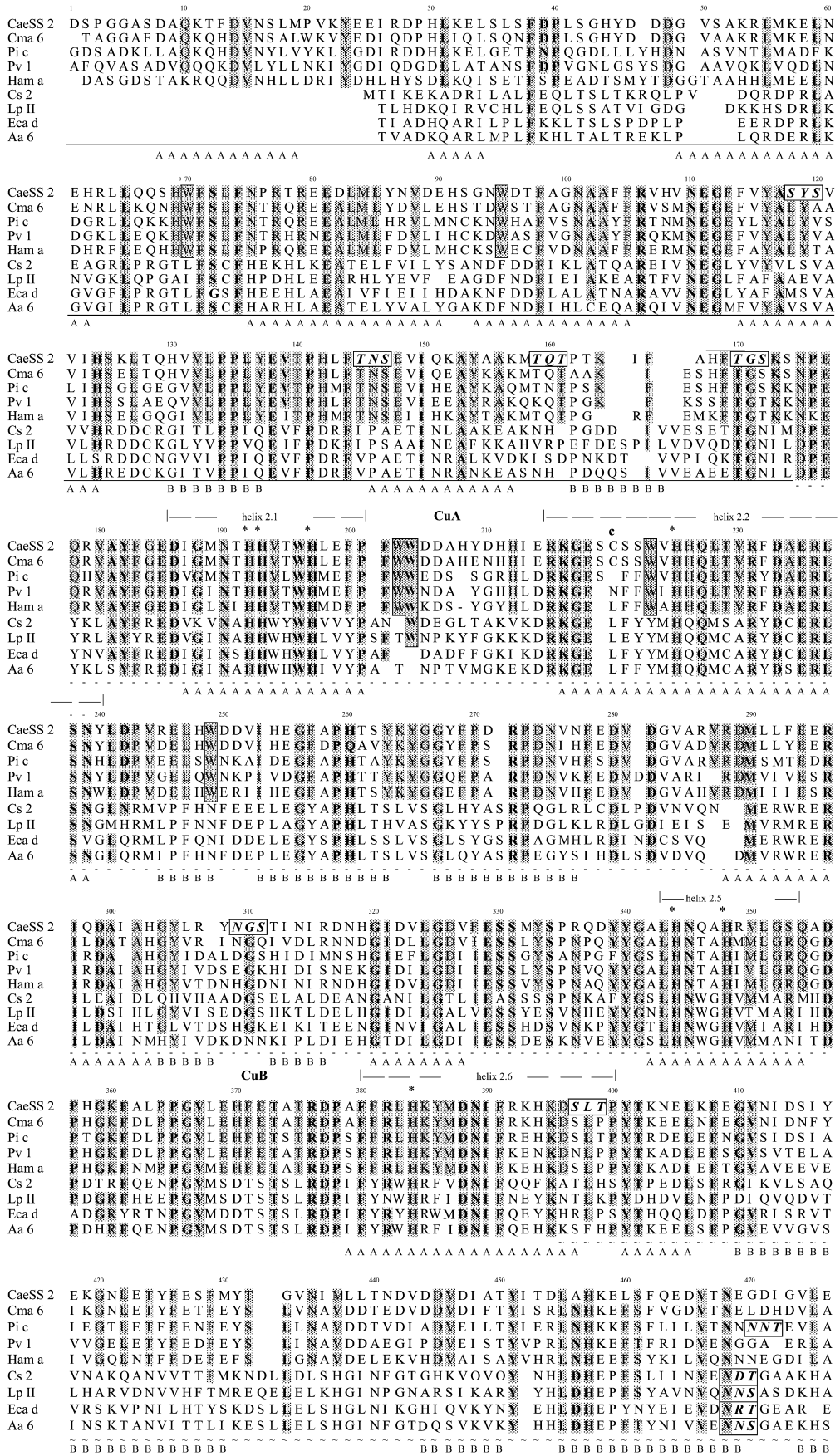
The accuracy in the excited state lifetime (τ determination was ± 0.2 ns. The decay curves contained 10^4 counts at the maxima. The time interval for these curves was 100 ps per channel.

RESULTS AND DISCUSSION

Fragmentation and Purification of Subunit CaeSS2 Peptides—The amino acid sequence of CaeSS2 was determined by automated Edman degradation and MALDI-MS analysis of protein fragments obtained by cleavage with chymotrypsin and trypsin. After reduction with 2-mercaptoethanol, modification with 4-vinylpyridine, and cleavage with chymotrypsin, the generated peptides/glycopeptides were fractionated on a Superdex peptide HR 10/30 column, yielding two fractions. After fractionation of the peptide mixture in these two fractions *via* RP-HPLC (Fig. 1), the peptides were analyzed by mass spectrometry and by sequenced. Similarly, enzymatic cleavage of subunit CaeSS2 was performed with TPCK-trypsin. The resulting peptides were separated by RP-HPLC, providing a well-resolved typical peptide map of this subunit. Although most of the isolated fractions contained pure individual peptides, mixtures of peptides were obtained in a few cases. About 40–45 HPLC fractions per digest were sequenced and used for alignment and overlap. Approximately 95% of the amino acid residues in the CaeSS2 structure were identified at least twice by sequencing different fragments. Occasionally, non-specific cleavages were observed. The sequences of tryptic peptides obtained by a combination of Edman degradation and MALDI-MS are summarized in Fig. 2 in comparison with sequences of known Hcs.

Primary Structure of Subunit CaeSS2—The primary structure including all structurally and functionally important sites of crab *Carcinus aestuarii* Hc subunit CaeSS2 was established by comparison with peptide sequences of the phylogenetically related *Cancer magister* Hc subunit 6 (*20*) (Fig. 2). The subunit is a polypeptide of 650 amino acids with a calculated molecular mass of 74,870 Da, which agrees with the results obtained by SDS-PAGE. The subunit was calculated to be acidic, with an isoelectric point [pI] of 5.59. The amino acid sequence shows the presence of high amounts of aromatic amino acids (tryptophan, phenylalanine and tyrosine) and low methionine content.

The complete hemocyanin sequence of subunit CaeSS2 was aligned with other crustacean hemocyanins: Dunge-



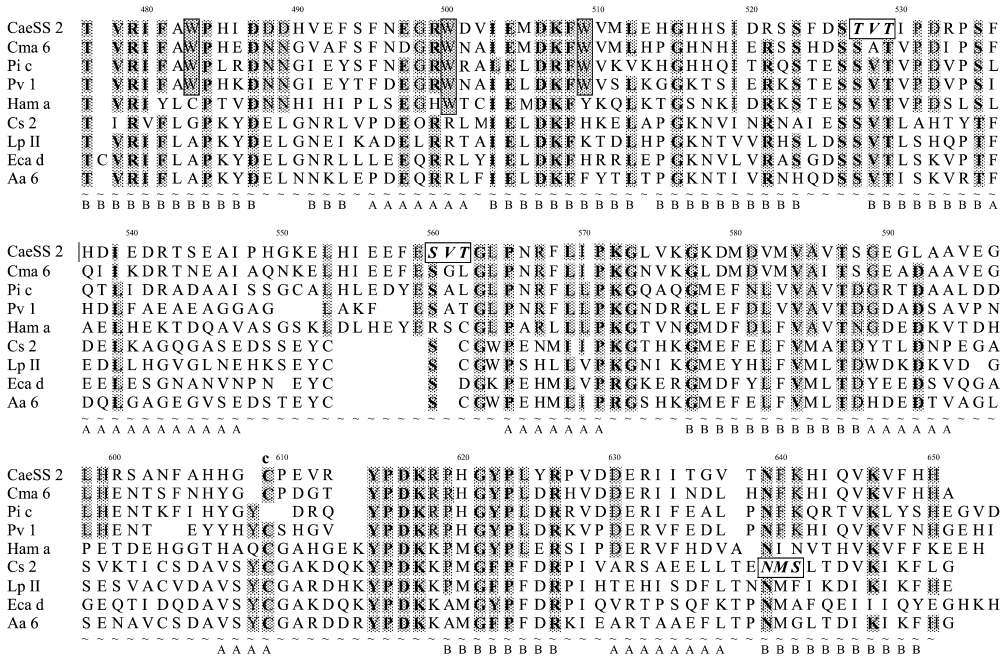


Fig. 2. **Alignment of the amino acid sequences of subunit CaeSS2 of *C. aestuarii* hemocyanin with selected crustacean hemocyanins: γ-type subunits of C. *magister* subunit 6 (Cma, AAA96966); *P. vulgaris* subunit 1 (Pv1, P80888); *P. interruptus* subunit c (Pic, P80096); α-type subunit of *H. americanus* (Ham, AJ272095) and cheliceratan Hcs: *C. salei* subunit 2 (Cs2, AJ307904); *L. polyphemus* subunit II (LpII, P04253); *A. australis* subunit 6 (Aa 6, P80476), *E. californicum* subunit d (Ecd, P02241).** Structural elements are related to *P. interruptus* Hc: domain I (—), domain II (- - -) and domain III (~ ~ ~). The putative *O*- and *N*-glycosylation sites are boxed and in italics; conserved residues are shaded; copper-binding histidine residues (*), disulfide bridges (c), α-helices (A) and β-sheets (B) are marked correspondingly.

ness crab *Cancer magister* subunit 6 (Cm a, Accession No. AAA96966) (73.2% homology) (*20*), *Palinurus vulgaris* subunit 1 (Pv1, P80888) (53.2%) (*18*), *Palinurus interruptus* subunit c (Pic, P80096) (56%) (*22*), *Homarus americanus* (Ham, AJ272095) (46.8%) (*21*), and cheliceratan Hcs: *Cupiennius salei* subunit 2 (Cs2, AJ307904) (31.1%) (*13*), *Limulus polyphemus* subunit II (LpII, P04253) (30%) (*34*), *Androctonus australis* subunit 6 (Aa6, P80476) (32%) (*19*), *Eurypelma californicum* subunit d (Ec d, P02241) (32%) (*35*).

Multiple sequence alignment of these primary structures reveals that the polypeptide chain lengths of many crustacean Hcs are roughly the same [CaeSS2 (650 aa), *H. americanis* a (654 aa), *C. magister* 6 (650 aa), *P. vulgaris* 1 (657 aa) and *P. interuptus* c (661 aa)], while the cheliceratan Hcs spider *E. californicum* subunit Ecal-a (630 aa), the scorpion *A. australis* subunit (626 aa) and the horseshoe crab *L. polyphemus* subunit LpII (628 aa) are also similar but shorter (*34–36*). Furthermore, the high degree of sequence similarity among arthropodan hemocyanins (30–70% sequence identity) suggests that the proteins have a common tertiary structure.

As expected, we observed a high degree of sequence identity between CaeSS2 and other crustacean hemocyanins, in particular to *P. vulgaris* subunit 1 (52.5% identity) and *P. interruptus* subunit c (56% identity). These three subunits belong to the γ-type hemocyanins according to the definition of Markl (*6*). CaeSS2 and *Cancer magister* subunit 6 seem to be more homologous to each other (73% identity) (Fig. 2). This value is very close to the homology reported between the subunits of *H. americanus* A and the hemocyanin subunits A and B of spiny lobster *P. interruptus* (70.2 and 70.4% identity respectively) and that of *P. vulgaris* (69% identity).

Crystallographic studies of hemocyanin from *Panulirus* and *Limulus* (*24–26*, *37*) have shown that arthropodan hemocyanins consist of three structural domains. A structure-based sequence alignment of the CaeSS2 amino acid sequence against homologous hemocyanins for which structural information is known suggests a rather strict conservation of structural features (Fig. 2). Domain 1 (residues 1–175 in CaeSS2) is mainly α-helical and is quite variable in sequence. Domain 2 (residues 175–395*)* is the most conserved and is sandwiched between domains 1 and 3. It is also mostly a-helical and contains the oxygen-binding Cu_A and Cu_B sites. Each copper ion is buried in the core of this domain and is ligated to three histidine side chains. The six histidines marked by an asterisk in Fig. 2 are implicated in Cu binding according to their highly conserved character among arthropodan hemocyanin subunits (*7*, *35*). The Cu_A helix pair in CaeSS2 extends from residue 187 to residue 201 (helix 2.1) and from residue 216 to residue 240 (helix 2.2). The Cu binding histidines are located at positions 193, 197, and 225. The Cu_B helix pair extends in the region 342 to 354 (helix 2.5) and 375 to 397 (helix 2.6). Histidine residues binding the copper ion at the Cu_B site are located at positions 344, 348, and 384. Domain 3 is rich in β-sheets and forms a β-barrel structure: it also contains a seven-stranded Greek-key motif with two long loops. One of these loops contains a disulfide bound that bridges and interacts with domain 2 and a Ca^{2+} binding site.

Two cysteine residues were identified in the crustacean Hcs CaeSS2, Cm6, Pic, and Pv1, while 4 or 6 of these

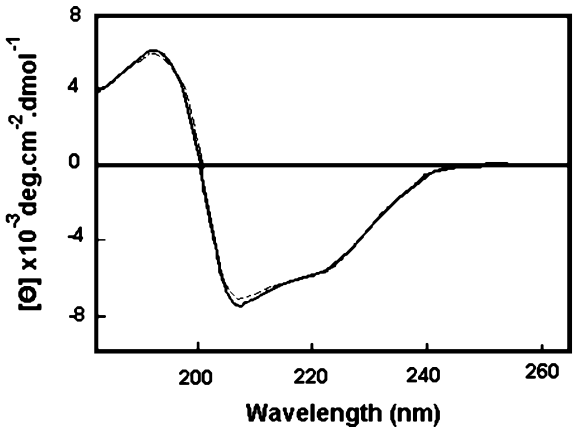


Fig. 3. **Circular dichroism spectra in the far-UV region of the structural subunit *Cae*SS2 in the holo- (dashed line) and apo-form (solid line).** All spectra were taken at 25°C in 50 mM Tris-HCl buffer, pH 7.5, protein concentration of 0.2 mg/ml, using 0.1 cm pathlength quartz cell.

residues were observed in chelicerata Hcs Cs2, LpII, Ecd, and Aa6. The cysteines forming the disulfide bridge and stabilizing domain 3 are conserved in CaeSS2 and in *C. magister* subunit 6 at positions 220 and 609, respectively, but are different from the cysteines forming disulfide bridges in LpII and in *E. californicum* subunits. The disulfide bridge was proven using DTT as a reducing agent for -S-S- bridges and applying CD spectroscopy. The CD spectrum of native CÑpÑuSS2 shows a typical negative Cotton effect at 208 and 221 nm. Loss of the native conformation occurred after disulfide reduction by DTT indicating that potential disulfide bridges indeed exist in CaeSS2. There are about 8 Cys residues in each subunit of *C. salei* hemocyanin, where four strictly conserved cysteine residues are present in domain 3. They most likely form two disulfide bridges that make up a flexible hinge stabilizing the three-dimensional structure of the subunit, as deduced from other hemocyanins (*13*).

As the amino acid sequence of CaeSS2 shows a rather high content of aromatic amino acids (tryptophan, phenylalanine and tyrosine), fluorescence spectroscopy is a useful tool to study the stability of this protein.

CD Spectroscopy—Structural characteristics of apo- and oxy-Hc and the subunit CaeSS2 were studied using CD spectroscopy (Fig. 3). The CD spectra in 50 mM Tris-HCl buffer, pH 7.5 of native and copper-free CaeSS2 are very similar. Only a small change in the secondary structure was observed. Analysis of the spectrum of apo-CaeSS2 yields the following average values of secondary structure components: α-helix 16.0%; β-sheet 30.1%; turns 20.0%; random coil 34.0%, which are similar to those of oxy-CaeSS2.

The spectra of the oxy- and apo-forms of the CaeSS2 were recorded in the temperature range of 20–90°C and in the 200–250 nm region, providing information on the backbone conformation of the protein. Sigmoidal curves obtained for the oxy- and apo-CaeSS2 were compared with the whole molecule (*27*). The thermal denaturation of Hcs was irreversible, and it was possible to follow only the forward reaction. For this reason the melting temperature (T_m), the midpoint in the sigmoidal denaturation curve, was used to explain the thermostability of the whole molecule of *Carcinus* Hc and its structural subunit 2. The melting temperature for apo-Hc is 69°C, 3°C lower than that of for the oxy-form (72°C). Removing the copper dioxygen system from the active site in structural subunit CaeSS2 leads to a decrease of the melting temperature by 5°C (58°C and 53°C, respectively for the oxy- and apo-form) (*27*), which shows that the copper ions bound in the binuclear active site of whole molecule and CaeSS2 play a stabilizing role on the tertiary structure of the protein. Upon removal of copper, the monomeric hemocyanin undergoes changes at the level of tertiary structure, while the secondary structure is mostly unaffected.

Fluorescence Properties—Our determination of the amino acid sequence of CaeSS2 identified 10 Trp and 27 Tyr residues. Upon excitation at 295 nm, the native oxy-Hc and oxy-SSs show an emission band with λ_{max} at 310–336 nm (Table 1), indicating the presence of "buried" tryptophan side chains. Indeed, analysis of the fluorescence decay (*38*) revealed that Trp residues of the CaeSS2 subunit can be classified into three classes, with fluorescence lifetimes of around 0.11–0.15, 0.33, and 3.1–3.5 ns, respectively. The short-lived component is mainly responsible for the decay of the holo-form. Its contribution to the overall fluorescence (around 70%) corresponds to that calculated as the most acrylamide-accessible fluorescence. The second class of Trp residues, with an intermediate lifetime (0.33 ns), is present only after copper removal. It can be identified with residues localized in the close surrounding of the active site, whose fluorescence is fully quenched by copper-related heavy atom and paramagnetic ion effects in the oxygenated form (*39*, *40*). To identify the contribution of Trp residues in CaeSS2, the structure of the *P. interruptus* subunit a was used to construct a tentative model of the structural subunit of *C. aestuarii* hemocyanin. Figure 4 shows the Trp distribution in subunit *CaeSS2*, based on this model. The Trp positions, together with the decay parameters, can give useful information on the fluorescence properties of these

Table 1. **Quenching constants (K_{sv}) of monomeric CaeSS subunits, native *C. aestuarii* Hc and the copper complex $Cu^{II}(PuPhPy)^{2+}$ in the presence of acrylamide and iodide ions.**

Sample	KI (K_{SV}^-) ($dm^3\ mol^{-1}$)		Acrylamide (K_{SV}) ($dm^3\ mol^{-1}$)		Copper complex (K_{SV}) ($dm^3\ mol^{-1}$)		Emission λ_{max} (nm) Excitation at 295 nm	
	oxy-	apo-	oxy-	apo-	oxy-	apo-	oxy-	apo-
Native Hc	–	–	–	3.64[a]			310 ± 1[a]	322 ± 1[a]
*Cae*SS1	1.0[b]	1.0[b]	2.2[b]	4.0[b]			329 ± 1[a]	331 ± 1[a]
*Cae*SS2	1.0[b]	1.0[b]	2.4[b]	3.5[b]	0.9×10^6	1.4×10^6	336 ± 1[a]	339 ± 1[a]
CaeSS3	1.0[b]	1.0[b]	0.9[b]	10.0[b]			332 ± 1[a]	335 ± 1[a]

[a]Ref. *27*; [b]Ref. *38*.

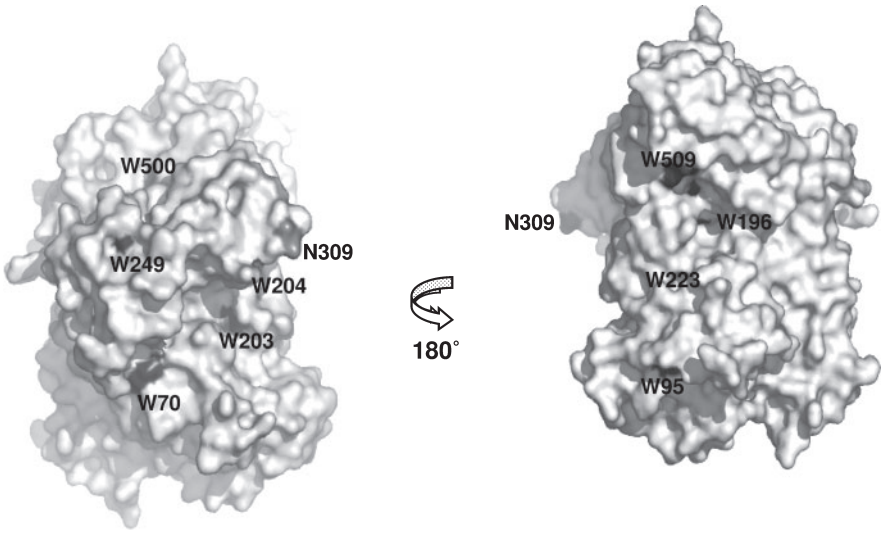


Fig. 4. **Putative localization of tryptophan residues and *N*-linked glycosylation site of CaeSS2.** Tryptophan residues are indicated in red. Trp483 is not shown because it is predicted to be completely buried in the structure. The *N*-linked glycosylation site (Asn309) is shown in blue. The model was generated using the structure of subunit 1 of *P. interruptus* hemocyanin (PDB code 1HC1) following a structure-based sequence alignment of the sequence of CaeSS2 against the sequence and crystal structure of *P. interruptus* hemocyanin. This figure was prepared and rendered with PyMol 0.95.

residues individually. The quenching of tryptophan fluorescence by copper ions (see above) is a very short-range process involving quenching interactions within chromophore-quencher distances of about 14 Å (*41*). As was calculated from the model, the distances of Trp196, 203, 223 and 249 are within this range (Fig. 5) and they should constitute the class of fluorophores, characterized by a τ value of about 0.3 ns, which become fluorimetrically active only after copper removal. Most of the tryptphans located near the active site are conserved in Hcs from other arthropods such as *Cancer magister* Hc subunit Cm6, *Palinurus vulgaris* subunit 1, *Palinurus interruptus* subunit c and *Homarus americanus* (Fig. 2). The third class of Trp residues with high τ values (3.0–3.5 ns) represents amino acids close to the active site. Trp196, with a distance to the active site of 8–11 Å, most probably belongs to this class. In most arthropodan Hcs only 50% of Trp residues are located in the vicinity of the metal centres (*8*). Most of the indole residues of structural subunit a of *E. californicus* Hc are located within a short distance (less than 1.1 nm) from copper A and B, which explains the exceptionally strong fluorescence quenching in the oxy-form.

As is seen from the model of CaeSS2 (Fig. 4), the emission of holo-CaeSS2 mainly arises from Trp70, 95, 500, and Trp509, which are exposed or are near the surface of the molecule. They belong to the first class of Trp residues having a very short fluorescence lifetime of 0.15 ns. Trp483 is about 19–20 Å removed from the copper ions, but is buried by the Cys residues and cannot be observed at the surface of the molecule. It can be concluded that the fluorescence emission of both oxy- and apo-forms of the CaeSS2 is dominated by tryptophans located in the hydrophobic core of the proteins and inaccessible to the solvent.

The static fluorescence parameters agreed fairly well with the data obtained by quenching experiments. Quenching experiments with neutral and ionic quenchers further supported this conclusion. Fluorescence quenching reactions have been widely used for studying the degree of exposure and electronic environment of aromatic amino acid residues. The quenching of the indole fluorescence of CaeSS1, CaeSS2, CaeSS3 and 16 S hexamer of *C. aestuarii* Hc is shown in Table 1, together with those obtained for Ac-Trp-NH_2 as a model for the completely solvent-exposed fluorophores. The decrease in fluorescence emission is a linear function of the concentration of quenching agent, *i.e.*, the quenching process follows the unmodified Stern-Volmer equation and can be described using single K_{sv} constants. The variation of the slopes (K_{sv}) of the plots reflects the different overall accessibility of the intrinsic tryptophan residues to solvent. The quenching efficiency K_{sv} of the oxy-form is lower than that of the apo-form. Removal of the copper ions from the active site increases the fluorescence intensity. The removal of the copper–peroxide complex causes structural rearrangement of the microenvironment of the fluorophores. As a result, this residue or residues probably become more inaccessible to the solvent. This conclusion is further confirmed by the values of the respective

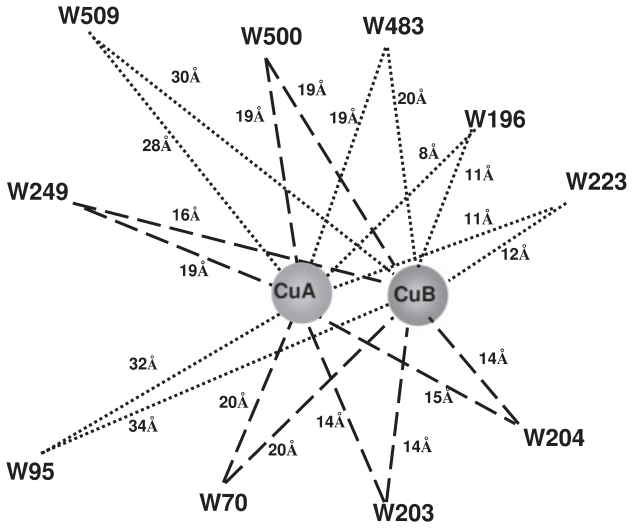


Fig. 5. **Approximate distances between tryptophan residues in *C. aestuarii* and the di-copper center.** The distances indicated are $Trp_{C\alpha}$-Cu values observed in the homology model that was used.

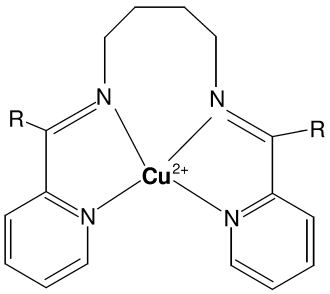
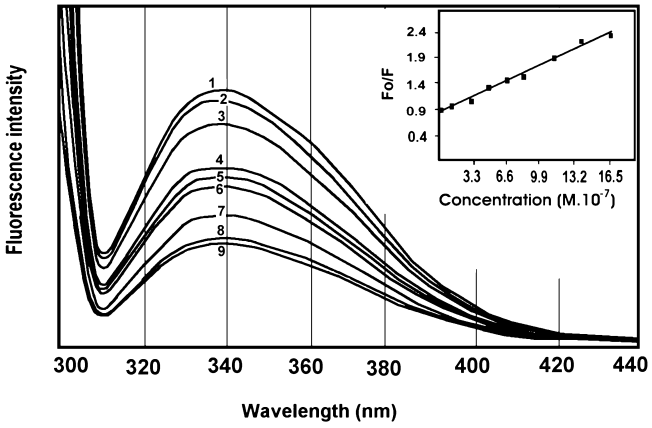


Fig. 6. **Copper complex [Cu^{II}(PuPhPy)$^{2+}$] including one Cu ion.**

Stern-Volmer constants (Table 1). The iodide ion (I^-) is able to quench only surface fluorophores and is effective in discriminating between "exposed" and "buried" chromophores, as well as in revealing charge effects. Considerable increase in the quenching efficiency is observed when iodide is exposed to different concentrations of acrylamide. Acrylamide is an efficient neutral quencher of tryptophyl fluorescence and provides topographical information about the emitting chromophores. The ability to quench collisionally the excited indole rings depends on its ability to penetrate the protein matrix. Acrylamide can discriminate between "exposed" and "buried" tryptophyl side chains, and the results are not influenced by the charge of the chromophore microenvironment. Different quenching constants (K_{sv}) of 3.64 (M^{-1}) for the native molecule and 4.0, 3.5 and 10.0 (M^{-1}) for the structural subunits CaeSS1, CaeSS2 and CaeSS3, respectively, were calculated (*27*, *38*) (Table 1) which are about 6 times less than for the copper complex [Cu ^{II}(PuPhPy)^2] (Fig. 6).

The copper complex [Cu ^{II}(PuPhPy)$^{2+}$], which includes one Cu ion in its center, was used as a quencher to study the effect of copper ions (Fig. 7). Figure 7, A and B, (inset) shows very high values of Stern-Volmer constants: 0.90×10^6 (M^{-1}) for oxy and 1.45×10^6 (M^{-1}) for apo-CaeSS2, respectively (Table 1). The difference of K_{sv} values calculated from holo- and apo-forms relates to the quenching effect on the indole emission of copper ions in the active site.

Oligosaccharide Composition—It has recently been reported that the hemocyanin of the crab *Carcinus aestuarii* contains carbohydrate moieties corresponding to 1.6% of the protein mass (*42*). This carbohydrate content is higher than that exhibited by other arthropodan hemocyanins investigated so far. Sugar analysis of the different subunits reveals that the subunit referred to as CaeSS2 is glycosylated, with a carbohydrate content of 6.3%. Our sequence results show three consensus sequences for *O*-glycosylation and one for *N*-glycosylation (Fig. 2). The putative *N*-linked site is observed at position 309 with a consensus sequence *Asn-Gly-Ser*, typical for *N*-glycosylation. From the mass difference between the experimental and the calculated value (based on the amino acid sequence of fragment), we found that a carbohydrate chain with a molecular mass of 1,348 Da and a potential $SO_4Man_4GlcNAc_3$ structure is connected to this site. One potential *N*-glycosylation site (*-Asp-Val-Thr-*) was observed at the C-terminal position 639 in *Cupiennius salei* subunit 2, which is, however, not conserved in other crustacean hemocyanin subunits. Four potential *N*-glycosylation sites are present in each of the *H. americanus* pseudohemocyanins[21]. One *N*-linkage site was identified at position 166 in the α-type subunits of *P. elephas* subunits 1, 2 and 3 (*23*), in *P. interruptus* subunits a and b, and at position 470 in β-type subunit c of *P. interruptus*. Several putative *N*-glycosylation sites [*-Asn-X-(Thr/Ser)-*] were observed in the primary structure of all seven subunits of *E. californicus* Hc, but they do not pass through the Golgi apparatus and no carbohydrate moiety was detected in the native tarantula hemocyanin (*35*).



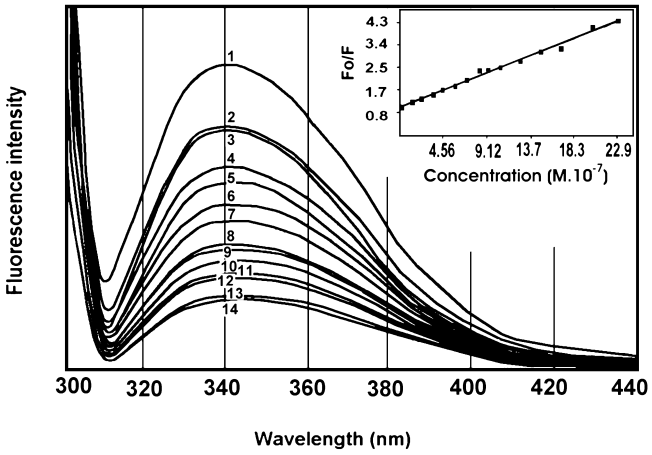


Fig. 7. **Fluorescence quenching of *Carcinus aestuarii* structural subunit 2 with copper complex [Cu^{II}(PuPhPy)$^{2+}$], in the concentrations range from 1×10^{-7} M to 20×10^{-7} M.** Emission spectra (1 to 9) with different concentration of the complex were recorded at a speed of 60 nm min^{-1}, and emission slit widths were adjusted to 5.0 nm band pasts. Fluorescence quenching data were analyzed according to the Stern-Volmer plots describing the quenching of oxygenated (A) and apo-form (B) of CaeSS2 by the copper complex. F_o is the emission without quencher, and F is the emission at different concentrations of the complex. The monomer was dissolved in 50 mM Tris/HCl buffer, pH 8.2.

The carbohydrate compositions of fractions 1–3 from CaeSS2 have been reported (*42*), but the glycosylation sites were not identified. Knowledge of the full amino acid sequence of this structural subunit now allows their identification. Though there exist several putative sites, the sequence *-Thr-Gly-Ser-* from glycopeptide 1 fits position 170–172 of the protein sequence. Therefore, the car-

bohydrate chain 1 with a molecular mass of 404 Da and the suggested structure of two GalNAc groups is *O*-linked to this site. The sequence -*Ser-Tyr-Ser*- of the second peptide (*40*) has an oligosaccharide attached with the suggested structure *N*-Acetyl-*O*-$NeuAc_2Gal_3GalNAc_2$ and a MW of 1525 Da and fits position 118–120. The position of the third *O*-linked carbohydrate chain (*42*) with the sequence *N*-Acetyl-*O*-di-$NeuAc_2Gal_2GalNAc_2$ and a molecular mass of 1,466 Da could not be identified.

Data deposition: The sequences reported in this paper have been deposited in the Swiss-Prot database: P84293. We express our gratitude to the DLR, Internationals Büro, BMBF, Bonn, Germany for financial support by Research Grant no. BGR 001/01. This work was supported by grants from North Atlantic Treaty Organization (NATO) (LST.CLG.978560) and from the NCSI (X-1202) of the Ministry of Education and Science, Bulgaria. P. D.-A. thanks the Belgian Office for Scientific, Technical and Cultural Affairs for a research fellowship (B/04530/01).

REFERENCES

1. Salvato, B. and Beltramini, M., (1990) Hemocyanins: molecular architecture, structure and reactivity of the binuclear copper active site. *Life. Chem. Rep.* **8**, 1–47
2. Meissner, U., Dube, P., Harris, J.R., Stark, H., and Markl, J. (2000) Structure of a molluscan hemocyanin didecamer (HtH1 from *Haliotis tuberculata*) at 12 Å resolution by cryoelectron microscopy. *J. Mol. Biol.* **298**, 21–34
3. Jaenicke, E., Decker, H., Gebauer, W., Markl, J., and Burmester, T. (1999) Identification, structure and properties of hemocyanins from diplopod Myriapoda. *J. Biol. Chem.* **274**, 29071–29074
4. Sanchez, D., Ganfornina, M.D., Gutierrez, G., and Bastani, M.J. (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
5. Markl, J. (1986) Evolution and function of structurally diverse subunits in the respiratory protein hemocyanin from arthropods. *Biol. Bull.* **171**, 90–115
6. Markl, J. and Decker, H. (1992) Molecular structure of the arthropod hemocyanins. *Adv. Comp. Environ. Physiol.* **13**, 325–376
7. Van Holde, K.E. and Miller, K.I. (1995). Hemocyanins. *Adv. Protein Chem.* **47**, 1–81
8. Linzen, B., Soeter, N.M., Riggs, A.F., Schneider, H.J., Schartau, W., Moore, M.D., Behrens, P.Q., Nakashima, H., Takagi, T., Nemoto, T., Vereijken, J.M., Bak, H.J., Beintema, J.J., Volbeda, A., Gaykema, W.P.J., and Hol, W.G.J. (1985) The structure of arthropod hemocyanins. *Science* **229**, 519–524
9. Burmester, T. (2001). Molecular evolution of the arthropod hemocyanin superfamily. *Mol. Biol. Evol.* **18**, 184–195
10. Brouwer, M., Bonaventura, C., and Bonaventura, J. (1978) Analysis of the effect of three different allosteric ligands on oxygen binding by hemocyanin of the shrimp, Penaeus setiferus. *Biochemistry* **17**, 2148–2154
11. Ellerton, H.D. and Anderson, D.M. (1981) In *Invertebrate Oxygen Binding Proteins: Structure, Active Site and Function* (Lang, J., ed.) pp. 159–170, Marcel Dekker, New York
12. Neuteboom, B., Dokter, W., Van Gijsen, J., Rensink, H., De Vries, J., and Beintema, J.J. (1989) Partial amino acid sequence of a hemocyanin subunit from *Palinurus vulgaris*. *Comp. Biochem. Physiol.* **94B**, 593–597
13. Ballweber, P., Markl, J., and Burmester, T. (2002) Complete hemocyanin subunit sequences of the hunting spider *Cupiennius salei*. *J. Biol. Chem.* **277**, 14451–14457
14. Burmester, T. (2002) Origin and evolution of arthropod hemocyanins and related proteins. *J. Comp. Physiol. B* **172**, 95–117
15. Bak, H.J. and Beintema, J.J. (1987) *Panulirus interruptus* hemocyanin. The elucidation of the complete amino acid sequence of subunit a. *FEBS Lett.* **204**, 141–144
16. Jekel, P.A., Bak, H.J., Soeter, N.M., Vereijken, J.M., and Beintema, J.J. (1988) *Panulirus interruptus* hemocyanin. The amino acid sequence of subunit b and anomalous behaviour of subunits a and b on polyacrylamide gel electrophoresis in the presence of SDS. *Eur. J. Biochem.* **178**, 403–412
17. Kusche, K. and Burmester, T. (2001) Molecular cloning and evolution of lobster hemocyanin. *Biochem. Biophys. Res. Commun.* **282**, 887–892
18. Jekel, P.A., Neuteboom, B., and Beintema, J.J. (1996) Primary structure of hemocyanin from *Palinurus vulgaris*. *Comp. Biochem. Physiol.* **115B**, 243–246
19. Buzy, A., Gagnon, J., Lamy, J., Thibault, P., Forest, E., and Hudry-Clergeon, G. (1995) Complete amino acid sequence of the Aa6 subunit of the scorpion *Androctonus australis* hemocyanin determined by Edman degradation and mass spectrometry. *Eur. J. Biochem.* **233**, 33–101
20. Durstewitz, G. and Terwilliger, N.B. (1997) cDNA cloning of a developmentally regulated hemocyanin subunit in the crustacean Cancer *magister* and phylogenetic analysis of the hemocyanin gene family. *Mol. Biol. Evol.* **14**, 266–276
21. Burmester, T. (1999) Identification, molecular cloning and phylogenetic analysis of a non-respiratory pseudo-hemocyanin of *Homarus americanus*. *J. Biol. Chem.* **274**, 3217–13222
22. Neuteboom, B., Jekel, P.A., and Beintema J.J. (1992) Primary structure of hemocyanin subunit *c* from *Panulirus interruptus*. *Eur. J. Biochem.* **206**, 243–249
23. Meissner, U., Stohr, M., Kusche, K., Burmester, T., Stark, H., Harris, R., Orlova, E.V., and Markl, J. (2003) Quaternary structure of the european spiny lobster (*Palinurus elephas*) 1×6-mer hemocyanin from cryoEM and amino acid sequence data. *J. Mol. Biol.* **325**, 99–109
24. Gaykema, W.P., Volbeda, A., and Hol, W.G. (1986). Structure determination of *Panulirus interruptus* hemocyanin at 3.2 Å resolution. Successful phase extension by sixfold density averaging. *J. Mol. Biol.* **187**, 255–275
25. Volbeda, A. and Hol, W.G. (1989). Crystal structure of hexameric hemocyanin from *Panulirus interruptus* refined at 3.2 Å resolution. *J. Mol. Biol.* **209**, 249–279
26. Hazes, B., Magnus, K.A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K.H., and Hol, W.G. (1993). Crystal structure of deoxygenated *Limulus polyphemus* subunit II hemocyanin at 2.18 Å resolution: clues for a mechanism for allosteric regulation. *Protein Sci.* **2**, 597–619
27. Dolashka-Angelova, P., Hristova, R., Stoeva, S., and Voelter, W. (1999) Spectroscopic properties of *Carcinus aestuarii* hemocyanin and its structural subunits. *Spectrochim. Acta Part A* **55**, 2927–2934
28. Dainese, E., Di Muro P., Beltramini M., Salvato B., and Decker H. (1998) Subunits composition and allosteric control in *Carcinus aestuarii* hemocyanin. *Eur. J. Biochem.* **256**, 350–358
29. Hristova, R., Dolashka-Angelova, P., Gigova, M., and Voelter, W. (2000) Fluorescence properties of the arthropodan hemocyanin *Carcinus aestuarii*. *Oxidation Commun.* **23**, 1, 145–152
30. Favilla, R., Goldoni, M., Mazzini, A., Di Muro, P., Salvato, B., and Beltramini, M. (2002) Guanidinium chloride induced unfolding of a hemocyanin subunit from *Carcinus aestuarii* I. Apo form. *Biochim. Biophys. Acta* **1597**, 42–50
31. Favilla, R., Goldoni, M., Del Signore, F., Di Muro, P., Salvato, B., and Beltramini, M. (2002) Guanidinium chloride induced unfolding of a hemocyanin subunit from *Carcinus aestuarii* II. Holo form. *Biochim. Biophys. Acta* **1597**, 51–59
32. Lehrer, S.S. (1971) Solute perturbation of protein fluorescence. The quenching of the tryptophan fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry* **10**, 3254–3263
33. Lange, J., Elias, H., Paulus, H., Muller, J., and Weser, U. (2000) Copper(II) and copper(I) complexes with an open-chain N4 Schiff base ligand modeling CuZn superoxide dismutase: struc-

tural and spectroscopic characterization and kinetics of electron transfer. *Inorg. Chem.* **39**, 3342–3349
34. Hazes, B., Magnus, K.A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K.H., and Hol, W.G. (1993) Crystal structure of deoxygenated *Limulus polyphemus* subunit II hemocyanin at 2.18 Å resolution: clues for a mechanism for allosteric regulation. *Protein Sci.* **2**, 597–619
35. Voit, R., Feldmaier-Fuchs, G., Schweikardt, T., Decker, H., and Burmester, T. (2000) Complete sequence of the 24-mer hemocyanin of the tarantula *Eurypelma californicum*. Structure and intramolecular evolution of the subunits. *J. Biol. Chem.* **275**, 39339–39344
36. Beintema, J.J, Stam, W.T, Hazes, B., and Smidt, M.P. (1994) Evolution of arthropod hemocyanins and insect storage proteins (hexamerins). *Mol. Biol. Evol.* **11**, 493–503
37. Nakashima, H., Behrens, P.Q., Moore, M.D., Yokota, E., and Riggs, A.F. (1986) Structure of hemocyanin II from the horseshoe crab, *Limulus polyphemus*. Sequences of the overlapping peptides, ordering the CNBr fragments, and the complete amino acid sequence. *J. Biol. Chem.* **261**, 10526–10533
38. Di Muro, P., Beltramini, M., Nikolov, P., Petkova, I., Salvato, B., and Ricchelli, F. (2002) Fluorescence spectroscopy of the tryptophan microenvironment in *Carcinus aestuarii* hemocyanin. *Z. Naturforsch.* **57c**, 1084–1091
39. Ricchelli, F., Beltramini, M., Flamigni, L., and Salvato, B. (1987) Emission quenching mechanisms in *Octopus vulgaris* hemocyanin: Steady-state and time-resolved fluorescence studies. *Biochemistry* **26**, 6933–6939
40. Boteva, R., Ricchelli, F., Sartor, G., and Decker, H. (1993) Fluorescence properties of hemocyanin from tarantula (*Eurypelma californicum*): a comparison between the whole molecule and isolated subunits. *J. Photochem. Photobiol. B* **17**, 145–153
41. Strambini, G.B. and Gabellieri E. (1991) Quenching of indole luminescence by copper ions: A distance dependence study. *J. Phys. Chem.* **95**, 4347–4352
42. Dolashka-Angelova, P., Beltramini, M., Dolashki, A., Salvato, B., Hristova, R., and Voelter, W. (2001) Carbohydrate composition of *Carcinus aestuarii* hemocyanin. *Arch. Biochem. Biophys.* **389**, 153–155