doi:10.1016/j.jmb.2008.10.080 J. Mol. Biol. (2009) 385, 963–983

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Keyhole Limpet Hemocyanin: 9-Å CryoEM Structure and Molecular Model of the KLH1 Didecamer Reveal the Interfaces and Intricate Topology of the 160 Functional Units

Christos Gatsogiannis and Jürgen Markl*

Institute of Zoology, Johannes Gutenberg University, D-55099 Mainz, Germany

Received 10 August 2008; received in revised form 19 October 2008; accepted 29 October 2008
Available online
5 November 2008

Hemocyanins are blue copper-containing respiratory proteins in the hemolymph of many arthropods and molluscs. Molluscan hemocyanins are decamers, didecamers, or multidecamers of a 340- to 400-kDa polypeptide subunit containing seven or eight globular functional units (FUs; FU-a to FU-h), each with an oxygen-binding site. The decamers are short 35-nm hollow cylinders, with their lumen narrowed by a collar complex. Our recently published 9-Å cryo-electron microscopy/crystal structure hybrid model of a 3.4-MDa cephalopod hemocyanin decamer [*Nautilus pompilius* hemocyanin (NpH)] revealed the pathway of the seven-FU subunit (340 kDa), 15 types of inter-FU interface, and an asymmetric collar consisting of five "arcs" (FU-g pairs). We now present a comparable hybrid model of an 8-MDa gastropod hemocyanin didecamer assembled from two asymmetric decamers [isoform keyhole limpet hemocyanin (KLH) 1 of the established immunogen KLH]. Compared to NpH, the KLH1 subunit (400 kDa) is C-terminally elongated by FU-h, which is further extended by a unique tail domain. We have found that the wall-and-arc structure of the KLH1 decamer is very similar to that of NpH. We have traced the subunit pathway and how it continues from KLH1-g to KLH1-h to form an annulus of five "slabs" (FU-h pairs) at one cylinder edge. The 15 types of inter-FU interface detected in NpH are also present in KLH1. Moreover, we have identified one arc/slab interface, two slab/slab interfaces, five slab/wall interfaces, and four decamer/decamer interfaces. The 27 interfaces are described on the basis of two subunit conformers, yielding an asymmetric homodimer. Six protrusions from the cryo-electron microscopy structure per subunit are associated with putative attachment sites for N-linked glycans, indicating a total of 120 sugar trees in KLH1. Also, putative binding sites for divalent cations have been detected. In conclusion, the present 9-Å data on KLH1 confirm and substantially broaden our recent analysis of the smaller cephalopod hemocyanin and essentially solve the gastropod hemocyanin structure.

Keywords: hemocyanin; quaternary structure; cryo-electron microscopy; 3D reconstruction; molecular model

Edited by W. Baumeister

Introduction

Many molluscs have blue blood, as their hemolymph oxygen carrier, hemocyanin, is a type 3 copper protein that is related to tyrosinases. Molluscan hemocyanins have been intensively studied for their structure, function, and evolution,[1] and for immunological and clinical applications.[2] They differ fundamentally from arthropod hemocyanins (for the latter, see Martin *et al.*[3] and Decker *et al.*[4]) and occur

**Corresponding author.* E-mail address: markl@uni-mainz.de.

Abbreviations used: FU, functional unit; NpH, *Nautilus pompilius* hemocyanin; KLH, keyhole limpet hemocyanin; cryoEM, cryo-electron microscopy; OdH, *Octopus dofleini* hemocyanin; 3D, three-dimensional; HtH, *Haliotis tuberculata* hemocyanin; CCF, cross-correlation function.

either as decamers (five subunit dimers assembled as a hollow cylinder), didecamers (face-to-face assembly of two decamers), or multidecamers (elongated cylinders formed from a didecamer with added decamers). The polypeptide subunit is a concatenation of seven or eight paralogous functional units (FUs; termed FU-a to FU-h), each with a single oxygen-binding site.

Recently, we published a subnanometer (9 Å) cryo-electron microscopy (cryoEM) structure and molecular model of *Nautilus pompilius* hemocyanin (NpH) that, at the level of the basic decamer, resolved many of the hitherto unanswered questions concerning the molluscan hemocyanin architecture.[5] Our present report on keyhole limpet hemocyanin (KLH) 1 should be primarily considered alongside this NpH study. We demonstrated that this cephalopod hemocyanin is composed of 10 identical 340-kDa subunits arranged as anti-parallel pairs; due to an oblique arrangement of the five subunit dimers, the cylinder wall appears as a right-handed helix. The first six FU types (NpH-a to NpH-f) form the wall of the decamer, and the previously unknown subunit pathway within the wall is revealed. The 10 copies of NpH-g constitute the internal collar complex consisting of five discrete "arcs", with each arc connecting two subunit dimers. Unexpectedly, this arc pentamer was found to be shifted towards one cylinder opening, thereby imposing overall C5 (rather than D5) point–group symmetry on the decamer. As a consequence, the subunit dimer of NpH is an asymmetric homodimer.[5]

All NpH FUs are arranged pairwise as morphological units, and we identified the following duplets: a↔b, d↔e, c↔f, and g↔g. We also detected that the pairing mode strikingly resembles the protomer–protomer association in the published crystal lattice of FU-g from *Octopus dofleini* hemocyanin (OdH).[6] In side-view orientation of the decamer, the wall appears as three tiers, with the FU repeat a-b-c-f in the two peripheral tiers and with d-e-e-d in the central tier. Moreover, 15 different types of molecular inter-FU interface were identified, and most of them are likely to transfer forces during progressive oxygen-binding allosteric interactions. From these data, NpH can be considered as a molluscan hemocyanin prototype.[5]

Gastropod hemocyanins such as KLH1 differ from hemocyanin molecules of the NpH prototype in four major aspects: (i) the native molecule is a didecamer, with two decamers assembled face-to-face at their open ends;[1] (ii) the gastropod hemocyanin subunit is a 400-kDa polypeptide due to the presence of an additional FU type (FU-h); (iii) the asymmetric collar complex of the decameric half-molecule is enlarged by an outer annulus of five "slabs" representing FU-h pairs;[7] and (iv) FU-h carries a C-terminal tail of ca 100 amino acids[8] that has recently been found to possess a cupredoxin-like fold important for h↔h anti-parallel pairing.[9] Our recent comparative cryoEM three-dimensional (3D) analysis at 11-Å resolution has shown that the wall of *Nautilus* hemocyanin is very similar to the wall of a gastropod hemocyanin [*Haliotis tuberculata* hemocyanin (HtH) 1], but the two collar complexes appeared to be substantially different; however, at 11-Å, a reliable molecular model of the didecamer could not be obtained.[7]

KLH that occurs in the marine gastropod *Megathura crenulata* in the two isoforms KLH1 and KLH2 is a widely used immunological tool and promising tumor vaccine carrier;[2,10–12] its protein structure and disassembly/reassembly behavior have been extensively studied.[13–31] More recently, we determined the complete amino acid sequence of KLH1 and KLH2.[32] The present cryoEM/crystal structure hybrid model of KLH1 was created to substantially clarify the difference between the cephalopod hemocyanin decamer of the NpH prototype and the more complex gastropod hemocyanin didecamer.

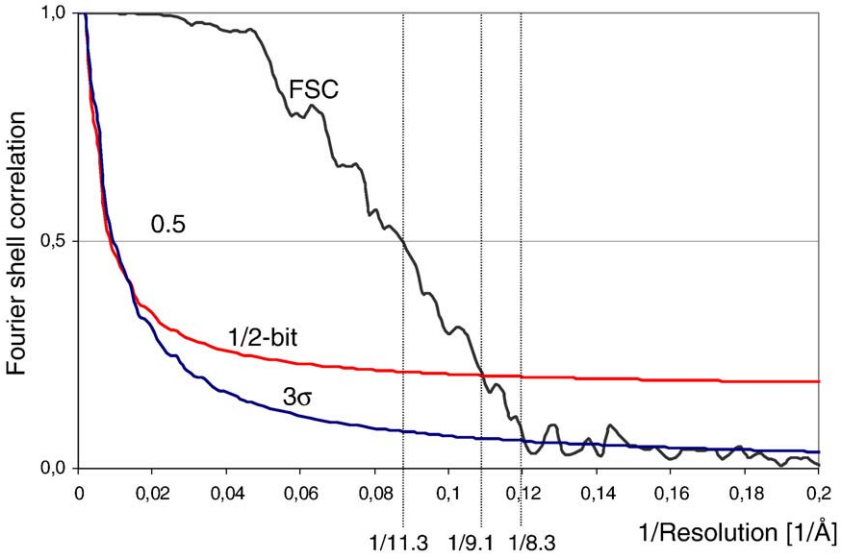

Fig. 1. Resolution determination of the final cryoEM structure of KLH1 by Fourier shell correlation. In this study, resolution is defined by the $FSC_{1/2\text{-bit}}$ criterion, but values for the more pessimistic $FSC_{0.5}$ and the more optimistic $FSC_{3\sigma}$ criteria are also given (for details, see Materials and Methods).

Results and Discussion

Wall architecture of the KLH1 decamer

Using ca 5000 randomly orientated KLH1 didecamers from cryoEM, we ultimately obtained a stable 3D volume with 9.1-Å resolution (according to the $FSC_{1/2\text{-bit}}$ criterion; Fig. 1) exhibiting D5 point–group symmetry. Compared to earlier 3D reconstructions of gastropod hemocyanin didecamers at lower resolution,[7,20,25,31,33–37] the present 9-Å cryoEM structure of KLH1 shows greater detail. The six types of wall FUs (KLH1-a to KLH1-f), as well as the FUs of arcs (KLH1-g) and slabs (KLH1-h), are clearly discernable (Fig. 2). Their extracted cryoEM densities, viewed in the same orientation, are closely comparable, but also show some specific differences that are particularly obvious for KLH1-a and KLH1-h (Fig. 3a).

From our previous comparative analysis of NpH and HtH1 at 11-Å resolution, it has been proposed that the principal wall architectures in cephalopod and gastropod hemocyanin are similar.[7] The present study shows that this correspondence is indeed much greater than expected. At 9 Å, the wall of the decameric half-structure of KLH1 is almost indistinguishable from the NpH molecule (see Fig. 2a–d; for comparison to NpH, see Figs. 2 and 3 in Gatsogiannis *et al.*[5]). This similarity encompasses the shape and orientation of the six wall FU types, and specific wall features (such as the anchor, and

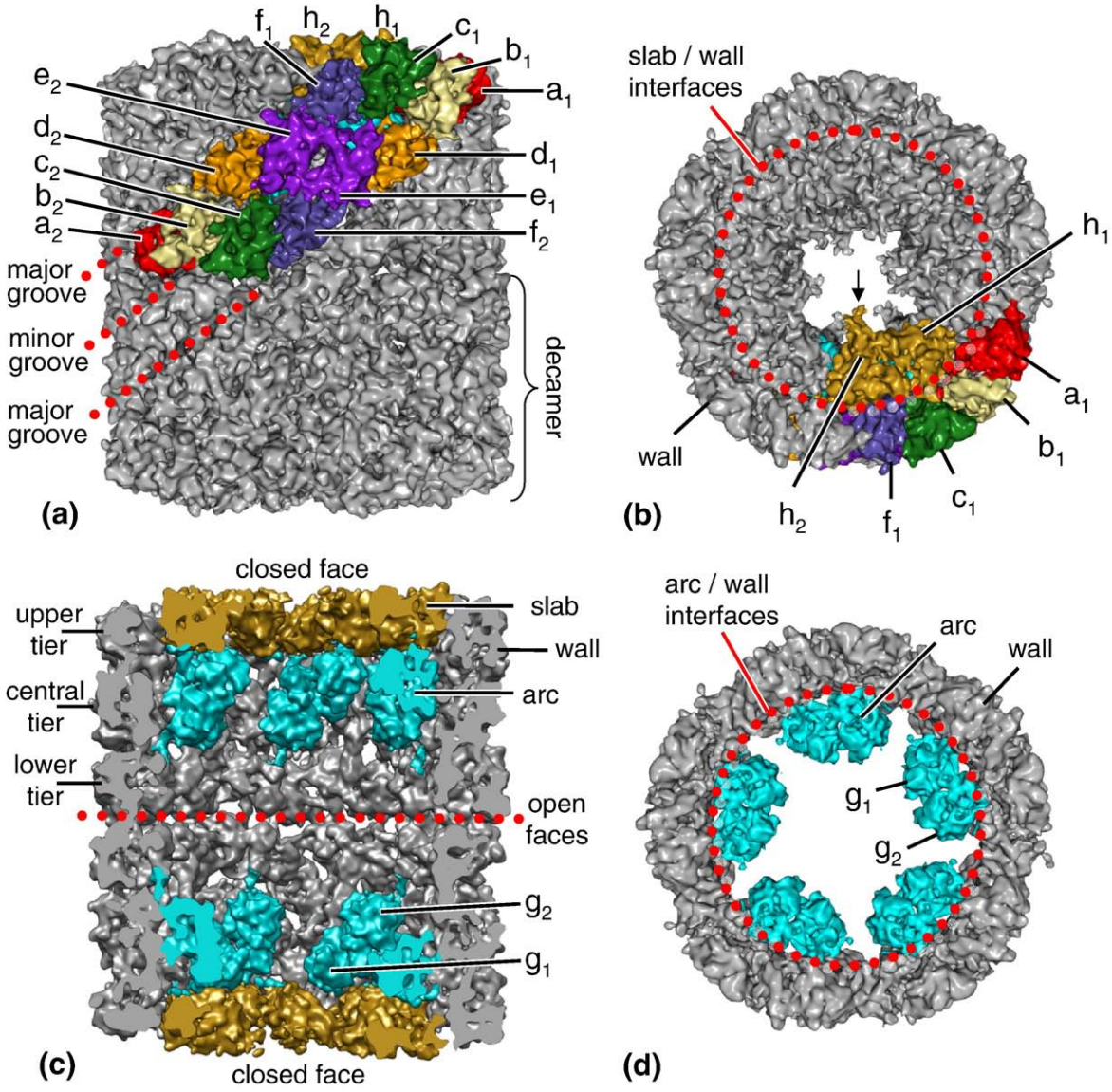

Fig. 2. 9-Å cryoEM structure of the KLH1. (a) Side view, (b) top view, and (c) cut-open view of the didecamer. (d) Top view of one extracted decamer, with the slab pentamer removed to reveal similarity with the NpH whole molecule (for comparison, see Fig. 3a in Gatsogiannis *et al.*[5]). The eight FUs are indicated (KLH1-a, red; KLH1-b, yellow; KLH1-c, green; KLH1-d, orange; KLH1-e, purple; KLH1-f, blue; KLH1-g, cyan; KLH1-h, gold). In (a), note the shape of the subunit dimer (colored). Also note the prominent protrusions from the collar narrowing the cylinder lumen (arrow in (b)), which are interpreted as N-linked glycans.

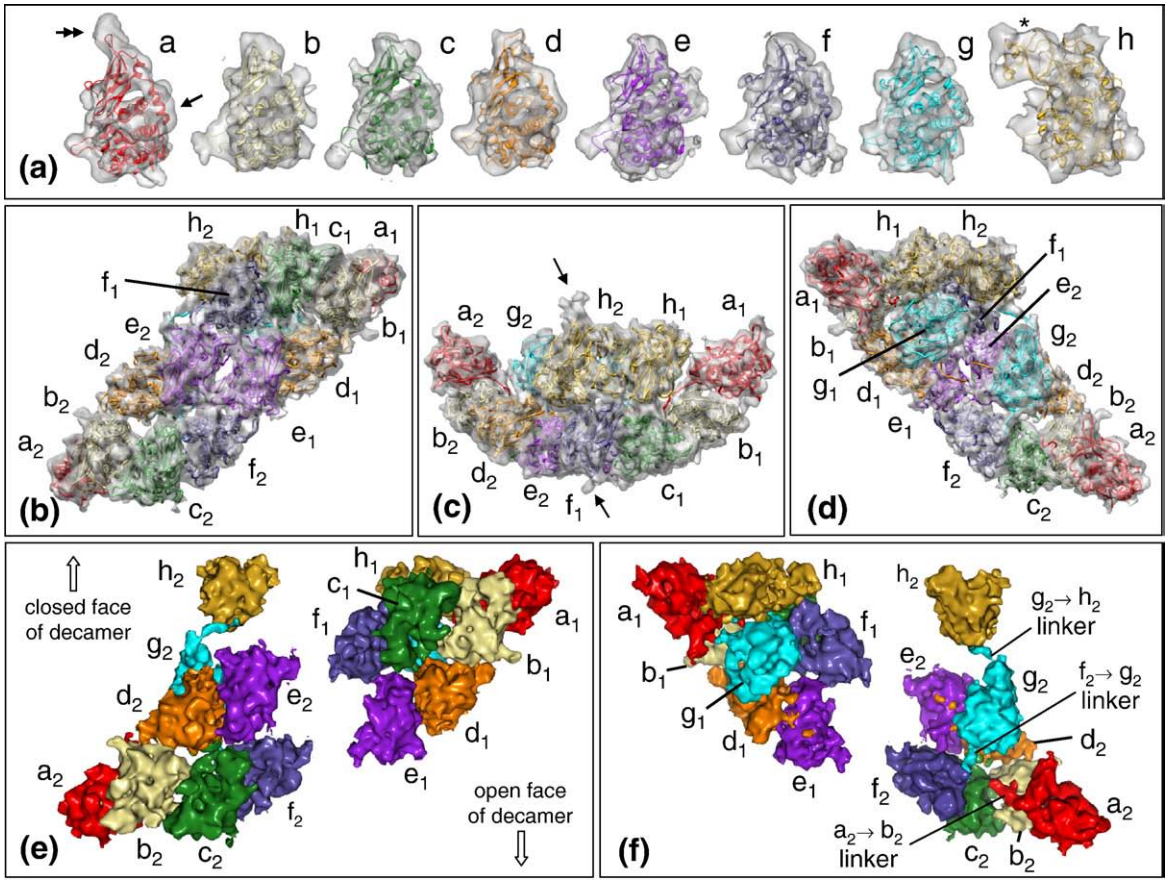

Fig. 3. Substructures of KLH1 extracted from the 9-Å cryoEM density map. (a) The individual FUs with their respective molecular models demonstrating their similarity and fitting quality. Note specially the appearance of FU-a (double arrow; enlarged β8→β9 loop as part of the anchor) and FU-h (asterisk; cupredoxin-like domain). Also note the representation of the exposed helix α15 (arrow) in all eight density maps. (b–d) Subunit dimer with the docked FUs shown in side view (b), top view (c), and side view rotated by 180° as observed from the cylinder lumen (d); each subunit dimer carries two separated arc FUs and one complete slab. Note the classical parallelogram (b and d) and "man-in-boat" profiles (c) of the subunit dimer; also note putative sugar trees (arrows in (c)). (e) The two subunit conformers in side view and arranged as in the subunit dimer. (f) The two subunit conformers rotated by 180°, as viewed from the cylinder lumen. For color coding of FUs, see Fig. 2.

the bridges that bury the inter-FU linkers). In this context, it is important to emphasize that, throughout the entire reconstruction process for KLH1, we did not use the cryoEM structure of NpH as reference. This means that both cryoEM structures have been entirely obtained independently. Features of the wall that differ in both cryoEM structures are protrusions from specific FU types, which are interpreted as glycan side chains (see below). Moreover, KLH1-a lacks a prominent protrusion that, in NpH-a, buries the slightly prolonged N-terminal part (see Fig. 2 in Gatsogiannis *et al.*[5]).

The basic repeating unit of the molluscan hemocyanin decamer is the subunit dimer,[34] but whether it is comparable in cephalopod and gastropod hemocyanins has been a long-standing question. The present description of the subunit dimer in KLH1 (Fig. 3b–d), together with our recent identification of the subunit dimer in NpH,[5] clarifies this point by showing that, in both hemocyanins, the subunit segment a-b-c-d-e-f-g dimerizes in the same manner. Another enigma that concerned hemocyanin researchers for decades has been the twisted pathway of the subunit, and whether it might be similar in the seven-FU cephalopod hemocyanin and the eight-FU gastropod hemocyanin.[1] The present identification of this intricate pathway in KLH1 (Fig. 3e and f), together with the corresponding

Fig. 4. Amino acid sequence of the 400-kDa polypeptide subunit of KLH1,[32] and a sketch of the crystal structure of OdH-g.[6] The sequence (AJ698341/CAG28309) is fragmented here to facilitate comparison of the eight FU types (a–h; KLH1-a to KLH1-h). In each FU sequence, the six copper-binding histidines (blue), the putative attachment sites for N-linked glycans (red), and the cysteines putatively involved in disulphide or thioether bridges (yellow) are highlighted. Also, the secondary structure elements in the respective molecular model are indicated (orange, α-helix; green, β-stand). Note that, compared to the crystal structure of OdH-g,[6] several short secondary structures are missing in the molecular models. Details of the cupredoxin-like domain are published elsewhere.[9]

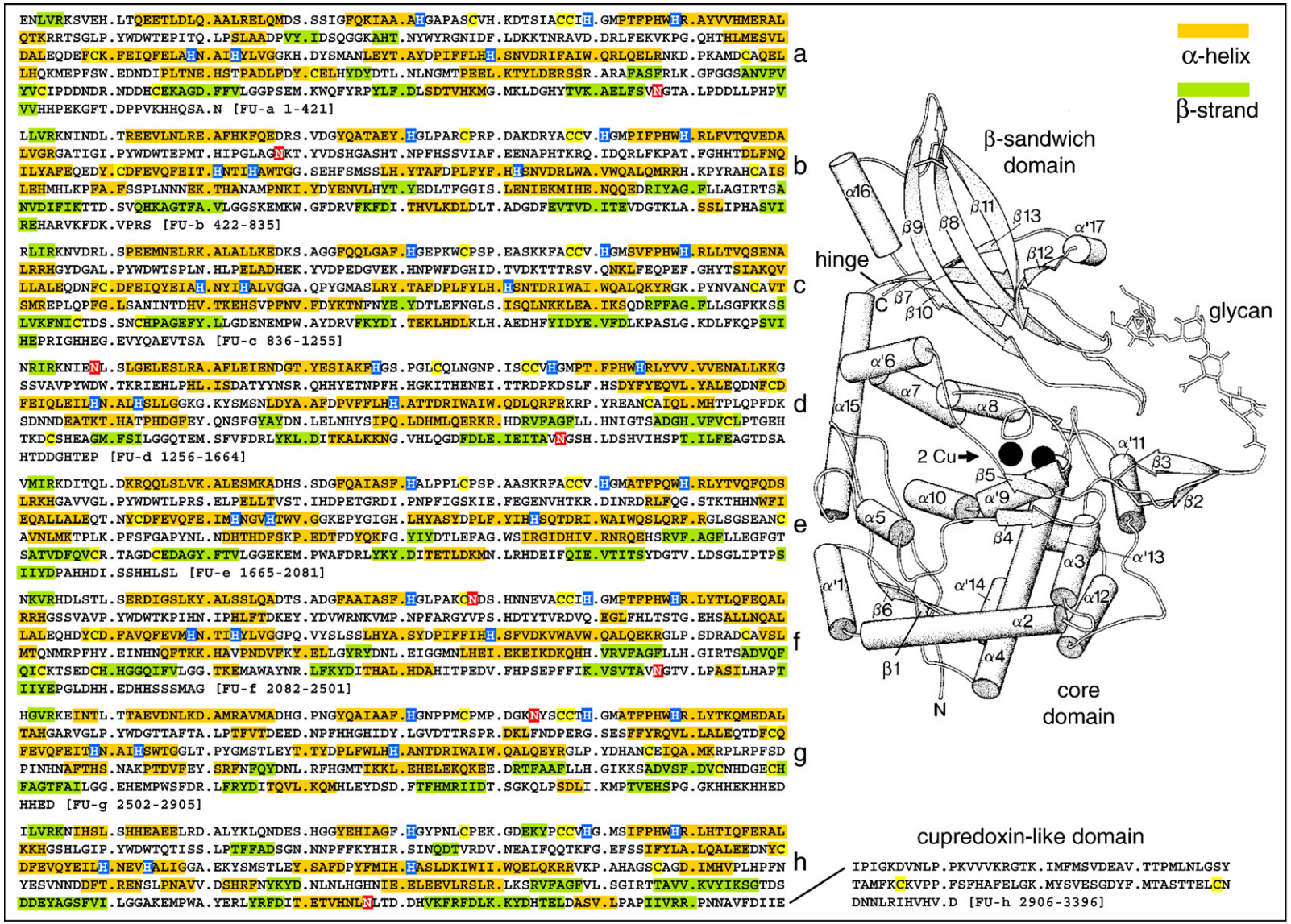

Fig. 4 (*legend on previous page*)

Table 1. Evaluation of the protein geometry of the finally used molecular models of the eight KLH1 FUs

Functional unit	Rotamer outliers (%)	Ramachandran outliers (%)	Ramachandran favored (%)	C^{β} deviations
KLH1-a	2.98	0.72	96.42	5
KLH1-b	3.06	0.97	96.60	5
KLH1-c	2.72	2.15	94.98	4
KLH1-d	3.61	0	95.82	4
KLH1-e	3.29	1.2	96.63	6
KLH1-f	1.69	0.49	95.8	2
KLH1-g	3.31	0.72	98.07	5
KLH1-h	2.37	0.52	95.12	3

The MolProbity software package was applied. In the case of KLH1-h, note that the values stem from the homology-modeled core/β-sandwich domain fragment, and not from the 4-Å crystal structure (for the latter, see Barends *et al.*[9]).

data on NpH,[5] also elucidates this point by revealing that the 3D structure of subunit segment a-b-c-d-e-f-g is highly conserved. Consequently, the major difference between the two hemocyanins concerns the additional features in KLH1: the contacts between the two decamers, and the intriguing slab pentamer. Details of the KLH1 structure will be highlighted below; for better orientation, its amino acid sequence[32] and a sketch of the tertiary structure of OdH-g[6] are given in Fig. 4.

Molecular model of KLH1 and definition of two subunit conformers

As already demonstrated for NpH,[5] our present subnanometer 3D reconstruction of KLH1 allows detection of predominant secondary structure elements within individual FUs, notably the six-stranded anti-parallel β-sandwich, but also the long and exposed helix α15 (see Fig. 3a). This, in turn, enabled us to perform high-quality rigid-body fitting of the homology-modeled individual FU types of KLH1 into their respective cryoEM densities (for quality assessment of molecular models, see Table 1). Additionally, we took advantage of the phenomenon that all FU types occur as morphological unit pairs; this enabled us to control and occasionally fine-tune docking by using the crystallographic dimers of OdH-g and KLH1-h as templates. Fitting quality was measured (correlation coefficients of 0.79–0.82; 0.81 in most cases) and is directly visible in Fig. 3. This docking was followed by loop refinements in order to improve the fitting of several prominent loops into appropriate cryoEM densities; during these final procedures, the locations of α-helices and β-strands remained unchanged.

The subunit pathway in the cylinder wall of the KLH1 decamer was analyzed by strategies developed for NpH.[5] A comparative analysis of the molecular models and cryoEM structures of KLH1 and NpH demonstrated that, in the wall, equivalent FU types occupy the same topological position and are connected *via* similar bridges in the density map (see Fig. 3f). This similarity allowed us to use the inter-FU linkers in the molecular model of NpH as templates for modeling their counterparts in KLH1. Some linker parameters are described in Table 2. The net result of this approach was that the upper and lower tiers of the KLH1 decamer are formed by the five copies of the FU repeat a-b-c-f, and the central wall tier contains five copies of the FU repeat d-e-e-d; the wall part of the subunit is triangular in shape, and the subunit dimer is formed in this region by anti-parallel association (see Figs. 2a and 3b–f). This is virtually the same situation as in the wall of NpH (for comparison, see Figs. 2 and 3 in Gatsogiannis *et al.*[5]). Indeed, rigid-body fitting of the molecular model of the extracted NpH wall into the cryoEM density of the wall of the KLH1 decamer yielded a correlation coefficient of 0.76, which underlines their similarity.

In our recent NpH study, we discovered that the two FU-g copies, united in one arc, differ in conformation with respect to their arc/wall interfaces; therefore, we designated them as FU-g_1 and FU-g_2.[5] In the present study, on the more complex situation in the didecamer, this concept has been expanded to the entire subunit; to facilitate description, we distinguish here a subunit conformer 1 (a_1-b_1-c_1-d_1-e_1-f_1-g_1-h_1) from a subunit conformer 2 (a_2-b_2-c_2-d_2-e_2-f_2-g_2-h_2). Conformer 1 occupies the five subunit positions at the closed face of the decamer (upper

Table 2. Available length of the inter-FU linker peptides

Linker peptide	a→b	b→c	c→d	d→e	e→f	f→g_1	f→g_2	g_1→h_1	g_2→h_2
Number of amino acids	20	13	19	16	14	17	17	17	17
Maximum length (3.8 Å per amino acid)	76	49	72	61	53	65	65	65	65
Potential length (2.8 Å per amino acid)	57	37	54	46	40	48	48	48	48
Minimum distance to be bridged (Å)	34	34	32	36	30	32	46	38	41

Note that utilization of the maximum length (the backbone fully extended to 3.8 Å per peptide bond) is highly unlikely. Here we assumed 75% of the maximum as the potential length. The minimum distance to be bridged in the molecular model means the distance between strand β13 of the preceding FU to strand β1 of the following FU (with the subunit pathway revealed and the contours of the cryoEM density neglected).

Table 3. Amino acid residues involved in the 27 types of interface between FUs

a↔b (wall morphological unit interface)
a: R308, 352WQFYRP357
b: 771WGFDRV776
d↔e (wall morphological unit interface)
d: L1554, H1556, 1597FVFDRL1602, H1642
e: E1980, 2016WAFDRL2021
c↔f (wall morphological unit interface)
c: 1185WAYDRV1190
f: L2389, H2391, 2427WAYNRL2432, H2478
b↔d (horizontal tier interface)
b: 440RE441, 444HK445, 448EDRS451, D453, R505, H540, N678
d: R1257, 1271ES1272, R1274, E1279, 1282ND1283, N1329, 1333KK1334
e↔f (horizontal tier interface)
e: K1684, 1687ES1688, 1691ADH1693, D1696, A1700, H1705, 1746RKH1748, 1926DH1927, H1929
f: H2086, L2091, 2093ERD2095, G2097, Y2101, S2105, 2163RH2164, S2167, 2195WR2196
b↔c (minor groove interface)
b: 426KN427, 429ND430, E434, T508, K568, F572, 574HH575, R649, D704, T706, 709GI710,E713, 831KVPRS835
c: 836RL837, K840, 844RLSPE848, E852, Y921, 986GH987, Y1062, N1113, 1117DT1118, E1120, 1125SISQ1128
e↔e (minor groove interface)
e/e: 1669KD1670, 1672TQLDK1676, T1760, 1776HD1777, E1779, S1815, K1817, 1819HHNWF1823, F1874, I1887, E1951
a↔f (major groove interface)
a: A106, 108PV109, 111ID112, 117KAH119, Y124, 136RAVDDR142, 189KHDY192
f: 2189KE2190, Y2192, 2197NKV2199, R2206, Y2215, V2217, D2219, 2270PQVY2273
d↔d (major groove interface)
d/d: 1359AT1360, 1362YN1363, 1367HHY1369, 1375HH1376, I1385, T1387, D1389, 1434KGKY1437
a↔e (major groove interface)
a: F384, 387NGTA390
e: S1713, 1715AASK1718
d↔b (major groove interface)
d: 1630AVNGSHL1636
b: R470, 472DAK474
a↔d (anchor interface)
a: C323, D327, 329DRNDDHC335
d: C1569, 1571PT1572, 1574EHTKDC1579, H1581
g_1↔g_2 (arc morphological unit interface)
g_1: 2837WSFDRL2842
g_2: 2837WSFDRL2842
g_1↔d_1 (arc/wall interface)
g_1: 2848TQ2849, 2852KQ2853, H2855, E2857
d_1: 1347KR1348, D1398, Y1401, 1471FR1472, 1529NHYSIPQ1535, 1537DHM1539, E1542
g_2↔d_2 (arc/wall interface)
g_2: Y2647, L2654, 2705DR2706, I2710, A2713, 2717YR2718, 2770NLRFH2774, 2776MTIKKLE2782
d_2: C1569, 1573GEHTKDCSH1581
a_2: N328, R330
g_1↔h_1 (arc/slab interface)
g_1: 2499MAGHG2503, E2760, 2762SR2763, N2765, Q2767
h_1: K3121, 3126PAHAGS3131, 3157EN3158, 3165VD3166, 3168HR3169
h_1↔h_2 (slab morphological unit interface)
h_1/h_2: C2952, E2958, Y2960, C2962, E3241, cupredoxin-like domain
h_2↔h_1 (slab/slab interface)
Numerous residues from helix α2 and the region around helix α5
h_1↔a_1 (slab/wall interface)
h_1: H2917
a_1: E296
h_2↔a_1 (slab/wall interface)
h_2: 2903HE2904, H2918, E2922, H2991
a_1: K148
h_1↔c_1 (slab/wall interface)
h_1: R3192, K3197, H3259
c_1: Y1067, E1098, D1107, H1094
h_1↔f_1 (slab/wall interface)
h_1: cupredoxin-like domain
f_1: 2404CK2405, C2410, 2449ED2450
h_2↔f_1 (slab/wall interface)
h_2: K2955
f_1: E2408
b_2↔a_2 (decamer/decamer interface)
b_2: 720HE721, 724QE725, 824HARV827, 829FD830
a_2: 1EN2, H261, 269DYC271, H274
c_2↔a_2 (decamer/decamer interface)
c_2: K855, R918
a_2: E14, R84
f_2↔c_2 (decamer/decamer interface)
f_2: 2394RT2395, S2463, A2466, 2470TVLP2473
c_2: W881, 883PSPEAS888, I1090
f_2↔f_2 (decamer/decamer interface)
f_2/f_2: E2408

If not specified as FU_1 or FU_2, the situations are identical in both FU conformers.

subunit in Fig. 3e and f), whereas conformer 2 occupies the positions at its open face (lower subunit in Fig. 3e and f; see also Fig. 10g). It should be pointed out that, from all our data, both conformers have the same amino acid sequence. However, due to their respective local environment, some of their interfaces with neighboring FUs are different. This is most pronounced in the collar complex, but also evident in the cylinder wall; for example, the FU repeat a_1-b_1-c_1-f_1 is connected to the slab pentamer, whereas the FU repeat a_2-b_2-c_2-f_2 couples the two decamers (see below). Pairing of conformers 1 and 2 yields an asymmetric homodimer.

The inter-FU interfaces in the cylinder wall

The molecular model of NpH revealed 12 different types of wall interface that have been described at the amino acid level and interpreted in terms of their possible contribution to allosterism.[5] As a common feature, these interfaces contact long α-helices from both adjacent FUs, and these α-helices carry an active site histidine at their distal end. This arrangement might transfer forces between the two active sites during the cooperative oxygen-binding process.[5] Compared to NpH, the present molecular model of KLH1 revealed similar molecular interfaces in equivalent positions. Of course, they are not identical with those in NpH, but they provide similar bonding and chemomechanical possibilities. Their comprehensive description is beyond the scope of the present report, but the participating amino acids are summarized in Table 3, and 4 of the 12 wall interfaces are shown in Fig. 5. It should be noted that due to the anti-parallel subunit arrangement in the wall, the 12 wall interfaces are identical in the two subunit conformers defined above; therefore, further distinction is not required.

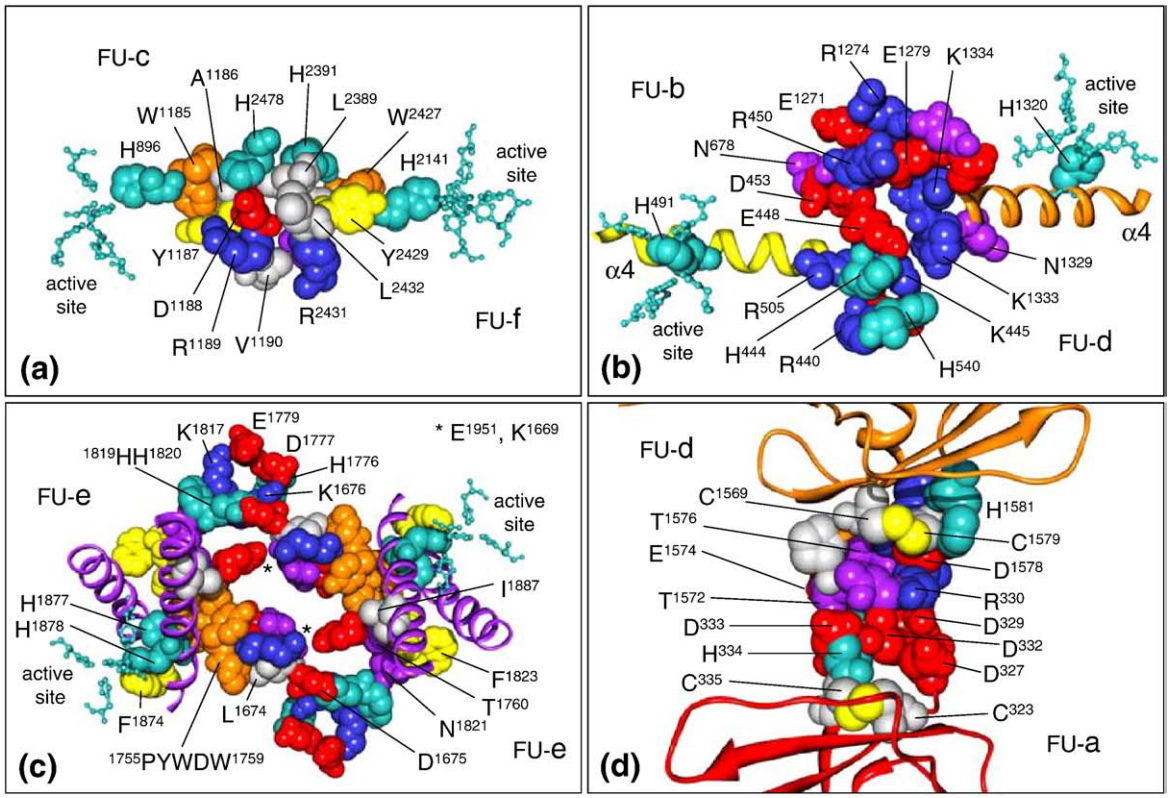

Fig. 5. Examples of molecular interfaces between FUs localized in the cylinder wall of KLH1. (a) Morphological unit interface c↔f; note the opportunity for allosteric force transfer between the two copper centers by mutually influencing the second active site histidine (H896 and H2141). For comparison to NpH, see Fig. 7d in Gatsogiannis *et al.*[5]; for the morphological unit interfaces a↔b and d↔e, see Table 3. Note that the three morphological unit interfaces of the wall connect FUs from the same subunit. (b) Horizontal tier interface b↔d; in this case, the neighboring active sites might interact by dislocating the long helices α4 that carry the third active site histidine (H491 and H1320). For comparison to NpH, see Fig. 8a in Gatsogiannis *et al.*[5]; for horizontal tier interface e↔f, see Table 3. (c) Minor groove interface e↔e connecting two subunits from the same subunit dimer; note possibilities for electrostatic bonding and allosteric interaction. For comparison to NpH, see Fig. 8d in Gatsogiannis *et al.*[5]; for minor groove interface b↔c, see Table 3. (d) Anchor interface a↔d crossing the major groove for connecting two subunit dimers; note opportunities for strong electrostatic bonding. For comparison to NpH, see Fig. 9d in Gatsogiannis *et al.*[5]; for the major groove interfaces a↔f, d↔d, a↔e, and d↔b, see Table 3. Color coding of residues: red, acidic; blue, basic; purple, polar; gray, hydrophobic; sea green, histidine; orange, tryptophan or PYWDW; yellow, phenylalanine; yellow/red, tyrosine; yellow/gray, cysteine.

The interface situation in the cylinder wall is complex, and we recommend consulting Fig. 6 in Gatsogiannis *et al.*[5] for better orientation. As mentioned, the cylinder wall is composed of 10 copies of the subunit segment a-b-c-d-e-f (see Fig. 2). Within this segment, the six FUs are arranged pairwise as three morphological units, thereby defining the morphological unit interfaces a↔b, d↔e, and c↔f (see Fig. 2a and b). They correspond structurally to the contact between the two protomers in the crystal lattice of *Octopus* FU-g.[5] In side-view orientation, the wall of the decamer exhibits three horizontal layers, or tiers, of 10 morphological units each (see Fig. 2c); they are interconnected by two types of horizontal tier interface (b↔d and e↔f; see Fig. 2a). Moreover, the cylinder wall shows two kinds of oblique fenestrated trenches, termed the major groove and the minor groove; the minor groove runs within each subunit dimer, whereas the alternating major groove separates adjacent subunit dimers (see Fig. 2a). The inter-FU contacts bridging these trenches are the minor groove interfaces (b↔c and e↔e) and the major groove interfaces (a↔f, d↔d, a↔e, and d↔b), respectively.

A prominent wall interface of NpH is the anchor (a↔d) that appears as a vertical rod at the rear of the large windows of the major groove where it connects two adjacent subunit dimers. It is formed between NpH-a and NpH-d by their joining β8→β9 loops stabilized by disulfide bridges.[5,7] The anchor is also part of the present cryoEM structure of KLH1 and appears to be very similar to the NpH anchor. As in NpH-a, also in KLH1-a (but not in any other FU type), the β8→β9 loop is enlarged by four aspartic acids (D327, D329, D332, and D333). The cryoEM structure of KLH1-a shows a prominent protrusion in this region that is lacking in the other FU types (see Fig. 3a). This acidic cluster offers strong bonding opportunities with polar and basic amino acids in the opposing β8→β9 loop of KLH1-d (see Fig. 5d). Such a conservation of primary and secondary structures in the two hemocyanins that diverged more than 520 million years ago is remarkable[38] and underlines the fact that the

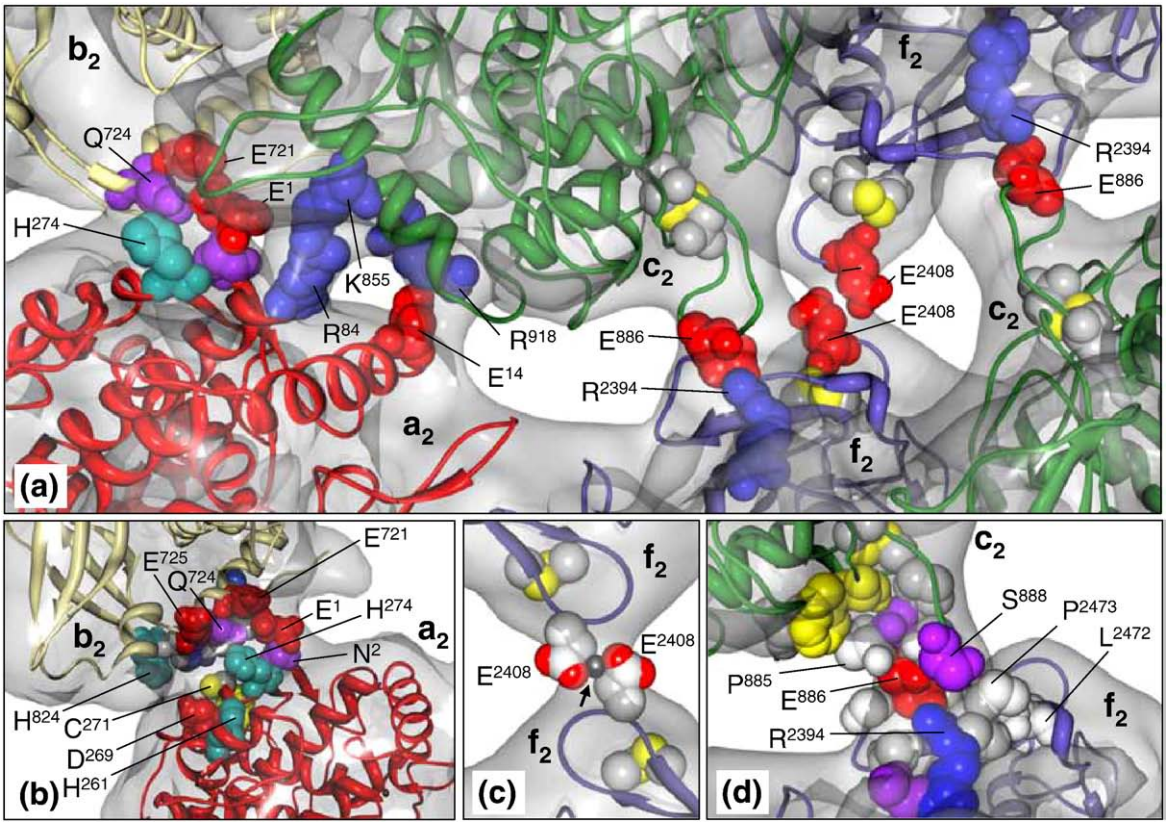

Fig. 6. The four decamer/decamer interfaces of KLH1. (a) Bridges between the two associated decamers visible in the cryoEM structure representing (from left to right) the interfaces $b_2 \leftrightarrow a_2$, $c_2 \leftrightarrow a_2$, $f_2 \leftrightarrow c_2$ (visible here in two copies), and $f_2 \leftrightarrow f_2$. Note that $c_2 \leftrightarrow a_2$ is represented by two bridges in the density map: one bridge is explained by electrostatic bonding; in the other bridge, two basic residues (R84 and K855) come together, which might indicate a bridging divalent anion. (b–d) Molecular details of interfaces $b_2 \leftrightarrow a_2$, $f_2 \leftrightarrow f_2$, and $f_2 \leftrightarrow c_2$, respectively. Note various charged residues and the solitary cysteine surrounded by three histidines in (b). Note opportunities for electrostatic bonding and hydrophobic shielding in (d). In (b) and (c), note opportunities for divalent cation binding between E1 and E721, and the two E2408 copies, respectively (arrow in (c); putative Ca^{2+} or Mg^{2+}).

anchor is indispensable as a connecting element between two adjacent subunit dimers; it might be responsible for overall stability and probably dampens movements during allosteric interaction.

Putative divalent cation-binding sites within the decamer

In NpH, carboxyl residues that are likely to stabilize the quaternary structure by bridging divalent cations such as Ca^{2+} and Mg^{2+} have been identified.[5] The effects of divalent cations on the quaternary structure of KLH1 and KLH2 have been carefully studied and found to be complex.[22–24] In the wall of the KLH1 decamer, we identified four carboxyl residue pairs, at a distance of no more than 7 Å, that are not neutralized by adjacent positively charged residues; therefore, they might bind a divalent cation (D449/E1279, 703ED704/D1117, E1244/E1478, and E1251/E1514). An additional site (E1246/E2492) represents contact between two linker peptides, $c_2 \rightarrow d_2$ and $f_2 \rightarrow g_2$, and probably supports stretching of the latter (in case of the differently folded $f_1 \rightarrow g_1$ linker, such contact is lacking). All five putative carboxyl sites of the KLH1 decamer are localized within the same subunit, which is also true for the majority of such sites in NpH.[5]

Evidence that divalent cations might link internal structures of the subunit, rather than reinforce subunit/subunit interfaces, seems to contradict observations that stabilizing the subunit dimer requires Ca^{2+} and/or Mg^{2+}.[1,39] However, it may well be that, for a stable subunit dimer, compaction of the subunit that is also supported by divalent cations is needed. Indeed, in dissociation/reassociation experiments of *Helix pomatia* hemocyanin, a "loose" conformation and a "compact" conformation of the subunit dimer have been distinguished; the loose → compact transition was achieved by lowering the pH and/or by adding Ca^{2+}.[34,40,41]

The decamer/decamer interfaces ($f_2 \leftrightarrow c_2$, $c_2 \leftrightarrow a_2$, $b_2 \leftrightarrow a_2$, and $f_2 \leftrightarrow f_2$)

The present cryoEM structure reveals, for the first time, molecular details of the contact zones between the two halves of a molluscan hemocyanin didecamer (Fig. 6a). Interdecamer bridges visible in previous 3D reconstructions of gastropod hemocyanins have remained ambiguous because of lower reso-

lution and unclear subunit topology.[7,25,31,35–37] The present cryoEM structure of KLH1 reveals 40 significant bridges representing three different types of interface, namely, $b_2 \leftrightarrow a_2$, $c_2 \leftrightarrow a_2$ (showing two separate contacts), and $f_2 \leftrightarrow c_2$ (see Fig. 6a); each of these bridges/interfaces is present in 10 copies. In addition, there are five narrow bridges formed by interface $f_2 \leftrightarrow f_2$ (see Fig. 6a). The four types of interface will now be described in relation to the molecular model.

Interface $b_2 \leftrightarrow a_2$ contains various charged residues at either side that might form contacts (Fig. 6b). A most interesting constellation in this interface is a single cysteine (C271) surrounded by three histidines (H274, H261, and H824). Apart from the single cysteine that forms (in all FU types) a thioether bridge with the second active site histidine, the cysteines in the KLH1 molecule are involved in disulphide bridges; the only exception is C271 (see Fig. 4), as C271 probably forms a thioether bridge with H274. On the other hand, histidine H824 is part of the hinge connecting the core and β-sandwich domain of KLH1-b_2, and hinge movements of the β-sandwich domain against the core domain might modulate oxygen affinity.[5,6] Therefore, this peculiar amino acid constellation, which is lacking in NpH,[38] might be involved in allosteric signal transfer between the two decamers.

Two carboxylate residues in interface $b_2 \leftrightarrow a_2$, E1 and E721, might bind a bridging divalent cation (Fig. 6b). A second potential site in this respect is interface $f_2 \leftrightarrow f_2$. It is localized at the luminal surface of the wall and shows a narrow bridge between two opposing copies of KLH1-f_2. In the molecular model, this contact is formed by the β8→β9 loop, which is stabilized by a disulfide bridge (the corresponding loop forms the anchor between FU-a and FU-d). There are only two available residues that could be responsible for bonding, namely, the two opposing E2408 copies (see Fig. 6a). These carboxylate residues are 4 Å apart and form an acidic spot that is not compensated for by positively charged amino acids; therefore, we assume that they bind a Ca^{2+} or Mg^{2+} (Fig. 6c). Indeed, it has been demonstrated in KLH1 that the decamer→didecamer transition can be controlled *via* the Ca^{2+}/Mg^{2+} level.[21–24] More specifically, reassembly experiments in KLH1 and KLH2 have demonstrated that didecamerization requires free decamers,[27] and ion-stabilized decamers reassemble into didecamers in the presence of 10 mM $CaCl_2/MgCl_2$.[23,24,29] Our data provide the structural basis for these observations, suggesting that decamer/decamer bonding in KLH1 is indeed reinforced by 15 bridging divalent cations. This opportunity is absent not only in the molecular model of NpH as one might expect but also in the second KLH isoform, KLH2.[32] The corresponding isoform of KLH1 in *H. tuberculata* (HtH1) shows the site,[8] whereas HtH2 lacks it.[42] This agrees with data from reassembly experiments showing that KLH1/HtH1 exhibit comparable didecamerization properties in the presence of divalent cations different from KLH2/HtH2.[29]

Interface $c_2 \leftrightarrow a_2$ offers various possibilities for electrostatic bonding, notably between E14 and R918, which would explain one of the two bridges visible in the cryoEM structure (see Fig. 6a). The second bridge is occupied by two opposing basic residues (R84/K855) that are not compensated for by negatively charged amino acids; this might indicate a binding site for a divalent anion. However, there is no evidence from reassociation experiments that divalent anions are required for reassembly; rather, such anions seem to influence the oxygen-binding function in some hemocyanins.[1] Interface $f_2 \leftrightarrow c_2$ shows an opportunity for a strong salt bridge shielded by hydrophobic residues, namely, between R2394 in the β7→β8 loop and E886 in the α3→α4 loop, which is stabilized by a disulphide bridge (Fig. 6d). Therefore, the 10 copies of this interface should provide a strong attractive force between the two decamers. Indeed, in NpH-c, the corresponding loop lacks a disulphide bridge, which might not be required in a hemocyanin restricted to solitary decamers.

The suggested strong ionic stabilization of the decamer/decamer contacts is also consistent with earlier ultracentrifugation studies on a bivalve hemocyanin in different media, assuming only few hydrophobic groups at the interdecamer interfaces.[43] The 45 interdecamer bridges, representing four types of interface, clearly appear as fixation points. Moreover, some of the contacts suggest that, in relation to allosterism, the associated decamers might not act as independent units. However, didecamers do not significantly differ in cooperativity from decamers;[1] therefore, we are still skeptical about the present indications for allosterism between the two decamers. In this context, it should be noted that, on one hand, oxygen binding studies have indicated that the allosteric units of *Octopus* hemocyanin might encompass not more than seven active sites.[44] On the other hand, the molecular model of NpH indicates that, in a nested way, allosterism might go beyond this limit,[5] and we presume the same for KLH1 from the structural similarities described here.

Topology and interfaces of the five arcs (KLH1-g pairs)

In *Nautilus* hemocyanin, five discrete arcs consisting of FU-g pairs constitute the collar complex; this $g \leftrightarrow g$ contact belongs to the morphological unit interfaces because it follows the association mode of the two protomers in the *Octopus* FU-g crystal. It was difficult to detect, and it was completely unexpected for this arc pentamer to be shifted towards one of the edges of the decamer, thereby disproving the established concept of D5 point–group symmetry for these cephalopod hemocyanins.[5] To explain this asymmetry, we suggested, as already mentioned, that the two FU-g copies of each arc represent two different conformations, designated as g_1 and g_2. In g_1 conformation, the FU is located at the center of the triangular subunit (see Fig. 3f); within

the decamer, 5 of the 10 FU-g copies occupy this position. During oligomerization, the alternating five FU-g copies are dislocated towards g_1 of the neighboring subunit dimer to form the arcs, thereby producing the asymmetric position of the collar complex within the hollow cylinder. This topology has been termed g_2 (see Fig. 3f); therefore, the arc interface is designated as $g_1 \leftrightarrow g_2$ (see Fig. 10c in Gatsogiannis *et al.*[5]).

Superposition of the 9-Å cryoEM structures of NpH and KLH1 showed that, at this resolution limit, the five arcs are topologically identical in both hemocyanins (see Fig. 7a). Comparison of the two molecular models shows that the KLH1 arc is shifted ca 5 Å towards the open face of the decamer and, indeed, the f→g linker in KLH1 is one amino acid shorter than in NpH. However, at 9-Å resolution, it is uncertain whether this shift is real. The arc interface $g_1 \leftrightarrow g_2$ (Fig. 7b), as well as the two arc/wall interfaces $g_1 \leftrightarrow d_1$ and $g_2 \leftrightarrow d_2$ (Fig. 7c and d), are fully comparable to those in NpH (for comparison, see Figs. 7a and 10 in Gatsogiannis *et al.*[5]); the amino acids involved in these contacts are summarized in Table 3. It should be noted that FU-g is primarily attached to the wall *via* these noncovalent g↔d contacts, whereas the longer covalent f→g linker appears to be a secondary link.

KLH1-f_2 and KLH1-g_2 are connected, in the cryoEM structure, *via* a well-defined coil (see Fig. 3f) that most likely represents the stretched $f_2 \rightarrow g_2$ linker (see Fig. 7a). In NpH, this linker is exceptionally long (51 Å) and was not fully correlated with the cryoEM density map, and this has previously been considered as the only flaw of the proposed subunit pathway.[5] Therefore, the additional mass in the cryoEM structure of KLH1 and the shorter linker distance to be bridged (46 Å) provide further strong arguments in favor of the suggested subunit pathway. Probably, the arcs are more flexible in NpH because FU-h, as well as the putative divalent cation bridge between the $f_2 \rightarrow g_2$ and $c_2 \rightarrow d_2$ linkers (see above), are absent as stabilizing elements; therefore, this long linker might be blurred in the 3D reconstruction of NpH, yielding a local resolution limitation.

Topology of KLH1-h in the slab pentamer, and the subunit pathway

The slab pentamer of the collar complex is exclusively composed of FU-h, as deduced from immunoelectron microscopy[29] and from the coincidence that some cephalopod hemocyanins (NpH and OdH) lack both FU-h and the five slabs. The

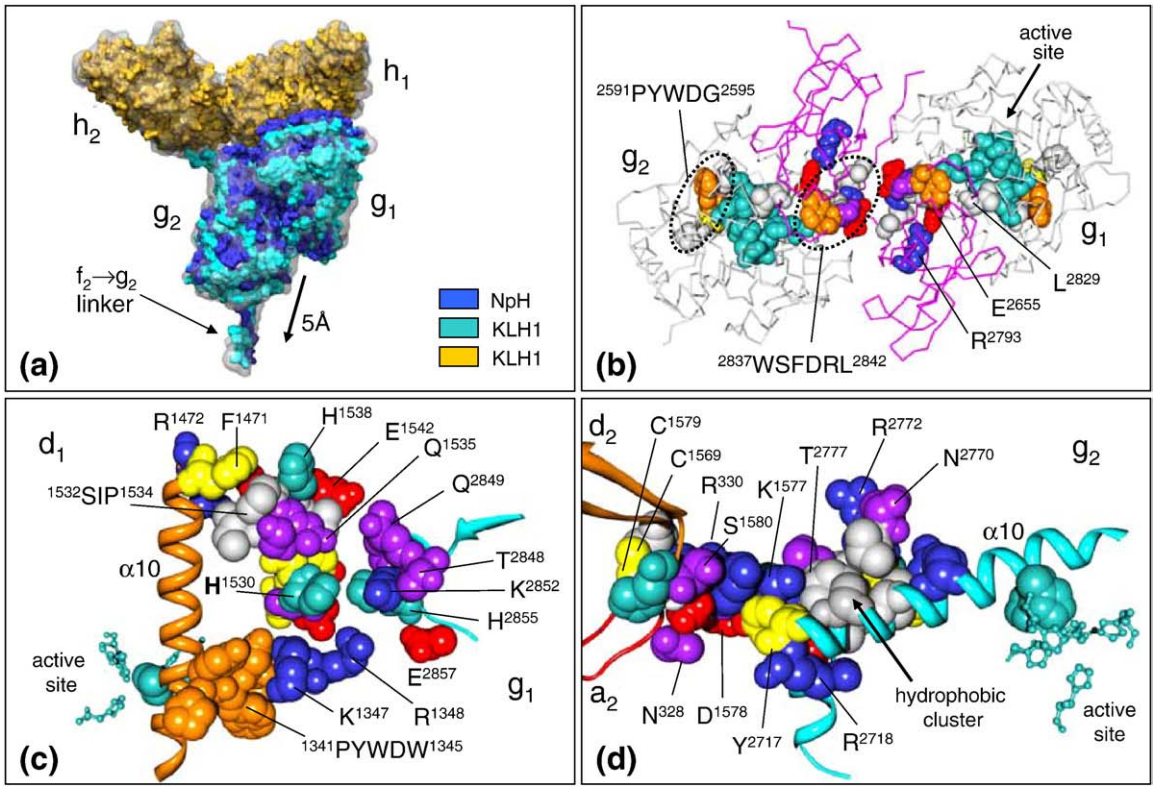

Fig. 7. Interfaces associated with the arc. (a) Comparison of the arc topology in the molecular model of KLH1 (cyan) and NpH (dark blue); note the 5-Å deviation, but identical association mode. The extracted cryoEM density (opaque) and the molecular models (gold) of two FU-h copies of KLH1 are also shown. Note that the visualized FU-h dimer is not a slab, but represents neighboring FU-h copies from two adjacent slabs (for details, see Fig. 8). (b) Morphological unit interface $g_1 \leftrightarrow g_2$; for comparison to NpH, see Fig. 7a in Gatsogiannis *et al.*[5] (c) Arc/wall interface $g_1 \leftrightarrow d_1$ and (d) arc/wall interface $g_2 \leftrightarrow d_2$; for the striking similarity of the corresponding interfaces in NpH, see Fig. 10a and b in Gatsogiannis *et al.*[5] In (d), note that the β8→β9 loop of KLH1-a_2 contributes two residues, R330 (hidden) and N328.

individual slabs are discernable at lower resolution,[7,25,31,35–37] but the present cryoEM structure reveals them in greatly improved detail (Fig. 8a), together with various slab/slab, arc/slab, and slab/wall bridges. Rigid-body fitting of the homology-modeled core/β-sandwich domain fragment of FU-h was efficient, but it left an empty space in the central region of the slab. This space could represent the peculiar tail of ca 100 amino acids (see Fig. 4), characteristic of FU-h.[8] Indeed, the recent 4-Å crystal structure of KLH1-h shows C-terminally a cupredoxin-like fold that exactly fills this space as detected using our previous 11-Å cryoEM structure of HtH1; moreover, the two FU-h copies in the slab seemed to be assembled similarly to the two protomers in the KLH1-h crystal.[9] This constellation could be fully confirmed using the present 9-Å cryoEM structure of KLH1. The possibility to dock KLH1-h as the crystallographic dimer confirmed that fitting of the single core/β-sandwich domain fragment into the slab was rather accurate and enabled a minor improvement (see examples in Fig. 8a). Nevertheless, even at 9-Å resolution, the intricate arc/slab architecture and the long g→h linker (17 amino acids) required careful topological analysis to determine how the subunit pathway continues, in both subunit conformers, from FU-g to FU-h.

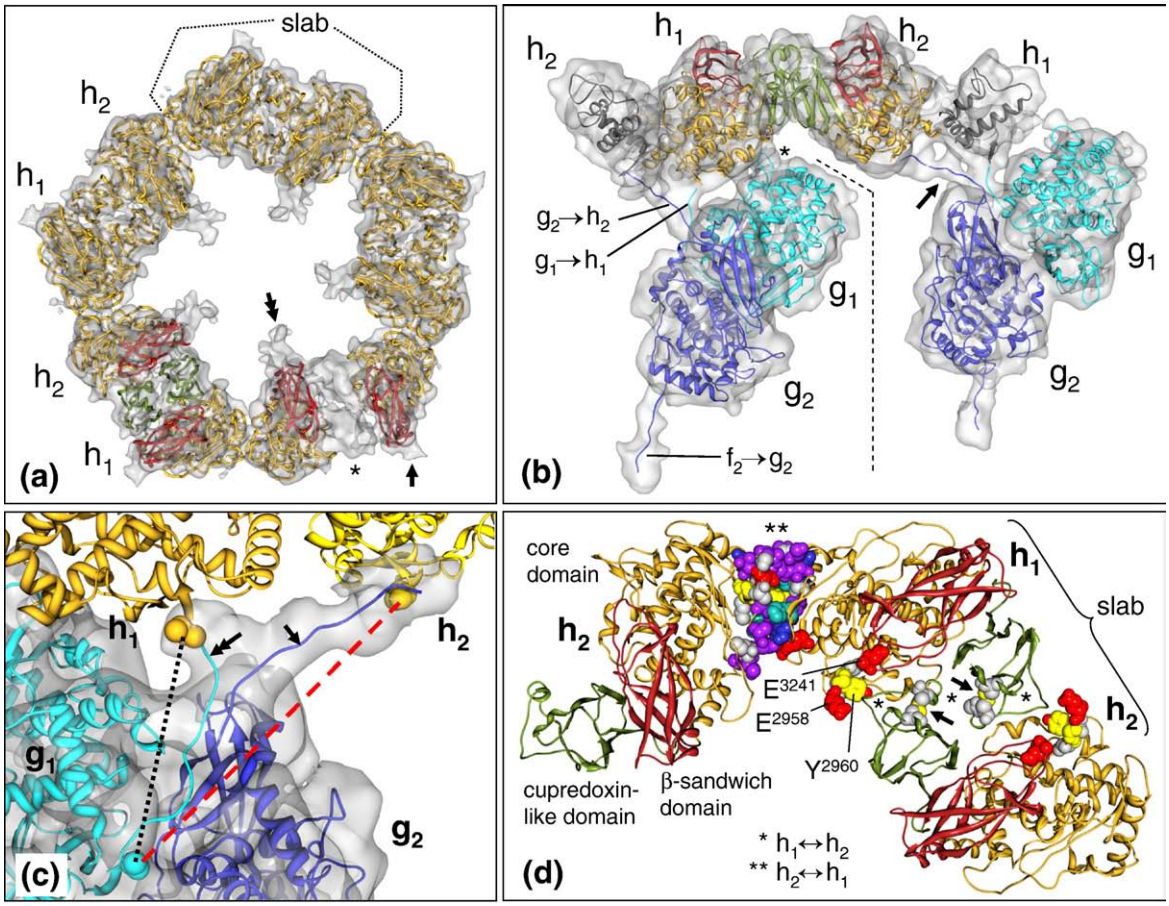

Fig. 8. Interfaces within and between slabs. (a) Extracted slab pentamer, with the crystal structure of KLH1-h docked in four slabs, and two molecular models of the core/β-sandwich domain fragment docked in the remaining slab (asterisk). Note the empty space in the density map that, in the other slabs, is occupied by the cupredoxin-like domain. Double arrow: putative carbohydrate side chain in KLH1-h_2; arrow: corresponding truncated protrusion in KLH1-h_1. (b) Extracted cryoEM structure and molecular model of a segment of the collar complex demonstrating the subunit pathway and why other possibilities for covalent linkage are excluded: a complete slab (gold, core domain; red, β-sandwich domain; green, cupredoxin-like domain), two adjacent N-terminal slab fragments (gray), and two associated arcs (cyan/blue). Note the $g_1 \rightarrow h_1$ interface (asterisk). Also note the cryoEM density containing both arc→slab linkers and its furcation (arrow). Alternative subunit pathways with linkers to cross the dashed line are excluded by insufficient linker length and lack of appropriate cryoEM densities. (c) Close-up of the cryoEM structure in the region of the curved bridge associated with the $g_2 \rightarrow h_2$ linker (arrow). Note that this far route is excluded for the $g_1 \rightarrow h_1$ linker (direct distance, 62 Å; red broken line), whereas the closer FU-h copy (termed h_1) is easily accessible (double arrow; direct distance, 38 Å; black broken line). This ultimately clarified the subunit pathway. (d) The two types of molecular interface in the slab pentamer. A complete slab is shown together with an associated KLH1-h_2 copy from the neighboring slab. Note that interface $h_1 \leftrightarrow h_2$ (asterisks) combines the two constituents of the slab at their cupredoxin-like domains (arrows; disulphide bridge); for details, see Barends *et al.*[9] Also note that the alternating interface $h_2 \leftrightarrow h_1$ (double asterisk) joins the two adjacent slabs by numerous polar residues provided by the core domains.

As a first step, we considered whether an individual slab would, like an arc, link two adjacent subunit dimers, or whether it would serve as a clamp within each dimer between the two constituent subunits. With its 17 amino acids, the g→h linker peptide would bridge a distance of 65 Å if it were fully stretched, but irregularly structured regions rarely exceed 75% of their maximum length. This means that the linker might not cover more than 48 Å (see Table 2). From this situation, it is completely clear that the two components of a given arc are not covalently linked to the two FU-h copies of the same slab; although one linker would of course fit, the other linker would have to bridge at least 100 Å of open space (Fig. 8b). Therefore, the two polypeptides from neighboring subunit dimers that join each other in one arc to form interface $g_1 \leftrightarrow g_2$ continue into two different but adjacent slabs. In other words, the subunit dimer contains one complete slab (see Fig. 2b–d).

The next step was to unravel how the polypeptide chain continues beyond the conformers KLH1-g_1 and KLH1-g_2. Although the two KLH1-h copies of each slab are completely identical in primary structure, they differ in their interaction with the arc and the wall. In one FU-h conformer, the tip of its β-sandwich domain contacts the wall; in the other conformer, it is directed into the cylinder lumen (see Fig. 8a). Therefore, it is well justified to distinguish KLH1-h_1 from KLH1-h_2, but which is which? Would the two g→h linkers run parallel or across (see Fig. 8b)? From the linker distances, KLH1-g_2 could lead to both FU-h conformers equally well (20 and 41 Å, respectively). From KLH1-g_1, the closely associated KLH1-h conformer (pointing towards the wall) is within a distance of 38 Å, whereas the more distant KLH1-h conformer (pointing into the lumen) would require 62 Å to be bridged (Fig. 8c). Such an extreme stretching of the linker peptide is unlikely, and it would yield a rather awkward subunit dimer. Moreover, the curvature of the corresponding bridge in the cryoEM structure would increase the distance to be covered from KLH1-g_1 beyond the 65-Å limit, whereas the linker emanating from KLH1-g_2 could be easily fitted into this bridge (see Fig. 8c). Therefore, we exclude the possibility of a crosswise connection and are fully convinced that the connection is indeed parallel as shown in Fig. 8b and c.

The subunit pathway identified in this study leads from KLH1-g_1 to the closely associated FU-h copy (pointing outwards in Fig. 8a), which therefore, per definition, is KLH1-h_1; the alternating FU-h copy, pointing inwards, is KLH1-h_2. These assumptions lead to a KLH1 subunit dimer with two half arcs as in NpH, but additionally equipped with a complete slab (see Fig. 3b–d). This shape is compatible with the "parallelogram" and "man-in-boat" profiles seen in

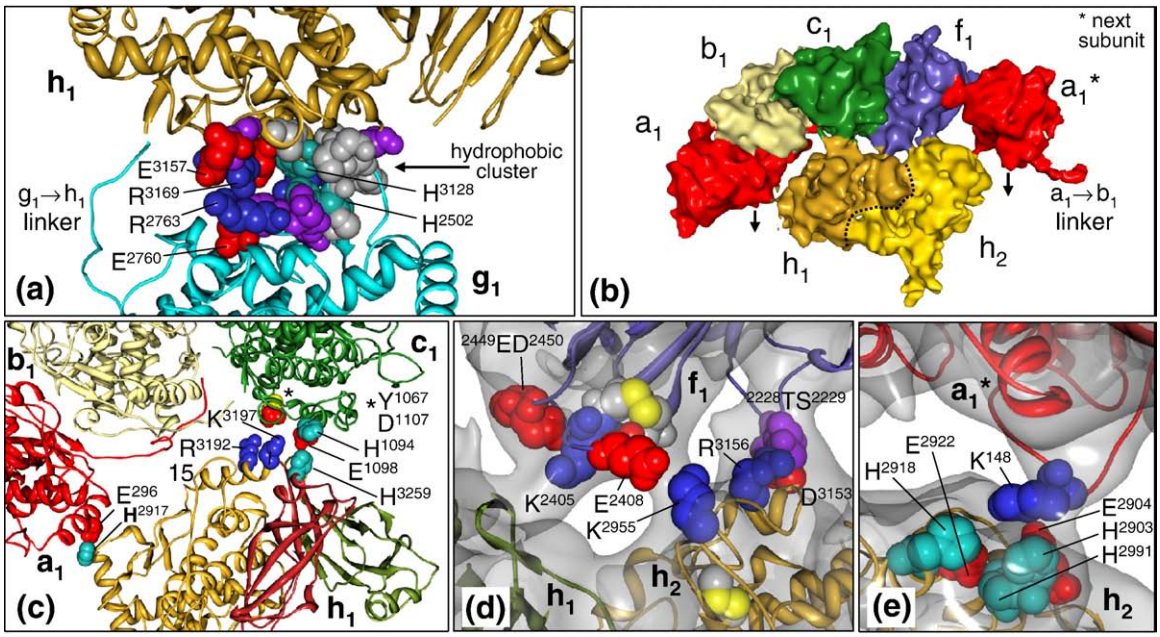

Fig. 9. The arc/slab and slab/wall interfaces. (a) Interface $g_1 \leftrightarrow h_1$; note the dense cluster of charged and hydrophobic amino acids indicating strong bonding, whereas the covalent $g_1 \rightarrow h_1$ linker appears more as a flexible contact. In the other conformer, the $g_2 \rightarrow h_2$ linker is the only contact (see Fig. 8b). (b) Extracted cryoEM densities of one slab and the adjacent segment of the upper wall tier to visualize the slab/wall bridges (dashed: borderline of the two FU-h copies; arrows: contacts with neighboring slabs). Note that a bridge between slab and KLH1-b_1 is absent. (c) Interfaces between KLH1-h_1 and the wall FU segment a_1-b_1-c_1. In interface $h_1 \leftrightarrow c_1$, note the two opposing histidines (H1094 and H3259) that might influence hinge movements of the β-sandwich domain of KLH1-h_1 towards its core domain. (d) Interface $h_1 \leftrightarrow f_1$, involving the cupredoxin-like domain of KLH1-h_1 (green), and interface $h_2 \leftrightarrow f_1$, involving the core domain of KLH1-h_2 (gold). (e) Interface $h_2 \leftrightarrow a_1$ that connects KLH1-h_2 to the KLH1-a_1 copy in the next subunit dimer (asterisk). Note that other wall FUs are inaccessible by this slab.

electron micrographs of the purified hemocyanin subunit dimer from *H. pomatia*, *Acanthochiton fasicularis*, and other species.[34,39] Furthermore, it fully confirms the subunit dimer model that we previously proposed for *Nautilus* hemocyanin.[5] Due to structural reinforcement provided by its FU-h slab, the KLH1 subunit dimer appears to be more compact than the NpH subunit dimer, suggesting greater stability, which might be required for didecamer and multidecamer formation. This is also supported by successful attempts to stabilize and to purify the subunit dimer from gastropod and chiton hemocyanins, whereas such experiments failed with *Nautilus* and *Octopus*; apparently, their subunit dimer is a more transient dissociation/reassembly intermediate.[45,46] Nevertheless, our previous data on NpH[5] and the present results on KLH1 confirm the subunit dimer as an asymmetric homodimer and as the smallest repeating unit of the native molluscan hemocyanin molecule.

KLH1-h interfaces within the collar complex ($h_1 \leftrightarrow h_2$, $h_2 \leftrightarrow h_1$, and $g_1 \leftrightarrow h_1$)

Within each slab, the two KLH1-h copies are in an anti-parallel arrangement and connected *via* their cupredoxin-like domains that form interface $h_1 \leftrightarrow h_2$ with the opposing core domains (Fig. 8d). A 4-Å resolution of the cupredoxin-like domain allowed tracing of its C^{α} backbone, prediction of its secondary structure elements, and identification of the single disulphide bridge, but the orientation of the side chains remained completely unsolved.[9] Homology modeling with structurally comparable cupredoxins is critical because of low sequence identities; therefore, molecular details of interface $h_1 \leftrightarrow h_2$ remain concealed (see Fig. 8d), and whether the two FU-h copies within the slab are allosterically coupled is still open. The cupredoxin-like domain seems to lack a special active site;[9] therefore, it might solely function as a structural stabilizer of the slab.

Adjacent slabs are interconnected *via* a large contact of two core domains, termed interface $h_2 \leftrightarrow h_1$. It utilizes from both FUs the long helix α2 and the region comprising the short helix α5, although at different sections (see Fig. 8d and Table 3). In this interface, each FU delivers a dozen polar and charged residues for electrostatic bonding. Adjacent to interface $h_2 \leftrightarrow h_1$ lies the invariant PYWDW motif, which has been proposed to influence the oxygen-binding behavior.[5,6]

Apart from the narrow bridge representing the $g_2 \rightarrow h_2$ linker, the cryoEM structure shows no physical contact between KLH1-g_2 and KLH1-h_2, meaning that a "$g_2 \leftrightarrow h_2$ interface" is lacking; in contrast, a prominent interface exists between KLH1-g_1 and KLH1-h_1 (see Fig. 8b). The amino acids at this interface $g_1 \leftrightarrow h_1$ offer possibilities for salt bridges (e.g., between E3157 and R2763) and hydrophobic interactions (Fig. 9a). Interestingly, two opposing histidine residues (H2502 and H3128) are involved in this interface, but whether they provide an opportunity for allosteric signal transfer is not clear. It appears that interface $g_1 \leftrightarrow h_1$ is the major plug between arc and slab, whereas the folded $g \rightarrow h$ linkers might be too flexible to provide a precise arc/slab adjustment.

Interfaces between slab and wall ($h_1 \leftrightarrow a_1$, $h_1 \leftrightarrow c_1$, $h_1 \leftrightarrow f_1$, $h_2 \leftrightarrow f_1$, and $h_2 \leftrightarrow a_1$)

This study unravels, for the first time, the contact zones between the slab and the wall, and the amino acids involved in these interfaces. The two KLH1-h conformers differ completely in their orientation relative to the wall and, therefore, constitute different slab/wall bridges. As deduced from the bridges visible in the cryoEM structure and supported by the molecular model, KLH1-h_1 is attached to KLH1-a_1, KLH1-c_1, and KLH1-f_1 of the same subunit, and this KLH1-f_1 copy is also fixed to KLH1-h_2; the latter forms an additional bridge to the KLH1-a_1 copy of the next subunit (Fig. 9b). A direct contact between slab and KLH1-b_1 is excluded; from the molecular model, connections *via* residues in the $a \rightarrow b$ or $b \rightarrow c$ linker might be possible, but corresponding bridges in the cryoEM structure are lacking. All other FUs are far beyond reach, notably those in the central wall tier.

Interface $h_1 \leftrightarrow a_1$ is a single electrostatic contact between H2917 and E296 (Fig. 9c). Opportunities for allosteric signal transfer are not obvious at this interface, although E296 is part of helix α15 that connects core and β-sandwich domain. Hinge movements of the β-sandwich domain against the core domain are believed to modulate oxygen affinity because it brings a conserved F/Y residue of the β-sandwich close to the copper active site.[5,6,47] In the present case, however, the position of the glutamate E296 is inefficient for influencing hinge movement.

Interface $h_1 \leftrightarrow c_1$ shows several opportunities for electrostatic bonding and involves two basic residues (R3192 and K3197) in helix α15 of KLH1-h_1, next to the hinge (see Fig. 9c). Moreover, there are two opposing histidines (H3259 and H1094); their close contact might influence hinge movement because H3259 is part of the β-sandwich domain and occupies a promising position for chemomechanical impact (see Fig. 9c). It should be noted that in *Limulus polyphemus* (an arthropod) hemocyanin, a panel of histidine clusters that might be crucial in allosteric signal transfer, due to the capability of histidine to bind or release protons under *in vivo* conditions, has been identified.[3] How force transfer might proceed between H1094 and the active site of KLH1-c_1 is less clear, but this histidine is also localized in a comparatively flexible part of the FU (see Fig. 9c).

So far, connections between KLH1-h_1 and the wall are restricted to the core and β-sandwich domain. In interface $h_1 \leftrightarrow f_1$, the curpredoxin-like domain of KLH1-h_1 comes into play. KLH1-f_1 delivers several charged residues (K2405, E2449, and D2450), but their opponents in the cupredoxin-like domain of KLH1-h_1 remain uncertain (Fig. 9d). In interface $h_2 \leftrightarrow f_1$, glutamate E2408 in the $\beta 8 \rightarrow \beta 9$ loop of

KLH1-f_1 provides an opportunity for a strong salt bridge with lysine K2955 in the α3→ α4 loop of the core domain of KLH1-h_2. In interface $h_2 \leftrightarrow a_1$, the same core domain offers two glutamates (E2904 and E2922) and three histidines (H2903, H2991, and H2918), whereas KLH1-a_1 delivers a lysine (K148) in the exposed β5→α7 loop (Fig. 9e). In conclusion, the slab/wall interfaces are completely dominated by electrostatic bonding.

Putative N-glycosylation pattern of KLH1

KLH is widely used in research and clinics as a biological response modifier and tumor vaccine carrier,[2,10–12] and its glycan side chains might be of prime significance for these functions. Analysis of the oligosaccharides of whole KLH revealed a rich collection of N-linked glycans, partly comprising novel structural elements.[48] In addition, it has been demonstrated that KLH shares a cross-reacting epitope with glycoconjugates from the trematode parasite *Schistosoma mansoni*,[49,50] and it carries the Thomsen–Friedenreich antigen that plays a role in bladder cancer carcinoma.[51] Besides N-linkage, in isoform KLH2, the presence of O-linked glycans has also been reported,[52] but it is unclear whether such carbohydrates are also bound to KLH1. The amino acid sequence of KLH1 shows eight potential attachment sites for N-linked glycans (NXT/S; see Fig. 4). KLH1-c and KLH1-e lack such sites, whereas the other FU types show either one or two of them. Also in other molluscan hemocyanins, N-glycans are lacking in case of FU-c, but FU-e is usually glycosylated.[53]

Indeed, the 9-Å cryoEM structure shows conspicuous protrusions that are not explained by the molecular model. Two of the eight asparagines in question (N1264 in KLH1-d, and N2468 in KLH1-f) are buried deeply in the cryoEM structure and not associated with any protruding mass; apparently, they are inaccessible for sugar attachment. However, each of the other six asparagines (N387, N529, N1632, N2129, N2555, and N3262) is close to the protein surface and associated with a prominent empty mass. Four of these combinations are localized at the outer surface of the cylinder wall in KLH1-a, KLH1-b, KLH1-d, and KLH1-f, respectively (Fig. 10a–d). Another site (N2555) is part of KLH1-g and, consequently, a pair of sugar trees might point from each arc into the cylinder lumen (Fig. 10e). Of course, we cannot predict from the cryoEM structure the exact shape of these glycans, and we assume that they are indeed considerably larger because their flexibility might blur the density

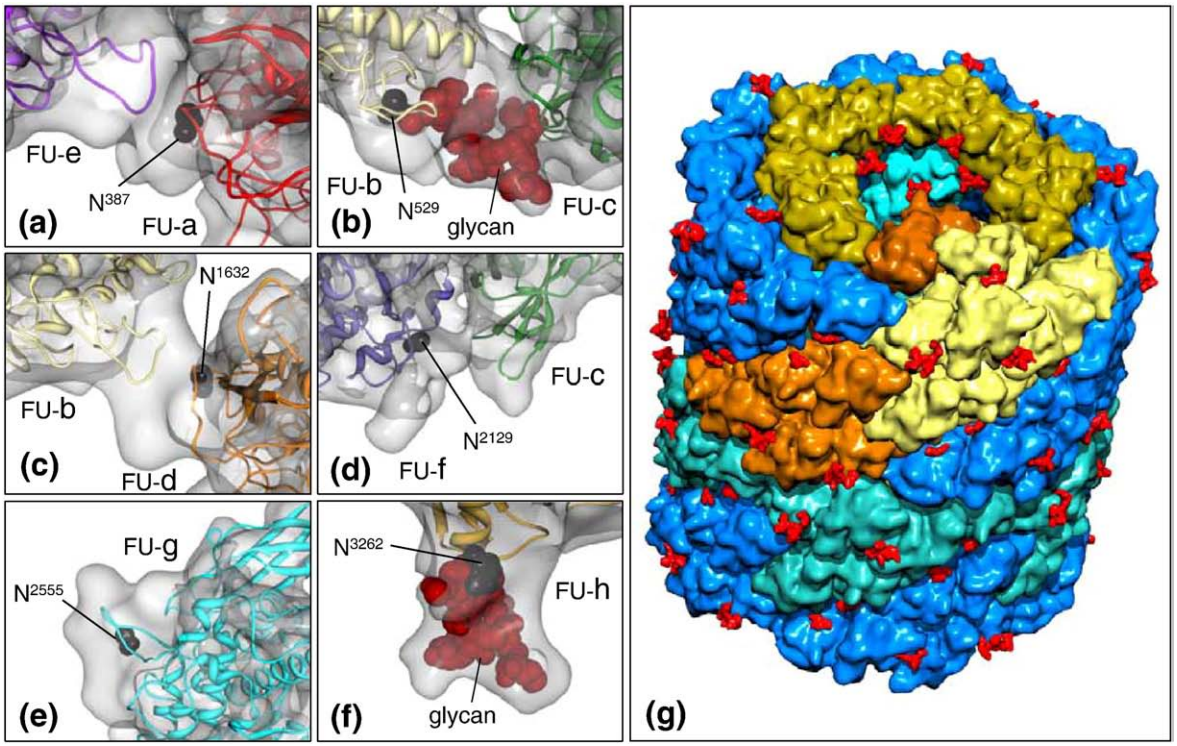

Fig. 10. Potential attachment sites for N-linked glycans. (a–f) Asparagines of NXT/S type, associated with protruding masses that are not occupied by the molecular model of the respective FU. In (b) and (f), short glycans (red) obtained from the crystal structure of OdH-g have been docked.[6] (g) Tilt view of a 9-Å volume calculated from the molecular model of the KLH1 didecamer, tagged with putative carbohydrate trees at 120 positions that correlate with empty masses in the cryoEM structure (red). In the lower decamer, subunits in conformation 1 (blue) are distinguished from those in conformation 2 (sea green). The constituents of one subunit dimer are highlighted by aberrant colors. Note the slab pentamer (gold) and the arc (cyan) in the cylinder lumen. Also note the five sugar trees on the rim of the wall that might sterically hinder multidecamer formation in KLH1.

map in these regions. Nevertheless, their principal localization and orientation are well confirmed because the FUs cannot be docked differently, and there are only these free spaces in the density map.

The most prominent protrusion of the density map is Y-shaped, is associated with N3262 in the β8→β9 loop of the β-sandwich domain of KLH1-h, and reaches in five copies from the slab pentamer deeply into the cylinder lumen (see Fig. 8a). This Y-shaped protrusion is likely to represent a diantennate carbohydrate tree. Indeed, a convincing docking of such a tree extracted from the crystal structure of OdH-g was achieved (Fig. 10f). This complex oligosaccharide of OdH-g is structurally related to several N-linked glycans found in a preparation of whole KLH.[48] It should be noted that the Y-shape is only evident in the five KLH1-h_2 copies of the slab pentamer, whereas in the alternating anti-parallel KLH1-h_1 copies, this protrusion contacts the rim of the cylinder wall where it merges with KLH1-c_1 (Fig. 10g).

The observation that each KLH1-h_1 copy appears to deposit a large sugar moiety on the rim of the cylinder wall, at the top of each KLH1-c_1 copy, is interesting because this specific N-glycan site is also present in HtH1-h, but missing in KLH2-h and HtH2-h. A characteristic feature of isoforms KLH2/HtH2 is that they form tridecamers and multidecamers *in vivo*, in contrast to KLH1/HtH1.[14,17,29] In a tridecamer, an additional decamer is attached with its open face to a "nucleating" didecamer. This assembly mode can continue on both sides of the didecamer to form long tubes. Of course, this capability might exclusively depend on specific amino acids in the FU repeat a-b-c-f, but it is also possible that the protruding N3262 sugars, notably those in h_1 position, considerably hamper such growth. For immunologists who apply KLH in immunological research or clinics, it may be worthwhile to see that isoform KLH1 of this giant invertebrate protein is almost evenly tagged by a total of 120 sugar trees (see Fig. 10g).

Concluding Remarks

In 1972, Mellema and Klug published their classical 3D model of a gastropod hemocyanin didecamer (based on superposition of EM images from negative staining).[33] In 1974, Siezen and van Bruggen presented EM images of solitary subunit dimers exhibiting "man-in-boat" and "parallelogram" views and, from such images, produced a more advanced model of the quaternary structure of the didecamer.[34] More than two decades later, several cryoEM-based 3D reconstructions of didecamers revealed the anti-parallel subunit arrangement in the wall, and defined the arcs and the slab pentamer; it then became clear that FU-g and FU-h form arcs and slabs, respectively.[7,20,25,31,35–37] From immunological and biochemical approaches and sequence analyses, it was also clear that the didecamer is a homopolymer.[8,17,26,42] However, at resolutions above 1 nm, docking of the individual FUs was impossible; therefore, an accurate molecular model of the didecamer could not be constructed. This deficit did prevent a definition of the subunit dimer, the twisted subunit pathway, and the molecular inter-FU interfaces. Notably, exactly how two identical copies of a single elongated subunit could start as an anti-parallel dimer to form the cylinder wall, and then continue in a parallel association to form the asymmetric collar complex, remained completely concealed.

Now, 35 years after Mellema and Klug, the present 9-Å cryoEM structure of KLH1 essentially solved all these questions because it allowed construction of a reliable molecular model of the entire didecamer (encompassing ~68,000 amino acids). A major surprise of this hybrid model was that, if the cryoEM structure of KLH1 is split into the separate decamers, followed by removal of the slab pentamer, the resulting fragment is almost indistinguishable from our previous 9-Å cryoEM structure of the native *Nautilus* hemocyanin decamer. Thus, it is justified to designate NpH as a molluscan hemocyanin prototype; in NpH and KLH1, the pathway of the subunit segment a-b-c-d-e-f forming the cylinder wall is conserved, and also the arcs composing FU-g are the same. With respect to allosterism, the surprising conservation in KLH1 of the 15 molecular inter-FU interfaces that also exist in NpH[5] suggests that basic functions are maintained similarly in NpH and the NpH-like core of KLH1. From the observed residue exchanges at these interfaces, at least some might be involved in the species-specific fine-tuning of the allosteric oxygen binding properties. From functional data, we would expect that, allosterically, the two decamers perform as independent units. However, possibilities for chemomechanical force transfer between the two associated decamers are indicated in the molecular model of KLH1. Considerations as to what extent the slab pentamer is included in allosteric interactions require greater structural detail of the cupredoxin-like domain, particularly concerning the side chains.

As already emphasized previously,[5] even with a detailed cryoEM structure and an accurate molecular model, tracing of the subunit pathway required careful strategy due to the length and presumed flexibility of the inter-FU linkers. In the present study, we also experienced this when tracing the two g→h linkers. In order to understand the intricate topology of the 160 FUs, definition of the two subunit conformers was prerequisite. Fortunately, the molecule cooperated with us in that the existing linker distances left only a single solution of the entire subunit pathway of both conformers. The observation that a single subunit type forms two different conformers during assembly is striking because the resulting asymmetric homodimer seems to contradict current models of homooligomer assembly. However, this is not a subunit as usual, but a "pearl chain" of eight covalently linked globular FUs. During assembly, the first six FUs

might form—in an anti-parallel manner—a stable cylinder wall. At the same time, the 10 FU-h copies might associate into a stable slab pentamer. This annulus has opportunities to form strong salt bridges with the edge of the cylinder wall, which might force the collar complex into its asymmetric position. Their long linkers allow the arc FUs to follow this movement and to complete and stabilize the structure.

In KLH1, the upper wall tier is connected to the slab pentamer, and the lower wall tier is fixed by the decamer/decamer interfaces. Lacking all this, the NpH molecule should be more flexible, which could cause local resolution limitations in 3D reconstruction. Indeed, compared to our recent 3D reconstruction of NpH, in the present cryoEM structure of KLH1, the inter-FU linkers, notably the exposed linker $f_2 \rightarrow g_2$, are more clearly defined; also the similarly exposed linker $g_2 \rightarrow h_2$ (absent in NpH) is completely represented in the density map. Moreover, the inter-FU bridges visible in the cryoEM structure of KLH1 are superior and can be fully explained by the molecular model. In most cases, these contacts indicate electrostatic bonding; often, there are clear opportunities for strong salt bridges, and binding sites for divalent cations are also present. This is fully compatible with the possibility of reversibly dissociating didecamers into subunits by controlling the pH and/or the levels of Ca^{2+} and Mg^{2+}.[1,21–24,26,27,29]

As deduced from molecular clock calculations, the eight molluscan hemocyanin FU types evolved ca 740 million years ago from a single proto-FU,[8,54] and, according to fossil records, gastropod and cephalopod hemocyanins separated at least 520 million years ago. After this primordial event, cephalopod hemocyanin lost FU-h and, thereby, the slab annulus. Moreover, it lost the capacity to form didecamers and multidecamers (an ancient feature, as deduced also from its occurrence in chitons and bivalves). Considering the clear demand for an efficient oxygen carrier in the highly active predatory cephalopods, we presume that this deconstruction is adaptive and somehow tailors the hemocyanin particle to the closed circulatory system with its high perfusion rates.

Vice versa, the more sluggish gastropods, with their low perfusion rates, almost invariantly use the didecamer/multidecamer hemocyanin version. The two distinct hemocyanin isoforms occurring in many species are apparently differentially expressed and play different functional roles in specific physiological situations.[17,26] In this context, it is still unclear why KLH2 forms multidecamers *in vivo* whereas KLH1 does not, although the present KLH1 study revealed the potential glycan side chain of FU-h_1 as a possible blocker. Another unexplored field constitutes oxygenation-dependent conformational changes in the quaternary structure.[55] Ultimately, to complete the molluscan hemocyanin picture, it might also be worthwhile to study the didecamers and multidecamers of KLH2 by cryoEM.

Materials and Methods

Animals and protein sample preparation

M. crenulata (Californian giant keyhole limpet) is a member of the Vetigastropoda. Animal maintenance, hemolymph collection, hemocyanin purification, and isolation of pure KLH1 from a KLH1/KLH2 mixture have been described;[19] this part was performed in our laboratory by K. Büchler. Sample grid preparation has been performed as in a preceding study,[7] with the exception that we tried to guarantee full oxygenation of the hemocyanin by using an atmosphere chamber containing 25% oxygen before cryoplunging and vitrification.

Cryo-electron microscopy

CryoEM was performed in our laboratory by F. Depoix using a Tecnai F20 transmission electron microscope equipped with a field emission electron source and operating at an accelerating voltage of 200 kV. Grids were cryotransferred using a Gatan cryoholder. Images were recorded as focal pairs under low-dose conditions using a Kodak SO-163 film at a nominal magnification of 50,000×, as previously described.[5]

Image digitalization, extraction, and correction

Selected micrographs were digitized using a PRIMESCAN rotating drum scanner (Heidelberger Druckmaschinen, Heidelberg, Germany) with a sampling size of 1.24 Å at the specimen level. The defocus value of each selected micrograph was determined by the program CTFFIND3.[56] Negatives with astigmatism and/or drift were discarded. From the digitized micrographs, 6100 molecular images of the KLH1 didecamer were extracted. Particle picking was performed semiautomatically within 512×512-pixel boxes using the module Boxer from the EMAN software package 1.7, which also allowed processing of the focal pair micrographs.[57] Single-particle images were then contrast transfer function (CTF) corrected (on the phases alone) by flipping phases of even rings using the module TRANSFER of the IMAGIC-5 software package (Image Science, Berlin, Germany).

Image processing and 3D reconstruction

All further steps of image processing were performed in the context of the IMAGIC-5 software system on a Linux cluster of 20×2 parallel Dual-Core AMD OPTERON processors (80), each associated with 4 Gb of memory. The CTF-corrected molecular images were bandpass-filtered (spatial frequencies below 1/35 Å and above 1/5 Å were suppressed) and normalized. Briefly, the process of 3D reconstruction consisted of iterative refinement of an initial 3D model against a set of CTF-phase-corrected images. Defocus of the selected particles varied from ~0.8 to ~4 μm (up to 75% of the selected particles had a defocus of ~0.8 to ~1.6 μm). As a starting model, we used the 11-Å cryoEM structure of HtH1,[7] downfiltered to a resolution of 30 Å, to avoid any reference bias. KLH1 and HtH1 are orthologous isoforms,[29,32] and at the quoted resolution level of 12 Å ($FSC_{3\sigma}$ criterion), their cryoEM structures are almost indistinguishable;[31,36] this would correspond to ca 14 Å if the present $FSC_{1/2\text{-bit}}$ criterion is applied.[58]

Projections of the initial reference covering the Euler angles at an interval of 2° were generated, taking the imposed D5 symmetry into account. After single-particle alignment, classification and angle assignment were performed *via* measurement of similarity between each single-particle image aligned to the corresponding reference projection from the 30-Å starting model. Then, class averages were used for 3D reconstruction by applying the exact filter backprojection method.[59] After this initial step, the use of an external model as reference was strictly avoided; instead, reprojections of the obtained KLH1 3D reconstructions were used as references for the subsequent iteration. After four iteration rounds, the 3D reconstruction achieved 10.5 Å resolution. The interval of generating references was changed to 1°, corresponding to a set of 4094 reference images. After five further refinement rounds, the final 9.1-Å reconstruction ($FSC_{1/2\text{-bit}}$ criterion) was achieved from 4762 high-quality images evenly distributed among 2271 classes. This data set contains 10% top views, 50% side views, and 40% tilted views ($25° < \beta\text{-angle} < 75°$). The balanced distribution of projections contributes to an isotropic resolution in the final 3D reconstruction. Some well-separated α-helices could be resolved after applying a high-pass filter. The final 3D reconstruction was then filtered to a nominal 9.1 Å, thus eliminating structural details beneath this cutoff.

Resolution determination

In this study, the resolution of the cryoEM structure is defined by the $FSC_{1/2\text{-bit}}$ criterion. The value obtained (~9 Å; see Fig. 1) conforms to the observed details in the density map, notably the shape of the individual FUs that allows high-quality fitting, the inter-FU bridges that are compatible with the molecular model, and protrusions associated with potential N-glycan sites. Furthermore, a 3D volume calculated from the molecular model of KLH1 and downfiltered to a nominal resolution of 9 Å shows a highly comparable level of detail (see Fig. 10g).

A resolution of 11.3 Å, as indicated here by the $FSC_{0.5}$ criterion (see Fig. 1), would not allow the definition of such details in the cryoEM structure; therefore, in our opinion, this approach is too pessimistic, particularly as conventional cross-correlation functions (CCFs) have been applied for the image alignment procedures. A parallel attempt to refine the present cryoEM structure by using, in the final iterations, mutual correlation functions instead of CCF (a procedure that underscores delicate details)[60] yielded virtually the same 3D volume (data not shown). However, with this approach, a nominal increase in the resolution (by ~2 Å) was obtained; in this case, the $FSC_{1/2\text{-bit}}$ criterion was clearly too optimistic, whereas the $FSC_{0.5}$ criterion yielded a reasonable value in the subnanometer range (~9 Å). After this experience, and using the CCF-derived 3D reconstruction, we decided to take our visual inspection of the cryoEM structure into account and to adhere to the $FSC_{1/2\text{-bit}}$ criterion.

Molecular modeling, rigid-body fitting, and visualization

Homology modeling of the individual FUs KLH1-a to KLH1-g, and of the corresponding fragment of KLH1-h, was performed by the MODELLER 9v4 software[61] using the crystal structures of OdH-g and *Rapana thomasiana* hemocyanin FU-e as templates.[6,62] Sequence alignment was performed with the ClustalW software package.[63] For each FU type, 10 different models were calculated. The final FU models were selected based on their stereochemical quality, which was determined using MolProbity.[64] The molecular models and the crystal structure of FU-h[9] were then semiautomatically rigid-body-fitted into their respective FU density using Chimera.[65] Linker peptides were modeled using the corresponding inter-FU linkers of our recent NpH molecular model as template;[5] this was considered to be justified from the present comparative analysis (see Results and Discussion). In some cases, loop refinement was performed by MODELLER 9v4 in order to increase the correlation with their corresponding densities. The two g→h linkers were also modeled by this method. In some interfaces, a few rotamers were corrected in Chimera for an improved fit with bridges in the density map; only rotamers with high probability values were accepted. For quality control of the various fitting stages, FU densities were extracted with the "color zone/split map" module of the Chimera software, and correlation coefficients were computed using the module COLACOR from the SITUS 2.2.1 software package.[66] Also, for visualization and analysis, and for preparation of the figures, we applied Chimera.

Our detailed interpretation of the cryoEM structure is based on high-quality rigid-body fitting of the molecular models of the individual FUs. This fitting is unique because the FUs cannot be docked differently, but how reliable are the models used? The available modeling and evaluation software enabled us to construct FU models of very high quality. The spatial orientations of their α-helices and β-strands are well confirmed because they are identical in the three existing crystal structures of single FU types.[6,9,62] Initially, the orientation of large exposed loops was uncertain, but refining their fitting into the cryoEM structure provided convincing opportunities for electrostatic bonding between FUs. Indeed, the entire range of inter-FU bridges visible in the cryoEM structure can be explained on this basis, which is compatible with results from dissociation–reassembly experiments and therefore strongly supports the present hybrid model. This approach even allowed, in certain loops, proposals for the orientation of some side chains, although their visibility is clearly beyond the resolution limit of 9 Å. We are well aware of this limitation of the present cryoEM structure of KLH1 and therefore avoided, for example, flexible fitting.

Database accession numbers

The cryoEM density map of *M. crenulata* hemocyanin isoform 1 (KLH1) has been deposited in the EMD database (EMBL-EBI) under accession number EMD-1569. The KLH1 sequence is available in the EMBL/Gene Bank database under accession number AJ698341/CAG28309.[32] The crystal structures of *O. dofleini* FU-g, *M. crenulata* FU-h, and *R. thomasiana* FU-e possess the RSCB Protein Data Bank ID codes 1JS8, 3EU2, and 1LNL.[6,9,62]

Acknowledgements

We thank Prof. Dr. J. Robin Harris (Institute of Zoology, Johannes Gutenberg University) for critical reading of the manuscript and language corrections. Moreover, we thank the *biosyn* company (Fellbach, Germany) for providing the animals, Dr. Kay Büchler (Institute of Zoology, Johannes Gutenberg

University) for purifying KLH1 from *M. crenulata* hemolymph, Dr. Frank Depoix (Institute of Zoology, Johannes Gutenberg University) for producing the cryoEM images, and Prof. Dr. Elmar Jaenicke (Institute of Molecular Biophysics, Johannes Gutenberg University) for details concerning the cupredoxin-like fold of FU-h. This work was supported by grants (to J.M.) from the Deutsche Forschungsgemeinschaft (GK 1043) and the Immunointervention Cluster of Excellence of Rheinland-Pfalz.

References

1. van Holde, K. E. & Miller, K. I. (1995). Hemocyanins. *Adv. Protein Chem.* **47**, 1–81.
2. Harris, J. R. & Markl, J. (1999). Keyhole limpet hemocyanin (KLH): a biomedical review. *Micron*, **30**, 597–623.
3. Martin, A., Depoix, F., Stohr, M., Meissner, U., Hagner-Holler, S., Hammouti, K. *et al.* (2007). *Limulus polyphemus* hemocyanin: 10 Å structure, sequence analysis, molecular modelling and rigid-body fitting reveal the interfaces between the eight hexamers. *J. Mol. Biol.* **366**, 1332–1350.
4. Decker, H., Hellmann, N., Jaenicke, E., Lieb, B., Meissner, U. & Markl, J. (2007). Minireview: recent progress in hemocyanin research. *Integr. Comp. Biol.* **47**, 631–644.
5. Gatsogiannis, C., Moeller, A., Depoix, F., Meissner, U. & Markl, J. (2007). *Nautilus pompilius* hemocyanin: 9 Å cryo-EM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units. *J. Mol. Biol.* **374**, 465–486.
6. Cuff, M. E., Miller, K. I., van Holde, K. E. & Hendrickson, W. A. (1998). Crystal structure of a functional unit from *Octopus* hemocyanin. *J. Mol. Biol.* **278**, 855–870.
7. Meissner, U., Gatsogiannis, C., Moeller, A., Depoix, F., Harris, J. & Markl, J. (2007). Comparative 11 Å structure of two molluscan hemocyanins from 3D cryo-electron microscopy. *Micron*, **38**, 754–765.
8. Lieb, B., Altenhein, B. & Markl, J. (2000). The sequence of a gastropod hemocyanin (HtH1 from *Haliotis tuberculata*). *J. Biol. Chem.* **275**, 5675–5681.
9. Barends, T., Büchler, K., Gatsogiannis, C., Schlichting, I., Markl, J., Decker, H. & Jaenicke, E. (2008). Hybrid methods in the structure determination of extremely large proteins. *Biophys. J.* **94**, 1571.
10. Harris, J. R. & Markl, J. (2000). Keyhole limpet hemocyanin: molecular structure of a potent marine immunoactivator. *Eur. Urol.* **37**, 25–31.
11. Kim, J. H., Lee, Y., Bae, Y. S., Kim, W. S., Kim, K., Im, H. Y. *et al.* (2007). Phase I/II study of immunotherapy using autologous tumor lysate-pulsed dendritic cells in patients with metastatic renal cell carcinoma. *Clin. Immunol.* **125**, 257–267.
12. Sabbatini, P. J., Ragupathi, G., Hood, C., Aghajanian, C. A., Juretzka, M., Iasonos, A. *et al.* (2007). Pilot study of a heptavalent vaccine–keyhole limpet hemocyanin conjugate plus QS21 in patients with epithelial ovarian, fallopian tube, or peritoneal cancer. *Clin. Cancer Res.* **13**, 4170–4177.
13. Senozan, N. M., Landrum, J., Bonaventura, J. & Bonaventura, C. (1981). Hemocyanin of the giant keyhole limpet, *Megathura crenulata*. In *Invertebrate Oxygen Binding Proteins* (Lamy, J. & Lamy, J., eds), pp. 703–717, Dekker, New York.
14. Markl, J., Savel-Niemann, A., Wegener-Strake, A., Süling, M., Schneider, A., Gebauer, W. & Harris, J. R. (1991). The role of two distinct subunit types in the architecture of keyhole limpet hemocyanin (KLH). *Naturwissenschaften*, **78**, 512–514.
15. Harris, J. R., Cejka, Z., Wegener-Strake, A., Gebauer, W. & Markl, J. (1992). Two-dimensional crystallization, transmission electron microscopy and image processing of keyhole limpet hemocyanin (KLH). *Micron Microsc. Acta*, **23**, 287–301.
16. Harris, J. R., Gebauer, W. & Markl, J. (1993). Immunoelectron microscopy of hemocyanin from the keyhole limpet (*Megathura crenulata*): a parallel subunit model. *J. Struct. Biol.* **111**, 96–104.
17. Gebauer, W., Harris, J. R., Heid, H., Süling, M., Hillenbrand, R., Söhngen, S. *et al.* (1994). Quaternary structure, subunits and domain patterns of two discrete forms of keyhole limpet hemocyanin: KLH1 and KLH2. *Zoology*, **98**, 51–68.
18. Harris, J. R., Chen, S. & Markl, J. (1994). Keyhole limpet hemocyanin (KLH): negative staining in the presence of trehalose and preliminary cryoelectron microscopy of unstained frozen-hydrated specimens. *Proc. Int. Conf. Electron. Mater.* **13**, 557–558.
19. Harris, J. R., Gebauer, W., Söhngen, S. M. & Markl, J. (1995). Keyhole limpet hemocyanin (KLH): purification of intact KLH1 through selective dissociation of KLH2. *Micron*, **26**, 201–212.
20. Dube, P., Harris, J. R., Orlova, E. V., Zemlin, F., van Heel, M. & Markl, J. (1995). Three-dimensional structure of keyhole limpet hemocyanin (KLH) by cryoelectron microscopy and angular reconstitution. *J. Struct. Biol.* **115**, 226–232.
21. Swerdlow, R. D., Ebert, R. F., Lee, P., Bonaventura, C. & Miller, K. I. (1996). Keyhole limpet hemocyanin: structural and functional characterization of two different subunits and multimers. *Comp. Biochem. Physiol. Part B*, **113**, 537–548.
22. Harris, J. R., Gebauer, W., Guderian, F. U. M. & Markl, J. (1997). Keyhole limpet hemocyanin (KLH): reassociation from Immucothel® followed by separation of KLH1 and KLH2. *Micron*, **28**, 31–41.
23. Harris, J. R., Gebauer, W., Söhngen, S. M., Nermut, M. V. & Markl, J. (1997). Keyhole limpet hemocyanin (KLH): characteristic reassociation properties of purified KLH1 and KLH2. *Micron*, **28**, 43–56.
24. Harris, J. R., Gebauer, W., Söhngen, S. M., Nermut, M. V. & Markl, J. (1997). Keyhole limpet hemocyanin (KLH): II. Characteristic reassociation properties of purified KLH1 and KLH2. *Micron*, **28**, 43–56.
25. Orlova, E. V., Dube, P., Harris, J. R., Beckman, E., Zemlin, F., Markl, J. & van Heel, M. (1997). Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 Å resolution by electron cryomicroscopy and angular reconstitution. *J. Mol. Biol.* **271**, 417–437.
26. Söhngen, S. M., Stahlmann, A., Harris, J. R., Müller, S. A., Engel, A. & Markl, J. (1997). Mass determination, subunit organization, and control of oligomerization states of keyhole limpet hemocyanin (KLH). *Eur. J. Biochem.* **248**, 602–614.
27. Harris, J. R., Gebauer, W., Adrian, M. & Markl, J. (1998). Keyhole limpet hemocyanin (KLH): slow *in vitro* reassociation of KLH1 and KLH2 from Immucothel. *Micron*, **5**, 329–339.
28. Gebauer, W., Harris, J. R., Geisthardt, G. & Markl, J. (1999). Keyhole limpet hemocyanin type 2 (KLH2): detection and immunolocalization of a labile functional unit h. *J. Struct. Biol.* **128**, 280–286.

29. Harris, J. R., Scheffler, D., Gebauer, W., Lehnert, R. & Markl, J. (2000). *Haliotis tuberculata* hemocyanin (HtH): analysis of oligomeric stability of HtH1 and HtH2 and comparison with keyhole limpet hemocyanin KLH1 and KLH2. *Micron*, **31**, 613–622.
30. Gebauer, W., Harris, J. R. & Markl, J. (2002). Topology of the 10 subunits within the decamer of KLH, the hemocyanin of the marine gastropod *Megathura crenulata*. *J. Struct. Biol.* **139**, 153–159.
31. Mouche, F., Zhu, Y. X., Pulokas, J., Potter, C. S. & Carragher, B. (2003). Automated three-dimensional reconstruction of keyhole limpet hemocyanin type 1. *J. Struct. Biol.* **144**, 301–312.
32. Lieb, B. & Markl, J. (2004). Evolution of molluscan hemocyanins as deduced from DNA sequencing. *Micron*, **35**, 117–119.
33. Mellema, J. E. & Klug, A. (1972). Quaternary structure of gastropod haemocyanin. *Nature*, **239**, 146–150.
34. Siezen, R. J. & van Bruggen, E. F. (1974). Structure and properties of hemocyanins: XII. Electron microscopy of dissociation products of *Helix pomatia* alpha-hemocyanin: quaternary structure. *J. Mol. Biol.* **90**, 77–89.
35. Lambert, O., Boisset, N., Taveau, J. -C., Preaux, G. & Lamy, J. N. (1995). Three-dimensional reconstruction of the α_D and β_C-hemocyanins of *Helix pomatia* from frozen-hydrated specimens. *J. Mol. Biol.* **248**, 431–448.
36. Meissner, U., Dube, P., Harris, J. R., Stark, H. & 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.
37. Cheng, K., Koeck, P. J., Elmlund, H., Idakieva, K., Parvanova, K., Schwarz, H. *et al.* (2006). *Rapana thomasiana* hemocyanin (RtH): comparison of the two isoforms, RtH1 and RtH2, at 19 Å and 16 Å resolution. *Micron*, **37**, 566–576.
38. Bergmann, S., Lieb, B., Ruth, P. & Markl, J. (2006). The hemocyanin from a living fossil, the cephalopod *Nautilus pompilius*: protein structure, gene organization, and evolution. *J. Mol. Evol.* **62**, 362–374.
39. Harris, J. R., Meissner, U., Gebauer, W. & Markl, J. (2004). 3D reconstruction of the hemocyanin subunit dimer from the chiton *Acanthochiton fascicularis*. *Micron*, **35**, 23–26.
40. Siezen, R. J. & van Driel, R. (1974). Structure and properties of hemocyanins: XIII. Dissociation of *Helix pomatia* α-hemocyanin at alkaline pH. *J. Mol. Biol.* **90**, 91–102.
41. Siezen, R. J. (1974). Structure and properties of hemocyanins: XIV. Reassociation of *Helix pomatia* α-hemocyanin. *J. Mol. Biol.* **90**, 103–113.
42. Altenhein, B., Markl, J. & Lieb, B. (2002). Gene structure and hemocyanin isoform HtH2 from the mollusc *Haliotis tuberculata* indicate early and late intron hot spots. *Gene*, **301**, 53–60.
43. Herskovits, T. T., Cousins, C. J. & Hamilton, M. G. (1991). The stabilizing influence of divalent ions and Na^+ on the di-decameric structure of *Yoldia limatula* hemocyanin. *Biochim. Biophys. Acta*, **1076**, 71–78.
44. Connelli, P. R., Gill, S. J., Miller, K. I., Zhou, G. & van Holde, K. E. (1989). Identical linkage and cooperativity of oxygen and carbon monoxide binding to *Octopus dofleini* hemocyanin. *Biochemistry*, **28**, 1835–1843.
45. Bonaventura, C., Bonaventura, J., Miller, K. I. & van Holde, K. E. (1981). Hemocyanin of the chambered *Nautilus*: structure–function relationships. *Arch. Biochem. Biophys.* **211**, 589–598.
46. van Holde, K. E. & Miller, K. I. (1985). Association–dissociation equilibria of *Octopus* hemocyanin. *Biochemistry*, **24**, 4577–4582.
47. Miller, K. I., Cuff, M. E., Lang, W. F., Varga-Weisz, P., Field, K. G. & van Holde, K. E. (1998). Sequence of the *Octopus dofleini* hemocyanin subunit: structural and evolutionary implications. *J. Mol. Biol.* **278**, 827–842.
48. Kurokawa, T., Wuhrer, M., Lochnit, G., Geyer, H., Markl, J. & Geyer, R. (2002). Hemocyanin from the keyhole limpet *Megathura crenulata* (KLH) carries a novel type of N-glycans with Gal(β1→6)Man-motifs. *Eur. J. Biochem.* **269**, 5459–5473.
49. Markl, J., Nour el Din, M., Winter-Simanowski, S. & Simanowski, U. A. (1991). Specific IgG-activity of sera from Egyptian schistosomiasis patients to keyhole limpet hemocyanin (KLH). *Naturwissenschaften*, **78**, 30–31.
50. Geyer, H., Wuhrer, M., Resemann, A. & Geyer, R. (2005). Identification and characterization of keyhole limpet hemocyanin N-glycans mediating cross-reactivity with *Schistosoma mansoni*. *J. Biol. Chem.* **280**, 40731–40748.
51. Wirguin, I., Suturkova-Milosević, L., Briani, C. & Latov, N. (1995). Keyhole limpet hemocyanin contains Gal(beta 1–3)-GalNAc determinants that are cross-reactive with the T antigen. *Cancer Immunol. Immunother.* **40**, 307–310.
52. Stoeva, S., Schütz, J., Gebauer, W., Hundsdörfer, T., Manz, C., Markl, J. & Voelter, W. (1999). Primary structure and unusual carbohydrate moiety of functional unit 2-c of keyhole limpet hemocyanin (KLH). *Biochim. Biophys. Acta*, **1435**, 94–109.
53. Lieb, B., Boisguerin, V., Gebauer, W. & Markl, J. (2004). cDNA sequence, protein structure, and evolution of the single hemocyanin from *Aplysia californica*, an opisthobranch gastropod. *J. Mol. Evol.* **59**, 1–10.
54. Lieb, B., Altenhein, B., Markl, J., Vincent, A., van Olden, E., van Holde, K. E. & Miller, K. I. (2001). Structures of two molluscan hemocyanin genes: significance for gene evolution. *Proc. Natl Acad. Sci.* **98**, *4546–4551.*
55. Hartmann, H., Bongers, A. & Decker, H. (2001). Small-angle neutron scattering reveals an oxygen-dependent conformational change of the immunogen keyhole limpet hemocyanin type 1 (KLH1). *Eur. Biophys. J.* **30**, 471–475.
56. Mindell, J. A. & Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347.
57. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97.
58. van Heel, M. & Schatz, M. (2005). Fourier shell correlation threshold criteria. *J. Struct. Biol.* **151**, 250–262.
59. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. (1996). A new generation of IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24.
60. van Heel, M., Schatz, M. & Orlova, E. (1992). Correlation functions revisited. *Ultramicroscopy*, **46**, 307–316.
61. Marti-Renom, M. A., Stuart, A., Fiser, A., Sánchez, R., Melo, F. & Sali, A. (2000). A comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325.
62. Perbandt, M., Guthohrlein, E. W., Rypniewski, W., Idakieva, K., Stoeva, S., Voelter, W. *et al.* (2003). The structure of a functional unit from the wall of a

gastropod hemocyanin offers a possible mechanism for cooperativity. *Biochemistry*, **42**, 6341–6346.
63. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882.
64. Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2004). MolProbity: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619.
65. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
66. Chacón, P. & Wriggers, W. (2002). Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.* **317**, 375–384.